# Rainbow Range Finder and its Implementation at the CVL[1]

Vladimír Smutný, Tomáš Pajdla
Computer Vision Laboratory
Czech Technical University
CZ-121 35 Praha 2, Karlovo nám. 13
smutny@vision.felk.cvut.cz

Research report No. K335-1996-K335-1996-130

March 5, 1996

Czech Technical University
Faculty of Electrical Engineering
Department of Control
Computer Vision Laboratory

[1]This research was supported by the Grant Agency of the Czech Republic, grants 102/93/0954, 102/95/1378 and European Union grant Copernicus No. 1068 RECCAD.

**Abstract**

This report describes the implementation of a 3-D scanner - Rainbow Range Finder. The sensor builds on the triangulation and active structured illumination using a special color projector. It delivers 3-D coordinates of points in the volume of size 20 cm $\times$ 15 cm $\times$ 15 cm with the precision about 3 mm. The depth is measured for all pixels in image.

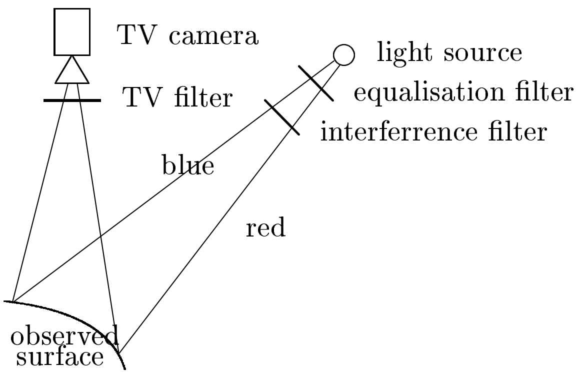



Figure 1: Basic principle of RRF and its components.

# 1 Introduction

Depth maps are needed for many tasks in science and industry (shape of human body in medicine, ergonomics, movie industry, advertisement, obstacle avoidance in robotics,...). The range finders for depth maps acquisition are based on various principles. See [1, 2] for the overview of the whole gamut of methods.

This text documents the implementation of the Rainbow Range Finder (RRF) at the Computer Vision Laboratory. The text is organized as follows. In Section 2, the principles of the sensor are explained. Constraints on the scene are described in Section 3. Concise description of the system is given in Section 4.

# 2 The principle of the sensor

Basic principle of a structured light 3-D sensor builds on the concept of stereo vision. Two perspective cameras placed at different locations in space look at the same 3-D point. Each camera therefore provides a light ray in the space that projects the 3-D point onto the retinas by connecting the point, the centers of projection, and the point in the retina. The intersection of two rays then readily gives the desired 3-D coordinates of the point in space. Computing the coordinates 3-D points by intersecting the rays passing through the point's image projections is usually called triangulation. Bottleneck of this approach consists in automatic extraction of corresponding projections from images which turns out to be rather difficult problem as it is not easy to predict how the projections are related without knowing where the point is [5].

The active vision approach eases the correspondence problem. Active illumination methods projects various patterns [1] onto the measured surface and the points of the patterns recognized in the image are then used for triangulation.

In the case of Rainbow Range Finder [7, 8, 9] one camera is replaced by a source of a color light rays and a their color observed by the color camera allows unique identification of the rays from the image of the illuminated scene (see Fig. 1).

The major advantages of this equipment comparing to other structured light range finders are the following:

- Short data acquisition time. The time required for one measurement is equal to grabbing of one image i.e. for color TV camera 40ms. This supports the use of this device in the applications measuring moving objects. One possible application is for medicine - capturing the shape of a human body.

- The information is dense. The 3-D coordinates are known for all pixels in the image excluding shadows.

### 2.1 Structured light projector

The Rainbow Range Finder uses for coding the pattern illumination monochromatic in each point but with varying wavelength. The wavelength is then measured by the color camera in each pixel.

The source of light (further called the projector) produces a color illumination with the following properties (see Fig. 2). :

1. The illumination rays are formed according the pinhole camera model, i.e. they pass through one common point — object focal point of the projector (further only focal point).

2. Each ray is monochromatic, i.e. only one wavelength has nonzero power.

3. All rays with the same wavelength form a plane passing through the focal point.

4. All monochromatic planes form a pencil of planes. These planes pass through one intersection line, which again passes through the focal point.

5. The wavelength of monochromatic planes is a continuous and monotonic function of the angle between planes and a line perpendicular to the intersection line.

6. The intersection line is not passing through the camera focal point.

Note that the role of the projector focal point $F_p$ is replaced by the intersection line and $F_p$ is not (and cannot be) found during the calibration procedure. However, the focal point $F_p$ is important from practical point of view because rays really pass through it. Therefore the position of $F_p$ on the intersection line determines scene selfocclusions for illumination (dark places on the measured surface).

