3D Recovery Using Calibrated Active Camera

Rüdiger Beß & Dietrich Paulus & Heinrich Niemann
Lehrstuhl für Mustererkennung (Informatik 5)
Friedrich-Alexander-Universität Erlangen-Nürnberg
Martensstr. 3, D–91058 Erlangen, Germany
{bess,paulus,niemann}@informatik.uni-erlangen.de

please cite as

3D Recovery Using Calibrated Active Camera

Rüdiger Beß & Dietrich Paulus & Heinrich Niemann

Lehrstuhl für Mustererkennung (Informatik 5)
Friedrich-Alexander-Universität Erlangen-Nürnberg
Martensstr. 3, D–91058 Erlangen, Germany
{bess,paulus,niemann}@informatik.uni-erlangen.de

ABSTRACT

In this contribution we describe a system to generate dense depth maps from image sequences recorded from a camera mounted on a robot. Two color images of the sequence are treated as a stereo pair. A color edge detector is applied, lines are segmented from the edge images, and approximations as sequences of straight lines and circular arcs are computed. An initial depth map results from lines of these compound segmentation objects. This map is refined to a dense map by a block matching algorithm. We describe results of a parallel implementation using object–oriented programming techniques. The results show substantial improvements in comparison to a monochromous system with respect to speed, accuracy, and completeness.

1. INTRODUCTION

One goal of the image analysis system which the presented work belongs to is the active exploration of an unknown environment with an emphasis on searching and recognition of particular objects, opening a wide range of applications. Several active camera devices are used for 3D reconstruction, object recognition, and object tracking (see [5] for an overview).

One subgoal of the image analysis system is to compute the complete surface of an object as a base for modeling, classification and 3D–pose recognition in this environment. Fig. 1 shows the architecture of the system for this subgoal. An active stereo camera supervises the scene, detects interesting objects, and thus determines the starting position for a monocular camera. This calibrated camera which is mounted to a robot's hand is moved actively to record an image sequence from known positions. Any two adjacent views have approximately 20^0 difference in the angle of their optical axis; they are treated as a stereo image pair.

In this paper we show the combined feature– and correlation–based stereo approach adapted to monocular color image sequences (block *"computation of depth"* in Fig. 1). A new feature based matching using a combination of straight lines, circular arcs and chain codes is explained here. We show results of a parallel implementation on a symmetric multiprocessor system. The other components of this system are described in [1, 2].
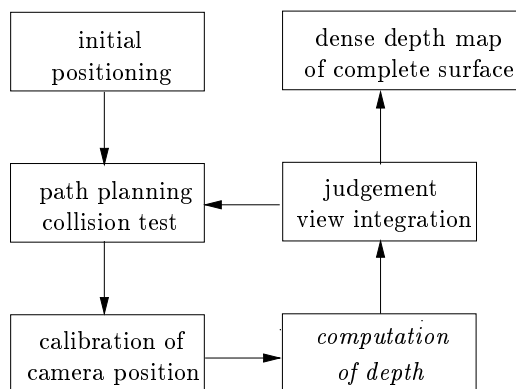

Fig. 1: System architecture

2. RELATED WORK

Feature based stereo algorithms can handle large disparity ranges with high accuracy but determine depth only for geometric primitives e.g. line segments or points, thereby generating sparse depth maps [10]. Recently, complex features such as polygons have been proposed for matching [12]. Block matching algorithms determine dense depth but their accuracy is highly influenced by illumination changes caused by different viewing angles of the moving camera. By use of the reliable sparse depth data of the feature based step the disparity range in the block matching step can be restricted, reducing the computational effort and the number of
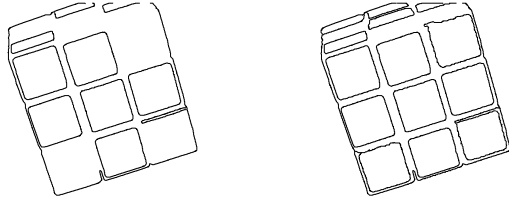
Fig. 2: Result of grey-value based segmentation (left), color segmentation (right)

erroneous matches.

Whereas traditional algorithms use gray level images, studies of color spaces showed improved performance [8]; the error in disparity has been reported smallest in RGB color space [7]. We employ color methods for all parts of a combined approach, line detection, feature based and area based stereo matching.

