Focus set based reconstruction of micro-objects

WEDEKIND, Jan

Published version

WEDEKIND, Jan (2004). Focus set based reconstruction of micro-objects. International IEEE Conference on Mechatronics & Robotics (MechRob'04), 754-756.

Focus set based reconstruction of micro-objects

Jan Wedekind*
Microsystems and Machine Vision Laboratory
School of Engineering, Sheffield Hallam University
Pond Street, Sheffield S1 1WB, UK.
j.wedekind@shu.ac.uk
http://www.shu.ac.uk/mmvl/

21.06.2004

Abstract - With a light optical microscope and a motorized Z-table, the focus depth of a micro-object can be determined. This paper presents an algorithm for determining the depth map of the surface of a non-translucent micro-object based on focus set data.

If a robot-manipulator wants to grip a known object, such a depth map can be used to determine its position. Surface data is also useful for automatically determining gripping points. Problems such as this are addressed by the current MiCRoN[1] project.

Many algorithms addressing the problem of extracting depth maps can be found. However, most of them suffer from systematic errors. The algorithm presented in this paper efficiently suppresses these systematic errors by using a multi-scale approach.

1 INTRODUCTION

1.1 MOTIVATION

There are many differences between macro- and micro-scale tasks using robots. While task planning has to cope with problems like adhesion, vision sensors encounter the problem of limited depth of focus.

In practice, the limited depth of focus can be used to obtain depth information. Two images with different focus settings are sufficient for measuring the surface of a non-translucent object[6]. This is more feasible than stereo-vision, exspecially in the micro-world.

If a micro-robot needs to perform a gripping operation, it has to sense the object's position. [2] describes a stable algorithm for recognizing objects using depth maps. The depth map could also be used for extracting information as where to grip the micro-object.

To use the algorithm described in [2], a set of points $\{P_j\}$, that are evenly distributed on the surfaces of the objects is required. Furthermore the corresponding normal vectors $\vec{n}_j$ have to be estimated:

$$\{(\vec{p}_j, \vec{n}_j)\}, \vec{p}_j := 0\vec{P}_j, |\vec{n}| \stackrel{!}{=} 1 \tag{1}$$

The *spin image* for a point Q with normal vector $\vec{m}$ is defined as a histogram of the function $S : \mathbf{R}^2 \mapsto \mathbf{R}$:

$$S_Q(\vec{x}) := \sum_j \delta(\vec{x} - \begin{pmatrix} \sqrt{|\vec{p}_j - \vec{q}|^2 - (\vec{m} \cdot (\vec{p}_j - \vec{q}))^2} \\ \vec{m} \cdot (\vec{p}_j - \vec{q}) \end{pmatrix}) \tag{2}$$

Spin images can be used as features for image recognition because they are rotationally invariant. The similarity measure for these features is the pixel-wise cross-correlation of two histograms.

Using these features, the 3D pose of objects in cluttered scenes can be obtained with a high level of confidence.

Object recognition using spin images combined with the algorithm presented in this paper could prove to be a powerful vision algorithm.

There are various algorithms for computing a depth map which maximise a local *sharpness measure* on a focus set [4], but every local *sharpness measure* suffers from systematic errors. [5] presents a more sophisticated algorithm, discarding pixels of the depth map which are considered to be biased.

The algorithm presented in this paper does not discard pixels. Instead, a heuristic method for suppressing these errors is introduced. This paper is based on research initially published in [7] and [1].

1.2 ENVIRONMENT

To obtain the images, a microscope with a CCD-camera was used. A set of gray-level images of the micro-object was taken whilst moving the focal plane upwards relative to the object. Using a piezo-driven Z-table it is possible to take a focus set with equidistant focal planes.

The focus set

$$g \begin{cases} \mathrm{Z} \times \mathrm{Z} \times \mathrm{Z} & \mapsto & \mathbf{R} \\ (x_1, x_2, z) & \rightarrow & \text{luminosity of pixel } (x_1, x_2) \\ & & \text{in } z\text{th image} \end{cases}$$

is the input for the algorithm. The algorithm has to compute depth map from this (see figure 1).

1.3 STATE OF THE ART

1.3.1 COMPUTING A DEPTH MAP

Maximising the *sharpness* of a local image region is a widely used approach to acquire a depth map of an object.

A simple but powerful *sharpness measure* is the variance of the set of gray-level values taken from a local region of an image.