Let us define the world coordinate system as an Euclidean coordinate system in which we want to determine the surface point coordinates. To have a calibrated camera means to know for the given image point $U$ with coordinates $(u, v)$ the world coordinates of the ray (straight line) which projects into $U$. The calibration of a camera can be obtained by the procedure described in [6]. To have a calibrated projector means to know for each monochromatic plane $\pi_\alpha$ its world coordinates. The calibration procedure will be briefly described in section 5.2. To have a camera with color calibration means to be able to determine plane $\pi_\alpha$ illuminating the projected surface point $X$ from the color information measured in one pixel.

Having camera and projector calibrated one can compute the coordinates of the surface point X in the world coordinate system (see Fig. 3). The mathematical procedure for the computation of 3-D coordinates from known projection of the point X into image and known projector plane $\pi_\alpha$ passing through (illuminating) the point X is called triangulation.

A new feature of our method is the determination of a unique function of a illuminating ray wavelength from measured pixel color. To encode planes by their color it is useful to have an asignment between a color $c$ detected on the image and the angle $\alpha$ of its plane, $\alpha = G(c)$. One is looking for the function $c = f(\bar{s})$ where $\bar{s}$ denotes the measured vector in each pixel of the camera and $c$ stands for scalar function of the pixel color. Note that both functions $\alpha = g(\lambda)$ and $\lambda = h(c)$ are monotonic. The finding of the function $\alpha = g(h(c))$ is a subject of color calibration as mentioned above.

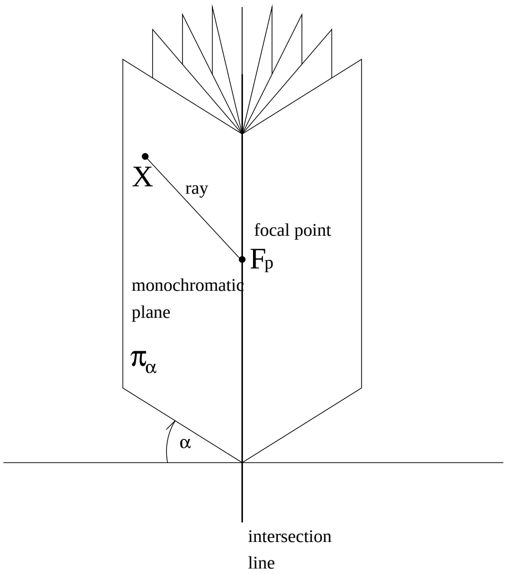



Figure 2: Illustration of the projector. $F_p$ is the object focal point of the projector, through which pass all rays. Rays with the same wavelength form a plane $\pi_\alpha$. All planes intersect in an intersection line passing through focal point $F_p$. The planes are parametrized by the angle $\alpha$ between the plane $\pi_\alpha$ and one common line perpendicular to the intersection line.

Let us look what are the basic requirements for such a transformation function. The function should be independent of a brightness of the light falling onto camera pixel because the brightness depends on the distance of the measured surface from the projector, on the intensity of the illuminator, on the albedo of the surface, and on camera, projector, and scene properties.

The measured vector $\bar{s}$ consists from different components for various combinations of cameras and frame grabbers. Some cameras give RGB signals, some PAL (NTSC) signal, some Y/C signal. Frame grabbers produce R, G, B components or Y, U, V components. One can think about the specialized hardware giving hue for input PAL signal. Generally, it holds that signals from various systems can be converted. The problem is that conversions are often done with roundoff error that produces distortion in the output signal and finally in 3-D coordinates. It is better to accommodate the function $f$ to the system of color camera—frame grabber what eliminates an unnecessary roundoff error.

The frame grabber used in our system can interface Y/C signal from the camera and digitize it into the YUV system. Y component reflects the brightness of the light falling into pixel. U is a component reflecting the difference between the B channel and the brightness Y, V is the difference between the R channel and the brightness. The function of four quadrant arctan of U and -V gives the monotonic function of the wavelength for monochromatic light. The function does not depend on the brightness even though U and V do. Note that for particular HW representation of U and V components the formula cannot be directly applied. The actual transformation of U and V before above described calculation should be found in the documentation of the frame grabber.