Other algorithms to compute dense 3D–information of an object, mainly for modelling, are known from the literature, some of them with impressing results, but none of them suitable for unknown environments, yet. Either the object is moved on a turn table e.g. in [13], or the object is placed in a calibration rig as in many photogrammetric approaches e.g. in [9].

3. INITIAL FEATURE BASED STEREO

In [10, 12] gray level image of normalized stereo pairs were segmented into straight line segments. Sets of these lines were matched in the left and right images by an iterative relaxation algorithm which used a score or quality measure on the lines.

We extended and enhanced this technique by color edge detection and line approximation by circular arcs and straight line sequences. The use of a color edge detector based on a Nevatia–Babu operator, (cmp. [10]) raised the number of discovered lines and thus lowered the number of erroneous matches caused by lines, which were detected in one image only. First, we apply an edge detector to each channel in RGB independently, in order to compute edge strength and edge orientation. The maximal strength value is selected and stored together with its orientation (see e.g. [11] for definition of strength and orientation). Fig. 2 shows a typical result for grey-value and color based segmentation. The segmentation of [12] was changed from straight line sequences to straight line–circular arc–sequences, thus reducing the number of primitives which have to be matched to 58.4% (average measured on a sample of 850 images [6]). Especially the effect of splitting

rounded lines into several short straight line segments was reduced significantly, an example can be seen in Fig. 5.

One problem of matching arcs is due to the feature attributes, such as contrast, which mean slightly different things for straight lines; the contrast for arcs is computed from the difference inside and outside the arc; contrast for lines results from the difference left and right to the line. The inner part of arcs with an horizontal tangent is left of the line, if seen from one side, and right, if seen from the other. To resolve this inconsistency, such an arc is split into two parts, leaving out the almost horizontal segment, which is not used for matching since its displacement can not be computed. This is illustrated in Fig. 3. This strategy allows for similar matching strategies for straight lines as well as circular arcs.

Fig. 3: Arc with tangent and result of splitting

The average edge strength of a line in one color channel is referred to as the color contrast of this line. Since we have three color channels we get three color contrast values, one for each channel. The color contrast values were introduced in the scoring for the matching strategy, reducing the number of potential matches at least to 50% in our experiments. This number can be further reduced by the use of an averaged color value as an attribute for matching. The color values in the areas left and right of the segmented line are averaged for each channel independently. Only those lines match which have similar values left or right of the line. 'Or' because due to occlusion an edge between two areas may change to an edge between one of these areas and the background. Wrong matches occurred only at very short line segments or if a line was detected in one image only.

To use a coarse line approximation with few primitives in the matching step while computing the disparity with the exact edge position, the following procedure was chosen, which generates a sparse depth map:

1. Chain coded lines in the two images are computed by applying a color edge filter, connecting adjacent edges and closing gaps.
2. A coarse straight line and circular arc approximation is obtained by a split and merge algorithm [6].
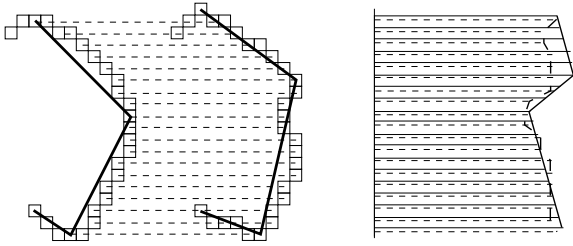
Fig. 4: Disparity computed from line sections (dashed lines) and computed from segmentation primitives (thin solid lines).
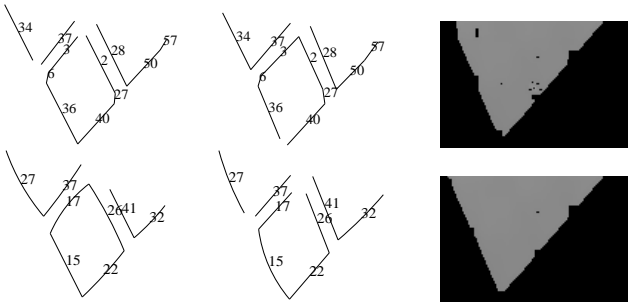

Fig. 5: Matched lines and result of interpolation step (detail) with straight line approximation (top), with straight line and circular arc approximation (bottom), each number referring to a segmentation primitive, equal numbers indicate matched primitives.