*Thanks to: Balasundram Amavasai, Manuel Boissenin, Axel Bürkle, Fabio Caparrelli, Arul Selvan, Jon Travis and everyone else who helped me writing this paper

[1] IST-2001-33567, http://wwwipr.ira.uka.de/~micron/

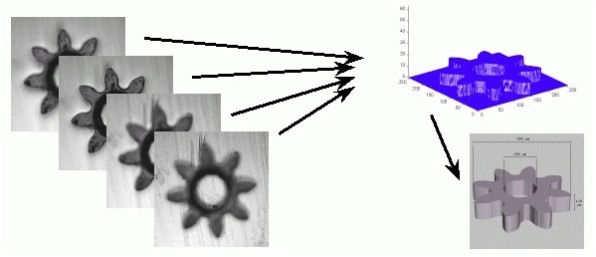

Figure 1: focus set ⇒ depth map

To generate the results shown in fig. 2, the weighted variance

$$\sigma_l(x_1, x_2, z) := \alpha_l^{(2)}(x_1, x_2, z) - \left(\alpha_l^{(1)}(x_1, x_2, z)\right)^2$$

with weighted moments $\alpha^{(1)}$ and $\alpha^{(2)}$

$$\alpha_l^{(i)}(x_1, x_2, z) := \frac{1}{n} \sum_{(x_1', x_2') \in \mathbf{Z}\times\mathbf{Z}} w_l(x_1' - x_1, x_2' - x_2)\, g(x_1', x_2', z)^i \quad (3)$$

was used, where w_l is defined recursively as follows

$$w_1(\mathrm{d}x_1, \mathrm{d}x_2) := \begin{cases} 4 & \mathrm{d}x_1 = 0 \wedge \mathrm{d}x_2 = 0 \\ 1 & \max(|\mathrm{d}x_1|, |\mathrm{d}x_2|) = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$w_l(\mathrm{d}x_1, \mathrm{d}x_2) := \sum_{\substack{-1 \le \mathrm{d}u_1 \le 1, \\ -1 \le \mathrm{d}u_2 \le 1}} w_{l-1}(\mathrm{d}x_1 + \mathrm{d}u_1, \mathrm{d}x_2 + \mathrm{d}u_2) \quad (4)$$

with l determining the size of the local region.

The depth map is obtained by maximising the *sharpness measure*:

$$dm_l(x_1, x_2) := \operatorname*{argmax}_{z \in \mathbf{Z}} \left(\sigma_l(x_1, x_2, z)\right) \quad (5)$$

1.3.2 EXTENDED DEPTH OF FOCUS

Using the depth map from equation (5), an image with extended depth of focus (=:"deep view") can be computed by simply taking the gray-level values of the pixels, which are considered to be on the surface of the object:

$$dv_l(x_1, x_2) := g(x_1, x_2, dm_l(x_1, x_2))$$

Often a depth map dm_l of poor quality can still result in a useful deep view dv_l.

2 MULTI-SCALE APPROACH

2.1 SYSTEMATIC ERRORS

Fig. 2 shows the reconstruction of a micro-sphere with a diameter of about $0.6\mathrm{mm}$. The depth map was computed from equation (5) for $l = 2$.

The resulting depth map is of poor quality.

By choosing higher values of l, a less noisy depth map can be generated while losing higher frequency details. Any

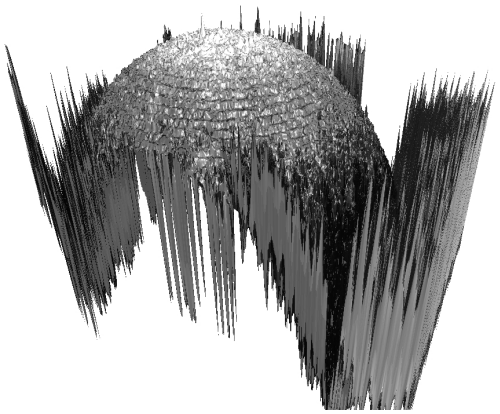

Figure 2: Reconstruction of a micro-sphere based on maximisation of a sharpness measure

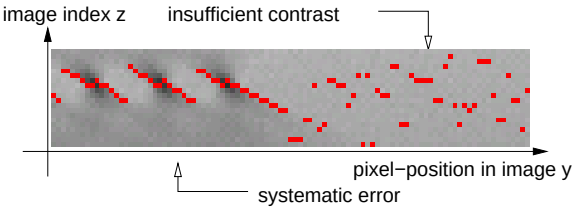

Figure 3: Systematic error of depth map

choice of l will always be a trade-off between noisiness and resolution. To tackle this problem, an adaptive algorithm could be used.

Fig. 3 shows a vertical slice through a focus set of a planar micro-grid. The red dots are the intersections with the computed depth map. There are errors not only resulting from *insufficient contrast* but also systematic errors, because the regions of *maximum sharpness* are not coinciding with the objects surface..