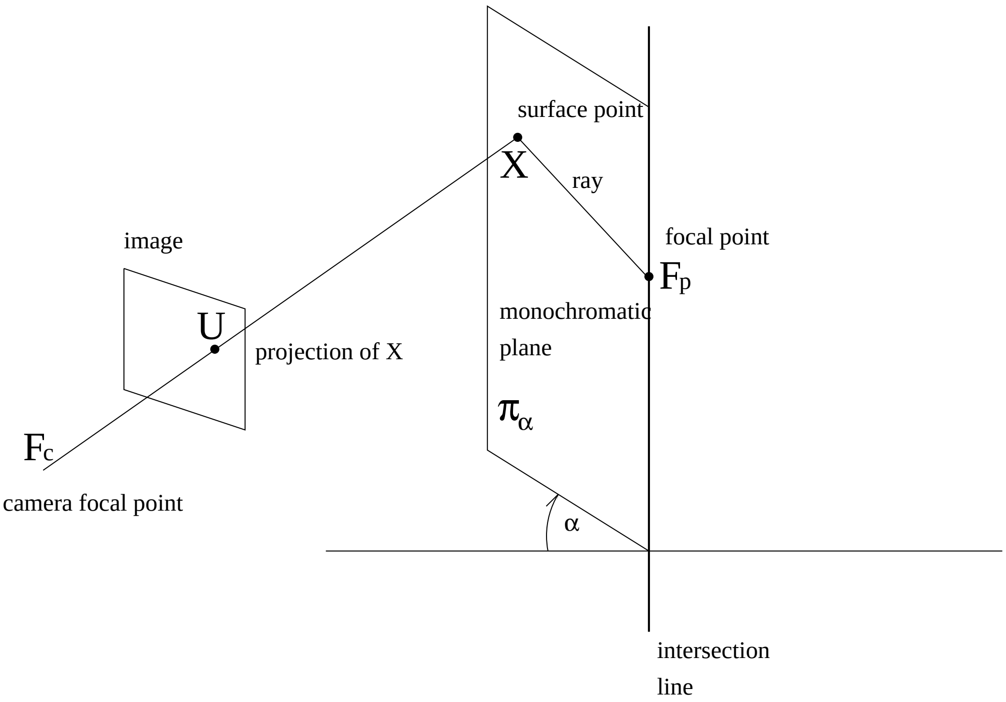



Figure 3: Triangulation of the surface point. Image point U defines the ray passing through surface point X as well as camera focal point $F_c$. Decoded color of the point U gives the monochromatic plane $\pi_\alpha$. The intersection of the camera ray and monochromatic plane gives the position of surface point $X$ in the world coordinate system.

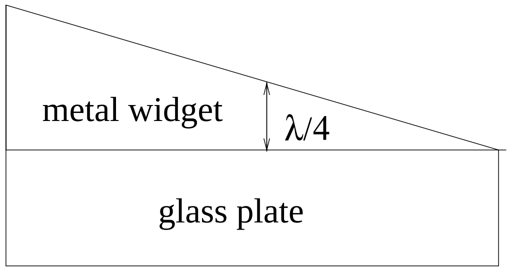



Figure 4: Cut through the interference filter. Interference filter consists of the base glass plate and a metal widget vapoured on it. The light with wavelength $\lambda$ passes only through place, where the thickness of the widget is $\lambda/4$.

### 2.2 Color illumination generation

The projector described above can be build from the light projector where light passing through interference filter (see Fig. 4) is projected to the scene. The interference filter is transparent in each place to only a narrow band of wavelengths and thus producing almost monochromatic illumination. The locus of the points with the constant wavelength is a straight line for each wavelength. All such straight lines are parallel.

### 2.3 Topology

The RRF captures whole scene from one color image. The reconstructed surface consists of discrete points each measured in one image point. The topology of reconstructed surface is the same as the topology of the original image upto the depth discontinuities.

Neighbouring pixels in image produce neighbouring points in the space unless they are very far in depth. To set the threshold on the depth discontinuity and eliminate the false neighbours is required from the user. The sensor itself does not have any additional information which is not already contained in the output data.

Therefore the topology of the output data is set to the full topology of the image with removed outliers and invalid data.

## 3 Constraints on the measured scene

The constraints on the scene for capturing its range image can be divided into two groups. First group (hard constraints) cover the principal physical constraints which cannot be overcome. Second group (soft constraints) describes the constraints of the technical nature which lead to the gradual corruptions of the range image. The technical advances in camera, frame grabber, and optics design as well as algorithm improvements can at least partially remove this constraints.

Physical constraints are:

1. The scene cannot be completely black for any wavelength used.

2. The scene should be reasonably opaque.

3. The measured part of the scene have restricted shape. Only the parts of the surface visible by camera and illuminated by projector are measured.

