

3D-Interpretation of Junctions from 2D-Correspondences in a Calibrated Stereo System

Marco Hahn, Norbert Krüger

Lehrstuhl für kognitive Systeme
Institut für Informatik

Christian-Albrechts-Universität zu Kiel

Preusserstrasse 1-9, 24105 Kiel, Germany

marco.hahn@elac-nautik.com, nkr@ks.informatik.uni-kiel.de

Abstract. We present a method for 3D junction interpretation. The interpretations include location, number of edges, and their orientations. Prerequisites are 2D interpretations of junctions in a general format found in stereo pairs. These junctions must be sets of intersecting edges and not occlusion events. We give a method for matching such 2D junctions that uses the semantics of the junctions. Using calibrated cameras we further show how to determine 3D junction location and 3D edge orientation. The final 3D interpretations allow for ambiguities.

1 Introduction

We present a system for matching 2D junctions and for interpreting the resulting 3D junctions. The interpretation consists of determining the location of the junction, the number of edges, and their orientations. Our methods allow for ambiguous junctions, i. e. junctions having edges with different degrees of certainty. The methods are described in the order of their application to images. Using sets of junctions derived from two images taken with calibrated cameras, we show how 3D junctions and edges can be obtained. Thereafter the results of experiments are given. Finally we discuss the results and possible extensions of the presented methods.

In [5] we describe our methods for the generation, filtering, and merging of 2D junction hypotheses. The preprocessing steps yield 2D junctions $\mathbf{j} = (\mathbf{x}; \{(\phi_1, c_1), \dots, (\phi_n, c_n)\})$ that consist of a location \mathbf{x} and an arbitrary number of edges (see figure 1). Each edge is described by its orientation ϕ and the confidence c . A confidence is always in the unit interval $[0, 1]$, with 1 for a certain edge and 0 for a non-existing edge. Any other method that produces junctions with such a structure can be used for detecting and evaluating 2D junctions. There exists a big variety of algorithms for 2D junction detection and 2D junction interpretation. Junction detection methods were proposed by Kitchen and Rosenfeld [6], Moravec [9], Felsberg and Sommer [3], among others. Junction

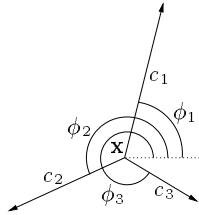


Fig. 1. General junction model for our method. Edge lengths are proportional to confidences.

classification methods have been proposed by Rohr [11], Michaelis and Sommer [8], Parida, Geiger and Hummel [10], and others. However, for 3D junction interpretation we have not found any work in the literature.

2 Generating 3D Junction Hypotheses and Analysis of Depth and Orientation

If two edges in two images are known to be images of the same real world edge, this can be sufficient to determine the orientation of the real world image. Under certain circumstances, which will be explained later, it is not fully possible to determine the orientation, but only the plane that contains the real world edge.

The first step to determine 3D location and orientation is to find corresponding features, in our case junction hypotheses and their edges, in at least two images. This correspondence problem involves a search within all used images and can be simplified, especially if the projection matrices are known. For example the epipolar line [7] gives information about the part of an image where the corresponding feature must be. After the correspondence problem is solved, the 3D location can be determined using standard methods [2].

To solve the correspondence problem, we developed two methods that use the semantic information as yielded above. The semantic 2D information is necessary to achieve a 3D interpretation. Other matching methods like template matching do often only evaluate gray scale information without assigning any semantics to it, so they cannot find edge correspondences. Thus it is not possible to derive edge orientations. Nevertheless, it can be shown that the use of gray scale information is advantageous for matching 2D junctions.