Fig. 4 shows an illustration of the problem. The pulse response of the light optical microscope[2] is a double cone. At the (x_1, x_2) position of the peak, the *sharpness measure* will have a maximum at the expected depth (see **b** and **d**). In contrast at neighboring positions the *sharpness measure* will have *two local maxima* (see **a** and **c**).

Any algorithm which solves these problems, will give results of much higher quality.

2.2 MULTI-SCALE APPROACH

The algorithm presented here computes a sharpness measure and a depth map for different sizes of the local region l. To save computing time the resolution is reduced whilst the filter-size is increased by sub-sampling with 2^{l-1}:

$$\sigma_l^*(x_1, x_2, z) := \sigma_l(2^{l-1}\, x_1, 2^{l-1}\, x_2, z) \quad (6)$$

[2]Here the pulse response is the focus set of a "tiny" dot. The vertical slice displayed in fig. 3 f. e. shows several structures, which are looking similar to double cones.

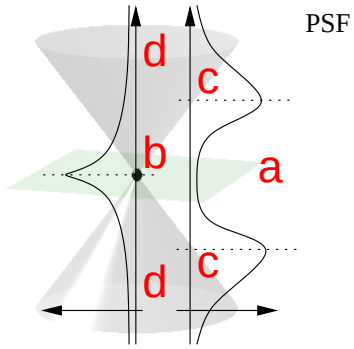

Figure 4: Schematic behaviour of the *sharpness measure* for the focus set of a peak

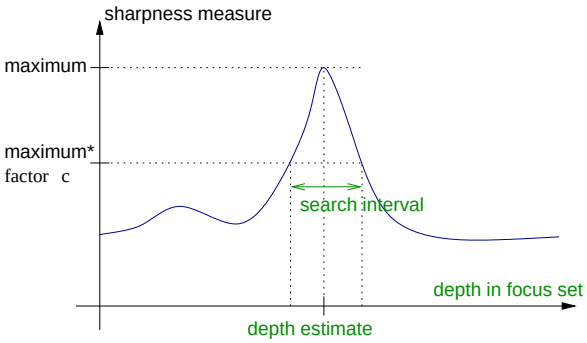

Figure 5: Definition of search interval

Also the possibility of recursively computing (6) using (3) and (4) is exploited. Using

$$idm_l^*(x_1, x_2) := \underset{z \in \mathbf{Z}}{\operatorname{argmax}} \left(\sigma_l^*(x_1, x_2, z)\right)$$

a set of intermediate depth maps

$$\left\{idm_l^* | l \in \{1, 2, \ldots, l_{max}\}\right\} \tag{7}$$

is generated afterwards.

To merge the depth maps, *sharpness intervals* are computed, as shown in figure 5, using a fixed parameter $c \in (0, 1]$:

$$I_l^*(x_1, x_2) := \left[\min\left(M_l^*\right), \max\left(M_l^*\right)\right]$$

with

$$M_l^* := \left\{z \middle| \sigma_l^*(x_1, x_2, z) \geq c\,\sigma_l^*\left(x_1, x_2, idm_l^*(x_1, y_1)\right)\right\}$$

Starting at the coarsest level l_{max} (f.e. $l_{max} = 8$), the depth map dm^* is computed by merging the set of depth maps from equation (7) as follows:

$$\begin{aligned} &dm_{l_{max}}^* := idm_{l_{max}}^* \\ &dm_{l-1}^*(x_1, x_2) := \\ &\quad \begin{cases} idm_{l-1}^*(x_1, x_2) & I_{l-1}^*(x_1, x_2) \subseteq \\ & I_l^*(\lfloor \frac{x_1}{2} \rfloor, \lfloor \frac{x_2}{2} \rfloor) \\ dm_l^*(\lfloor \frac{x_1}{2} \rfloor, \lfloor \frac{x_2}{2} \rfloor) & \text{otherwise} \end{cases} \\ &dm^* := dm_1^* \end{aligned} \tag{8}$$

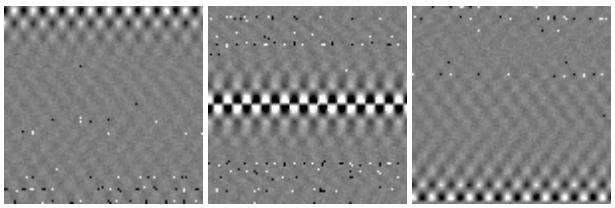

Figure 6: 1th, 50th and 99th image of a simulated focus set generated with *povray*[3]

Table 1: Results of test on simulated focus set

(c, l_{max})	variance
(0.0,1)	21.6531
(0.4,6)	1.41354

3 RESULTS

It is difficult to quantise the quality of the algorithm, because the accuracy of the resulting estimation depends on the spatial frequencies of the object's surface and texture. It also depends on the parameters c and l_{max}.