4. Scene should not be illuminated by any other source of light than projector described[1].

5. The surface can have any color but cannot be fluorescent or phosphorescent, i.e. cannot reflect the light of another wavelength than is the wavelength of the light falling onto it.

Regarding item 5 note that for monochromatic illumination the reflected light consist always from the same wavelength as is the wavelength of the falling light. This means that the red blob illuminated by monochromatic green illumination appears either black (restricted by the item 1) or green.

Technical constraints are:

1. The range cannot be calculated for any image point where any component of $\bar{s}$ is saturated.

2. The range accuracy is low for poorly illuminated image points. This constraint applies particularly when the surface is, for example, red and it is illuminated by blue light. The intensity of the reflected light is then small and the angle determined by the small noisy integer number is ill conditioned.

3. Interreflections in the scene cause wrong estimates of surface shape. This can be overcome but it is not implemented at the device.

# 4 Sensor Description

The setup of the sensor (see Fig. 5) consists from of color camera and the slide projector with the interference filter. They are both mounted on an optical bench what allows to change their relative position for measuring objects with different sizes.

The orientation of the camera and projector is adjusted by hand so that the intersection line of the projector is almost parallel to the $y$ axis, and projections of the $x$ and $y$ axes to the image are almost parallel to rows and columns respectively. The angle between the projector and camera optical axes is about 30 degrees.

The interference filter is a a metal wedge vapoured onto the glass plate with the size of a standard slide ($50 \times 50$ mm).

## 4.1 Color estimation

The color image of size $576 \times 768$ is captured by the camera but U and V components are measured only for pairs of pixels on a row. Therefore the resolution of the final image is at maximum $576 \times 388$ pixels.

Each pixel is processed independently. The arctan of transformed U and V is calculated ($c = \arctan(-U/V)$). The coordinates of appropriate monochromatic plane are found through a lookup table. The coordinates of pixels in image are used for the triangulation.

The value of brightness is used to express belief in the measured value. The saturation in any value is used as invalidation of the output data.

[1]This constraints can be overcome by different approach to the color decoding but this is not the case of our equipment.

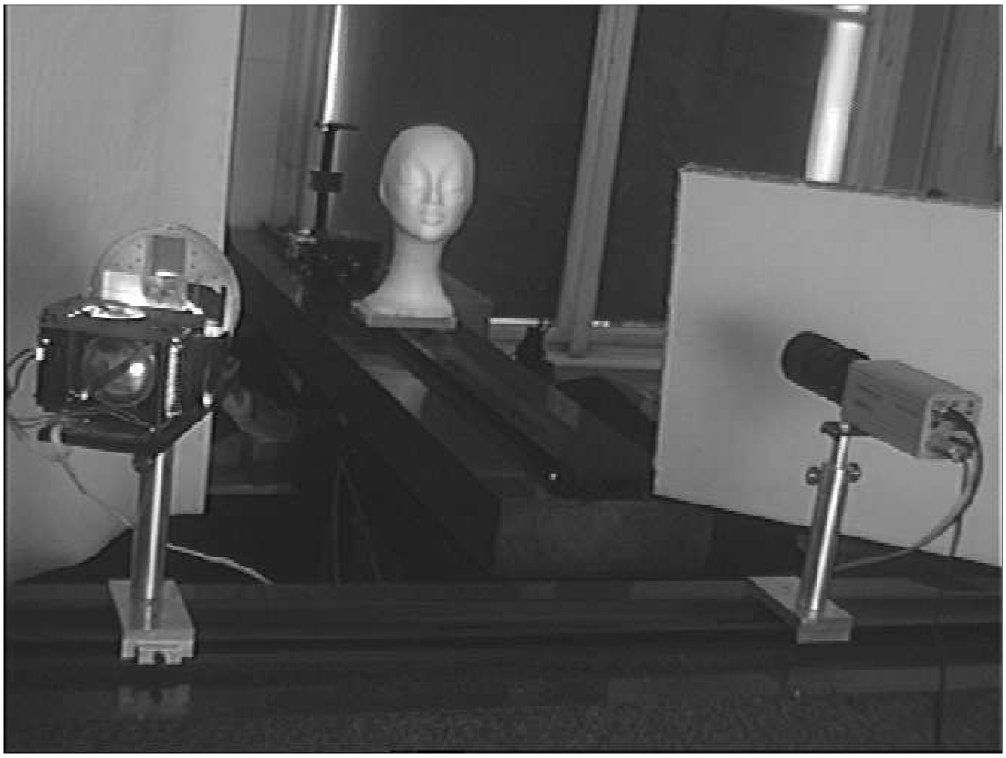


Figure 5: Setup of the sensor. Color camera is on the right, structured light projector on the left side of the image. The object (sculpture of the human head) is in the middle of the image.