3. After the matching step for lines, the common rows of each two matched segmentation primitives are determined.

4. In each common row the positions of the chain coded line section belonging to the matched primitives are identified.

5. The disparity is computed using the chain code positions. In case of horizontal chain code links, the mean position of these codes is used.

An example illustrating the disparity computation is shown in Fig. 4. Thick lines show two pairs of matched segmentation primitives. Each box shows a pixel of the associated line segment. Fig. 4 shows an example for the algorithm above. The disparity computed for line segments is shown as dashed lines; the disparity computed from segmentation objects is depicted as thin continuous lines.

A result of the procedure is shown in Fig. 5; even for approximately straight lines the number of primitives is reduced.

4. COMBINED STEREO MATCHING

The minimal and maximal disparity and an initial disparity map computed by the feature based approach (Sect. 3) are now used as initialization of a correlation based area matching. If a local maximum of the correlation is found within two pixel distance of the estimated disparity value and this maximal value is above a threshold (typically, a value of two times the average correlation is used), the position of this maximum is taken as the correct disparity. In two cases a disparity value will remain undefined: first if no point with a correlation above the threshold is found; second if there are several local maxima for one point, the difference between these maxima is low, and there is no initial value for this point. The former indicates a point which is probably visible in one image only, the latter a homogeneous region for which borders were not found.

The computational effort needed for this algorithm is reduced to less than 20% by a quality measure, estimating the reliability in the matching of one particular block. If the quality measure is below a threshold, the disparity for this block will not be computed. If an initial disparity value is defined for this point its value will be used, otherwise the value will remain undefined. This reduces the number of points where the disparity is computed and immediately discarded because of poor reliability, as typically happening in homogeneous background regions.

The disparity computed by this approach reached an average accuracy of 1.93 pixels.

5. EXPERIMENTS, IMPLEMENTATION, AND INTEGRATION

The algorithms were implemented using ἵππος (HIP-POS) [11], an NIHCL based object oriented class library designed for image analysis. Computationally expensive parts of the stereo algorithm are parallelized on MEMSY, a modular, expandable multiprocessor system [4]. Fig. 6 shows efficiency and speedup for the parallel execution of the block matching procedure; p shows the number of processors, T_p shows the required computation time using p processor units. The parallelization of the remaining part of the algorithm is subject to current research.

For several objects the camera was positioned to record 40 views each. Adjacent views of the image sequences are subject to stereo matching. Results of the combined stereo matching for a Rubic's cube are shown in Fig. 7, In this image light grey indicates undefined regions, the disparity in pixels is mapped modulus three on white, grey and black.
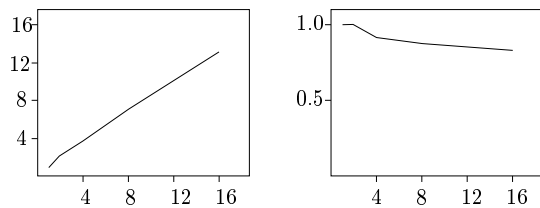
Fig. 6: Speedup $s_p = T_1/T_p$ (left) and Efficiency $e_p = T_1/(pT_p)$ (right) for the blockmatching step

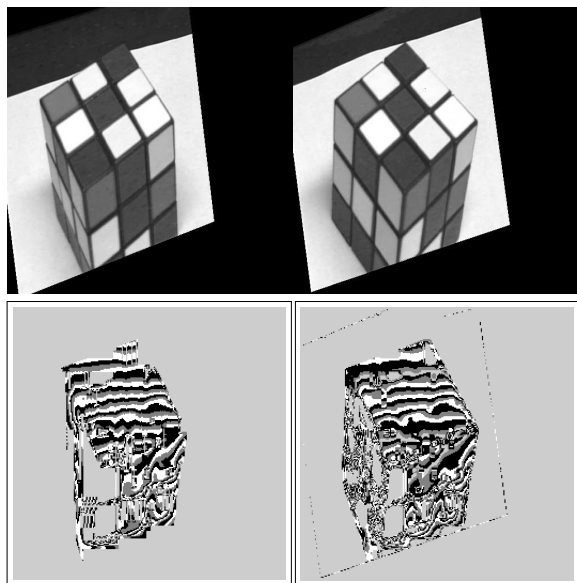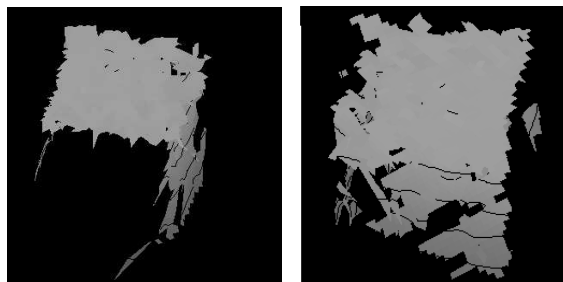Fig. 8: Resulting depth data of the cube.