The algorithm was tested on a simulated focus set of a plane with an inclination of $\frac{0.8 \text{ focal planes}}{\text{pixel}}$ and a checker board as texture (see fig. 6). To measure its performance the algorithm was parametrised once with $(c, l_{max}) = (0.0, 1)$ and once with $(c, l_{max}) = (0.4, 6)$. A regression plane was fitted to the results and the variance of the difference between the result and the fitted plane was computed. The results are shown in table 1.

Fig. 7 shows the reconstruction of the micro-sphere using the improved algorithm. No filtering was applied so that the result could be compared to fig. 2 more easily. The algorithm also does not interpolate between the focal planes so that the depth map is showing steps.

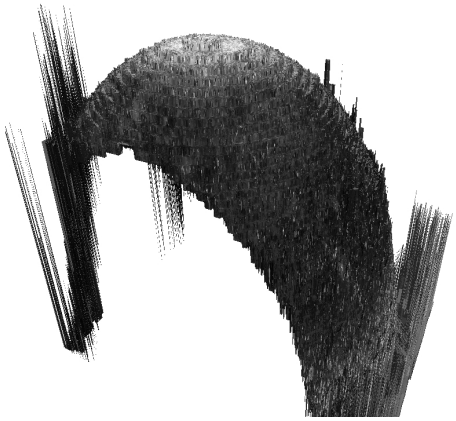

Figure 7: Reconstruction of the micro-sphere using the improved algorithm

4 CONCLUSION

Simply maximising a *sharpness measure* is not sufficient in most cases, because the regions of *maximum sharpness* and the object's surface are not necessarily overlapping. Already by applying the multi-scale approach presented in this paper, the quality of the reconstruction is improved significantly.

By improving the method of computing and composing the depth maps (see equ. (8)) or by filtering the result, the quality of the reconstruction could be improved even further.

Whilst aiming to improve the overall quality of the 3D object reconstruction, this method is already being used in the MiCRoN project to address the object recognition subtask as part of a more complex micromanipulation task. This demonstration consists of a pick-and-place manipulation of micron-sized 3D parts by using cm^3-sized robots equipped with specialised manipulation tools. The whole manipulation task is supervised by a global positioning system (used to track the position of the robots) and a local mobile miniature CCD camera mounted on a larger robot. Because of the constrained environment and the relative size of the objects with respect to the CCD camera, other 3D reconstruction methods such as depth from stereo are not physically achievable. In this case, as well as in many other applications at the microscale, 3D object reconstruction by using single viewpoint images is an obliged choice.

The software can be downloaded from `http://www.shu.ac.uk/mmvl/~engjw/surftex-1.6.tar.gz`.

References

[1] Axel Bürkle. *Optische 2,5D-Rekonstruktion mikroskopischer Objekte*. Logos Verlag Berlin, 2004. ISBN 3-8325-0468-0.

[2] Andrew E. Johnson. Spin-images: A representation for 3-d surface matching. PDF-Dokument, March 1997. `http://www.ri.cmu.edu/pub_files/pub2/johnson_andrew_1997_3/johnson_andrew_1997_3.pdf`.

[3] Persistence OF Vision RAYtracer. WWW-Seite. `http://www.povray.org/`.

[4] Gerrit Rönneberg. Mikrostrukturen im Fokus. *F&M Zeitschrift für Elektronik, Optik und Mikrosystemtechnik*, 104(10):715–720, 1996. Also see `http://www.emk.e-technik.tu-darmstadt.de/forsch/publikationen/ro_int_ver.pdf`.

[5] Torsten Scheuermann. *Berührungslose Gestaltvermessung von Mikrostrukturen durch Fokussuche*, volume 13 of *Wissenschaftliche Schriftenreihe des ICT*. Fraunhofer Institut für Chemische Technologie (ICT), Joseph-von-Fraunhofer-Straße 7, 76327 Karlsruhe, February 1997. engl. title: Profilometry of microstructures by depth from focus techniques.

[6] Masahiro Watanabe and Shree K. Nayar. Rational Filters for Passive Depth from Defocus. *International Journal of Computer Vision*, 27(3):203–225, 1998. `http://www.cs.columbia.edu/CAVE/publinks/watanabe_IJCV_1997.pdf`.

[7] Jan Wedekind. Fokusserien-basierte Rekonstruktion von Mikroobjekten. Master's thesis, University of Karlsruhe (TH), May 2002. `http://www.ubka.uni-karlsruhe.de/vvv/2002/informatik/2/2.pdf`.