# 5 Sensor calibration

Geometrical relation between the camera, the projector, and the world coordinate system is established through the explicit calibration of the camera-projector. The calibration uses known scene, i.e., the set of points coordinates of which are a priory known, to obtain necessary parameters for building the mapping $(u,\ v,\ c) \rightarrow (x,\ y,\ z)$, where $u$ and $v$ stand for image coordinates of the evaluated pixel and $c$ parametrizes the position of the monochromatic plane from the projector. The coordinates $x$, $y$, and $z$ are then 3-D coordinates of the reconstructed point in the world Euclidean coordinate system.

The calibration of the sensor consists of three steps. First, the camera is calibrated with respect to the world coordinate system. Then, the projector is calibrated with respect to the world coordinate system. Finally, the LUT for the relationship between measured color $c$ and position of the plane in the space is build.

## 5.1 Camera calibration

The pinhole camera model is used. Three calibration panels with known parameters and automatically detectable reference points are sequentially placed in front of the camera. The whole procedure is described in [10].

## 5.2 Projector calibration

The projector is calibrated using a pinhole camera model. Three panels with known positions are captured and colors are computed. The points with the same color should lie in a monochromatic plane. Therefore a Least Squares method for finding the world coordinates of the plane is used. Because this method does not proved stable results we added to the data an approximate position of the projector focal point measured by meter. This is used for making estimate of the parameters of a pencil of monochromatic planes more consistent. The actual sensitivity of the calibration to the focal point position estimate is small compared to the positive contribution in consistency of the calibrated planes. The more elaborate method will be added soon.

The LUT from decoded color to the parameters of decoded plane is then build from the captured images. Figure 6 shows the RGB components, YUV components and decoded color c for varying wavelength.

# 6 Errors

The nature and origins of RRF specifical errors in recontructed surface are described in detail in [7, 8]. Here we want to briefly mention the most important problem. It is the uncertainty in color measurement which causes the uncertainty in monochromatic plane determination and finally the uncertainty 3-D reconstruction.

The color uncertainty is worst in the flat parts of the function inverted function $c = h^{-1}(\lambda)$ shown on Fig. 6 (c). This causes that some parts of the reconstructed scene are much more noisy than others. The origin of this error can be found in noise of the camera.

The calibration of the projector is dificult because the measured data are noisy. The monochromatic planes calibrated actualy do not form a pencil and even in the measured space the planes are not well ordered. The addition of the projector focal point to the data produce a smooth monotonic pencil of calibrated planes as it is shown on Fig. 7(d).