Fig. 7: Combined matching: normalized input images (top), initial depth map from line based matching (bottom left), result of combined matching (bottom right).

As noted in Sect. 1, the matching results are input to further processing. Resulting range images are registered and combined to a complete 3D object [2]. Fig. 8 shows views of the final result. The missing side surface (left picture) was not visible in the image sequence.

6. REFERENCES

[1] R. Beß. Kalibrierung einer beweglichen, monokularen Kamera zur Tiefengewinnung aus Bildfolgen. In W. G. Kropatsch and H. Bischof, editors, *Tagungsband Mustererkennung 1994*, volume 5 of *Informatik Xpress*, pages 524 – 531, Berlin, 1994. Springer.

[2] R. Beß. Registering depth maps from multiple views recorded by color image sequences. In *The 4th Open Russian–German Workshop on Pattern Recognition and Image Analysis*, pages 19–24, Valday, The Russian Federation, 1996.

[3] M. Dal Cin et al. Architecture and realization of the modular expandable multiprocessor system memsy. In *Proc. of the conf. on Massively Parallel Computer Systems — MCPS '94*, Ischia, 1994.

[4] J. Denzler, R. Beß J. Hornegger, H. Niemann, and D. Paulus. Learning, tracking and recognition of 3D objects. In V. Graefe, editor, *International Conference on Intelligent Robots and Systems – Advanced Robotic Systems and Real World*, volume 1, pages 89–96, 1994.

[5] M. Harbeck. *Objektorientierte linienbasierte Segmentierung von Bildern* "Object–oriented line–based image segmentation" . Dissertation, Technische Fakultät, Universität Erlangen–Nürnberg, Erlangen, 1996.

[6] P. Koller. Computation of range images from color–stereo–images by 'simulated annealing'. In N. Pavešić, H. Niemann, and D. Paulus, editors, *Proceedings of the German–Slowenian Workshop on Image Processing and Stereo–Analysis*, pages 119–130. Arbeitsberichte des IMMD der Universität Erlangen–Nürnberg, Band 26/1, Erlangen, 1993.

[7] A. Koschan. Chromatic block matching for dense stereo correspondence. In S. Impedovo, editor, *Progress in Image Analysis and Processing III (Proceedings of the International Conference on Image Analysis and Processing)*, pages 641 – 648, Singapore, 1993. World Scientific.

[8] F. Leberl, M. Gruber, P. Uray, and F. Madritsch. Trade–offs in the reconstruction and rendering of 3–D objects. In W. G. Kropatsch and H. Bischof, editors, *Tagungsband Mustererkennung 1994*, volume 5 of *Informatik Xpress*, pages 58 – 73, Berlin, 1994. Springer.

[9] G. Medioni and R. Nevatia. Segment-based stereo matching. *Computer Vision, Graphics and Image Processing (CVGIP)*, 31:2–18, 1985.

[10] D. Paulus and J. Hornegger. *Pattern Recognition and Image Processing in C++*. Advanced Studies in Computer Science. Vieweg, Braunschweig, 1995.

[11] Stefan Posch. *Automatische Tiefenbestimmung aus Grauwert-Stereobildern*. Dissertation, Deutscher Universitäts Verlag, Wiesbaden, 1990.

[12] W. B. Seales and O. D. Faugeras. Building three–dimensional object models from image sequences. *Computer Vision and Image Understanding*, 61(3):308–324, May 1995.