2.1 Edge-Based Junction Matching

To match a junction in a first image with another one in a second image, we compare the confidence and orientations of the edges of both junctions. Under the assumption that both cameras are close together, have similar orientations, and the imaged junction is in a sufficient distance from both cameras, it follows that corresponding edges will have similar appearance in both images, thus having

similar orientations and confidences. The junction matching process involves several steps. The first one is to match the edges. Thereafter, the similarity of the edge characterizations is determined and the effect of unmatched edges is assessed. If the epipolar geometry was established, we check the distance of the junction from the epipolar line. The weighted distance and the edge set similarity yield a measure for the probability of a junction match.

To match the edges of two junctions, we process only the orientations (see figure 2). For this we use a matrix. Each row of this matrix represents an edge of the first junction, and each column represents to an edge of the second junction. The elements of the matrix are the angular differences between the orientations of the edges denoted by row and column. The minimal entry of this matrix gives the first edge correspondence. The used columns and rows are discarded and a new minimal element is selected, yielding another edge correspondence. This process is repeated until the matrix becomes fully marked. Note that this scheme also handles junctions with different numbers of edges.

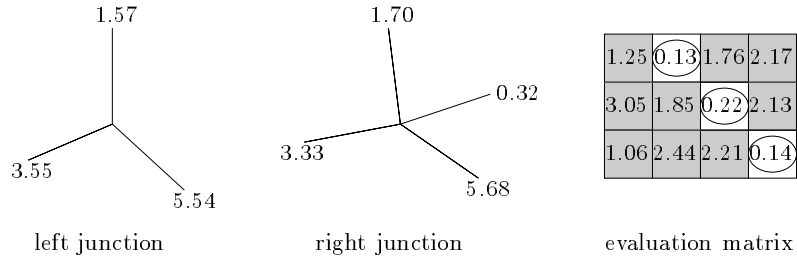


Fig. 2. Sample of an edge matching: The orientations of the edges are given in radians. The matches, in order of their finding, are (1.57, 1.70), (5.54, 5.68), and (3.55, 3.33). (Rows 1, 3, and 2.)

Thereafter we assess the edge matches and mismatches. The assessment of a given edge correspondence should be high, if both orientation and confidence of the edges are similar. This criterium is implemented using a simple heuristic. On all assessments, we apply an ordered weighting averaging (OWA) operator [12]. OWA operators can be steered to behave like the minimum or the maximum function, or any weighted average in between. In this case, we set it to behave almost like the minimum function. The rationale behind this reasoning is that all correspondences should be good, thus yielding high assessments.

If the junctions had different numbers of edges, the assessments of the edges for which no correspondences were established is combined with a maximum-like OWA operator. Finally, we subtract the second value from the first one. A high positive value represents a good match.

If the epipolar geometry was established, the distance of the junction location to the epipolar line is also computed. This distance is weighted by a Gaussian bell curve. Thus only junctions that are within a few pixels distance to the

epipolar line yield a significant weighting. The distance assessment and the edge correspondence assessment are of different nature. Therefore we use a weighted sum to combine both for the final evaluation of the match.

This method is much faster than template matching. Template matching requires roughly as many neighborhood comparisons as the image diameter is in pixels, while there are usually only up to two dozen junctions found in an image. Furthermore, this method does not have significant runtime penalties if epipolar geometry cannot be established due to uncalibrated cameras. In such a case, however, more mismatches will be made. In the following section we extend our method to incorporate gray scale information, which can be advantageous in determining junction matches.

2.2 Feature-Based Junction Matching Using Gray Scale Information

The method given above does not make any use of the underlying gray scale information. However, this information may be very useful in distinguishing between match or mismatch, see figure 3(a). Therefore we use the edges to divide the neighborhood of a junction into ‘sectors’. We do not evaluate small areas close to the edge orientations due to smearing effects.