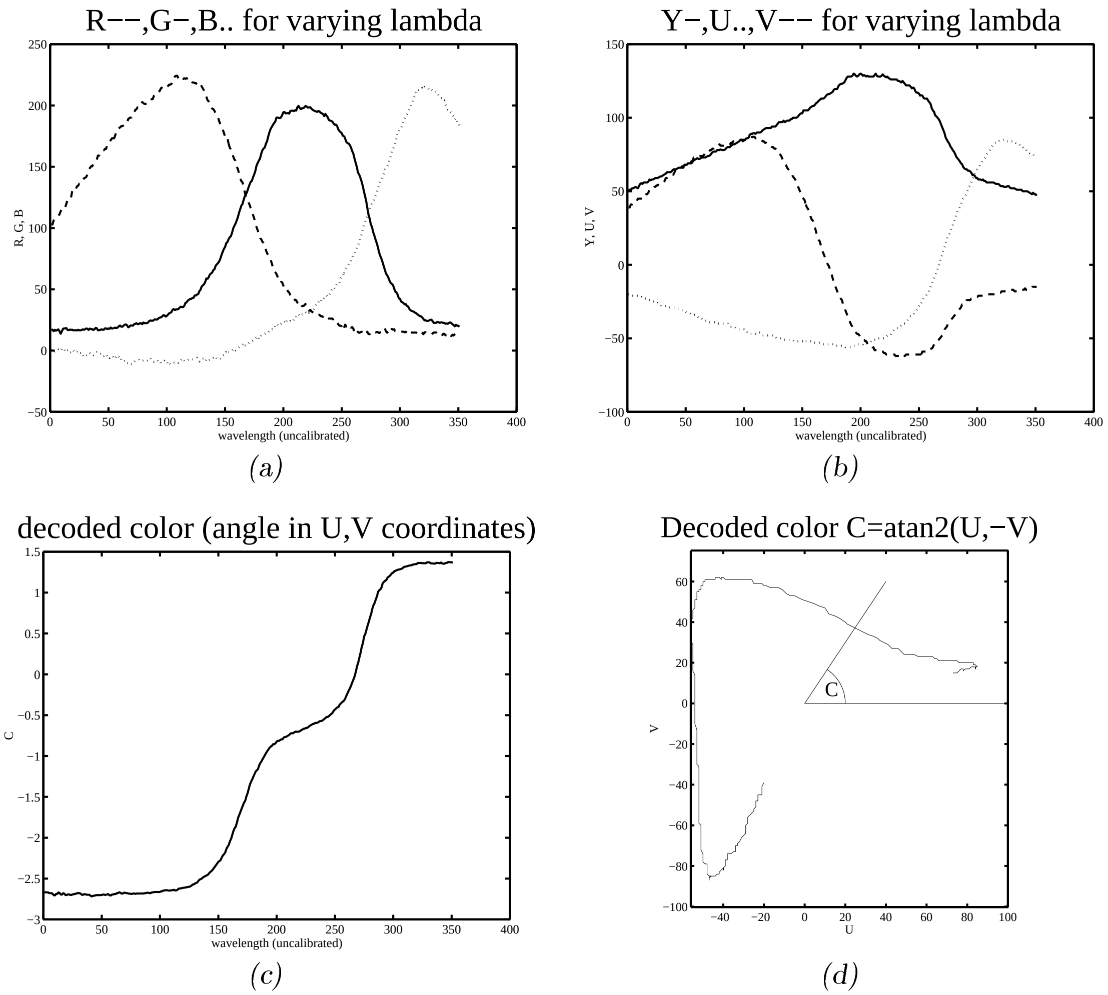



Figure 6: The example of RGB (a) and YUV (b) components as function of wavelength, (c) shows decoded color. Horizontal axis (wavelength) is not calibrated, red part of the spectra left, blue right. The figure (d) shows the locus of the $U$, $V$ values measured and how the color $C$ is determined from them as an angle.

Fig. 7 shows the statistical results of the calibration procedure. The procedure measures three calibration plane panels. The data shows the errors of calibrating points on these panels. The data can be interpreted as a error measurement on testing object consisting from the plane perpendicular to $z$-axis. One can see that the standard deviation is about 3 mm.

# 7 Technical Data

| | |
|---|---|
| Camera | color 3 chip CCD Hitachi HV-C20, CCIR,<br>$582 \times 752$ pixels, Y/C output |
| Lens | Computar, zoom 12-75mm |
| Projector | Slide projector with interference filter, f=80mm |
| Frame grabber | miroMovie, miro Computer Products,<br>$576 \times 768$ pixels, YUV in 4:2:2 format |
| Resolution x/y/z | 0.5 mm / 0.5 mm / 3 mm |

## 7.1 Format of 3-D data

3-D data are stored to an ASCII file arranged to lines delimited by CR-LF characters. Each line of the file contains data describing one point. Data are stored columnwise, i.e. one scanline follows the other. The format of each line, i.e. point, is the following:

x y z north nbr. west nbr. south nbr. east nbr. certainty

The first three items hold the coordinates, next fourhold links to the four neighbours as the indices of the rows in the file, 0 stands for missing neighbour. The last item provides the measure of the certainty of point detection.

The alternative representation of the output data relates to the fact that the output of the RRF is a dense range image. This means the output data arrangement corresponds to the input image. The output ASCII file then consist of four matrices. The matrices correspond to the x, y, z, and certainty. The lines of matrices are separated by CR-LF characters. Matrices are stored in column. The zeros in certainty matrix stand for invalid measurements.

# 8 Experimental results

# 9 Conclusion

The prototype of Rainbow Range Finder provides data at high resolution and low precision when compared to best commercially available. However, a number of improvements are still thought and currently developed.

# References

[1] R. Jarvis. Range sensing for computer vision. In A.K. Jain and P.J. Flyn, editors, *Three Dimensional Object Recognition Systems*, volume 1 of *Advances in Image Communication*, pages 17–56. Elsvier Science Publishers B.V., 1993.

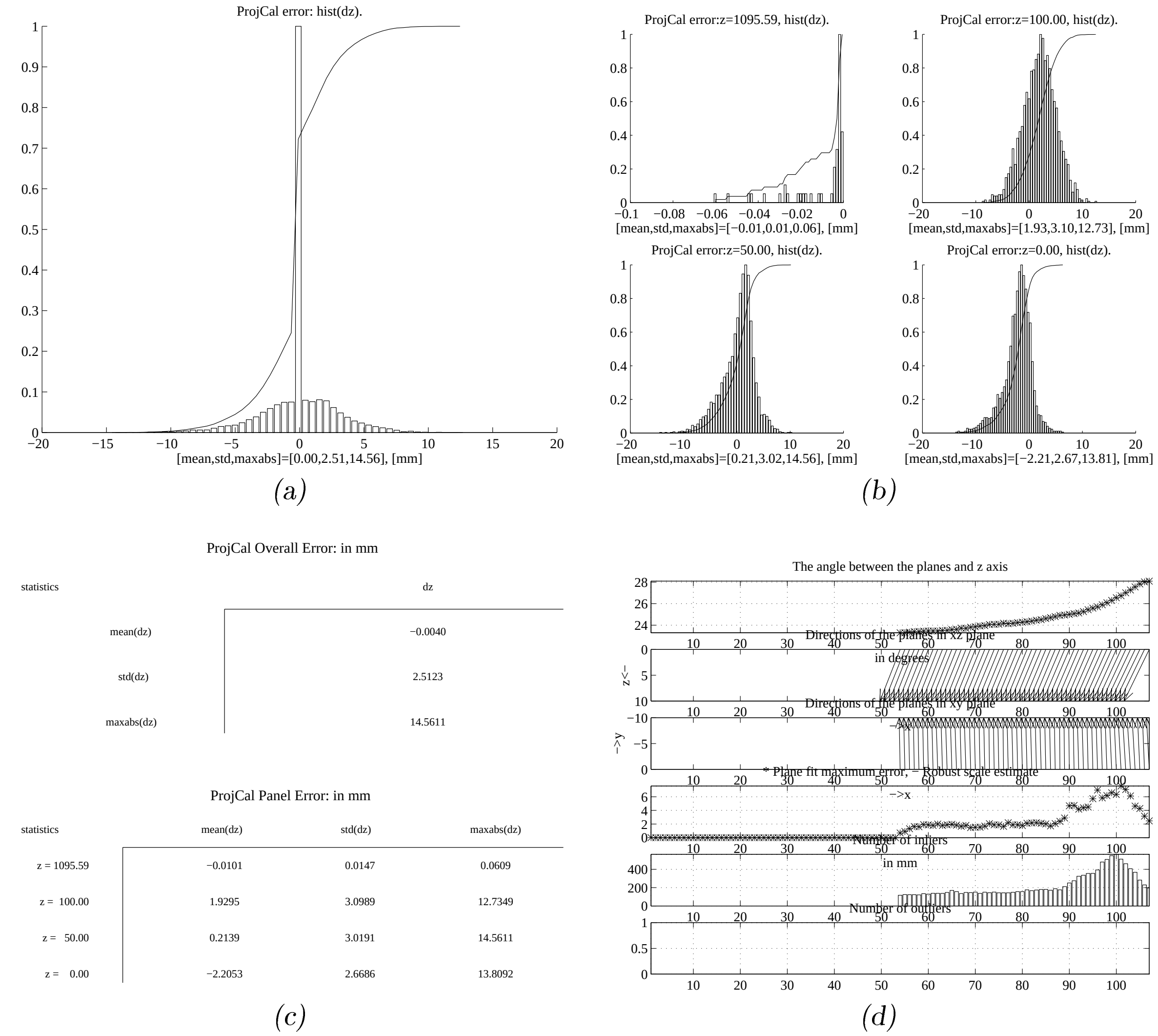
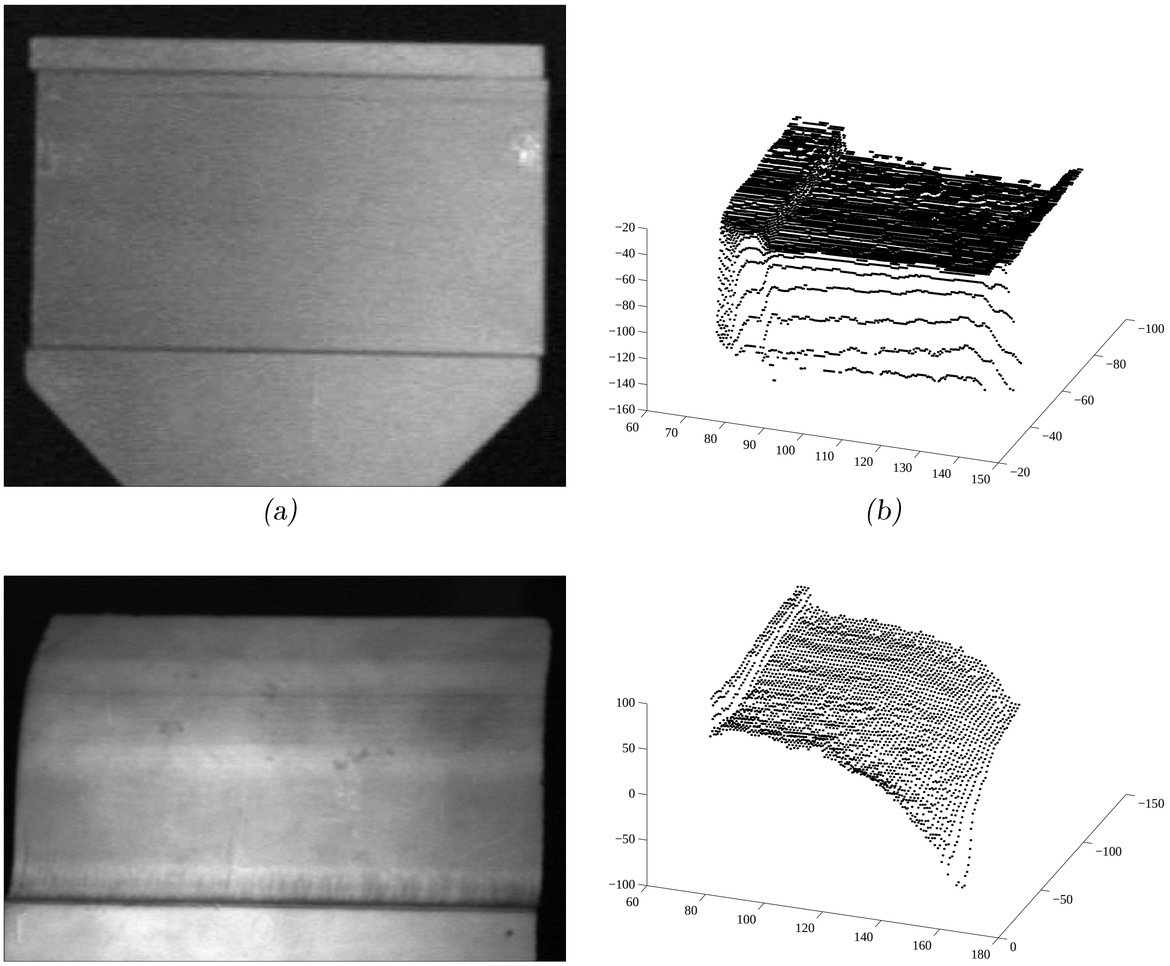



Figure 7: Calibration results. (a) shows the histogram of the distances between the points used for calibration and found planes, (b) shows the same figures for each calibration panel separately, (c) documents the errors in figures. (d) describes (from top to bottom) the angle between the monochromatic planes and z axis, orientation of each planes (two graphs), maximum error for each particular plane, number of inliers (used points for calibration of the plane), number of outliers.




(a) (b)

(c) (d)

Figure 8: Measured data. The brightness image (left) and some measured points on the object surface (right). For demonstration purposes only each fourth point in row and each four row are displayed.

[2] P. Vuylsteke and A. Oosterlinck. Range image acquisition with single binary-encoded light pattern. *IEEE Pattern Analysis and Machine Intelligence*, 12(2):143–164, 1990.

[3] R.Y. Tsai. A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323–344, 1987.

[4] Trond Melen. Geometrical modelling and calibration of video cameras for underwater navigation, trond.melen@itk.unit.no. Master's thesis, Norges Tekniske Høugskole Trodheim, November 1994.

[5] Tomáš Pajdla. Laser plane range finder the implementation at the CVL. Computer Vision Laboratory, Prague, Nr. K335-95-98, 1995.

[6] Tomáš Pajdla. Camera calibration from know scene. Computer Vision Laboratory, Prague, Memo CVL-CAMCAL-94, November 1994. pajdla@vison.felk.cvut.cz.

[7] Vladimír Smutný. Analysis of rainbow range finder errors. In Hlaváč V. and Pajdla T., editors, *Czech Pattern Recognition Workshop*, pages 59–66. CTU, Prague, November 1993.

[8] Smutný V. Optimization of color filter for rainbow range finder. In *CTU Workshop 95*, pages 173–4. CTU, Publishing House of CTU in Prague, 1995.

[9] Johji Tajima and Masato Iwakawa. 3-d data acquisition by rainbow range finder. In *10th Int. Conference on Pattern Recognition*, pages 309–13, June 1990.

[10] Tomáš Pajdla. BCRF - Binary Illumination Coded Range Finder: Reimplementation. ESAT MI2 Technical Report Nr. KUL/ESAT/MI2/9502, Katholieke Universiteit Leuven, Belgium, April 1995.