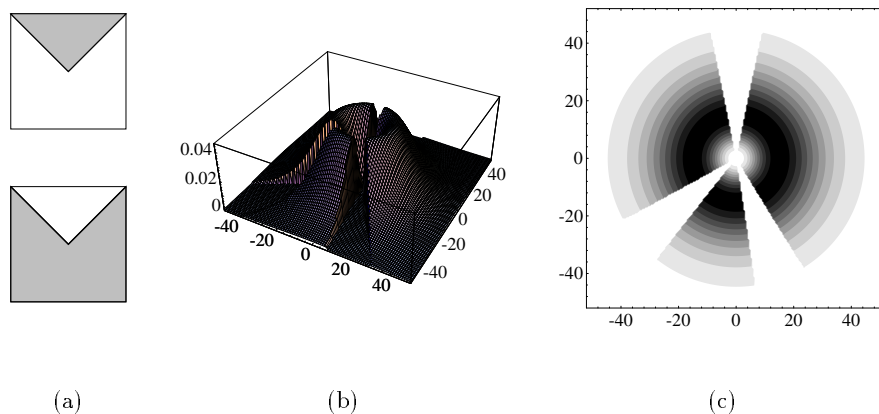


Fig. 3. (a) Two junctions with similar edges, but very different gray scale information between the edges. Usually real world junctions do not yield such different images under the given assumptions. (b) Example of the weighting function for a junction with three edges with orientations of 90° , 220° and 290° , a border area of 11.5° to both sides of the edges and parameters $b = 3$ and $c = 5$ as a 3D plot; (c) As gray scale coded (small values: light, high values: dark). The distance is measured in pixels.

For each sector a weighted average gray level is computed. The weighting factor depends on the distance to the location of the junction and is a generalized,

rotated form of Planck’s formula (see figure 3(b), (c) for an example of the weighting function for a junction with three edges). Planck’s formula originates from physics [1]; our use is motivated by its general shape, which is close to zero near the origin, thereafter rises sharply, and finally decreasing slowly to zero. This takes into account the smearing effects of differently colored areas close to the junction and the decreasing likelihood of distant regions contributing to the description of a junction.

We establish correspondences between the set of sectors using the same technique as in the previous section. Furthermore we check whether the sectors have the same order around both junctions. Finally we correlate the average gray scale levels of the matched sectors of the junctions. Since this correlation constitutes a refinement of the first matching method, its assessment is only factored into the semantic evaluation of the edges, not into the locational match derived from the distance of the 2D junctions from the epipolar line.

2.3 Computing the Orientation of Edges

Once a match is established using the methods of the previous subsection, we use the calibration information from the cameras to compute the location of the junction in 3D. This is done with standard methods [2]. If two corresponding edges are found, their location in the image, together with the optical center of the camera, allow the computation of the 3D orientation of the 3D edge. The edges of the corresponding junction must all meet in one 3D point, junctions that arise from occlusion events cannot be handled by our method. Each imaged edge plus the optical center of the camera spans a plane that contains the real world edge. The intersection of these planes is parallel to the real world edge (see figure 4(a)). If by coincidence both planes are parallel, it is not possible to determine the orientation exactly, but only the plane that contains the edge. Translations of the cameras or the real world junction do not change the setup. The orientation of the cameras has to be taken into account. We assume that the pixels are quadratic, otherwise a simple correction has to be applied [4].

The retinal plane of the camera has to be embedded into the 3D world coordinate system. For this the retinal plane becomes the xy -plane of the 3D world coordinate system and the optical center is on the positive z -axis. Note that the world coordinate system is right-handed while the embedded camera coordinate system is left-handed. Thus the sign of the y component has to be changed.

To define the orientation of a plane we need either a normal vector or two vectors \mathbf{v} and \mathbf{w} that span the plane (see figure 4(b)). We choose \mathbf{v} to be in the direction of the imaged edge and \mathbf{w} to be in the direction of the optical center from the location of the junction in the image. These vectors can be computed with the intrinsic parameters of the camera. The vector \mathbf{v} is easily computed from the orientation ϕ of the imaged edge.

The second vector \mathbf{w}' is the vector from the location of the junction to the intersection of the optical axis with the retinal plane. The components of this vector are measured in pixel. Since the third component is the focal length,

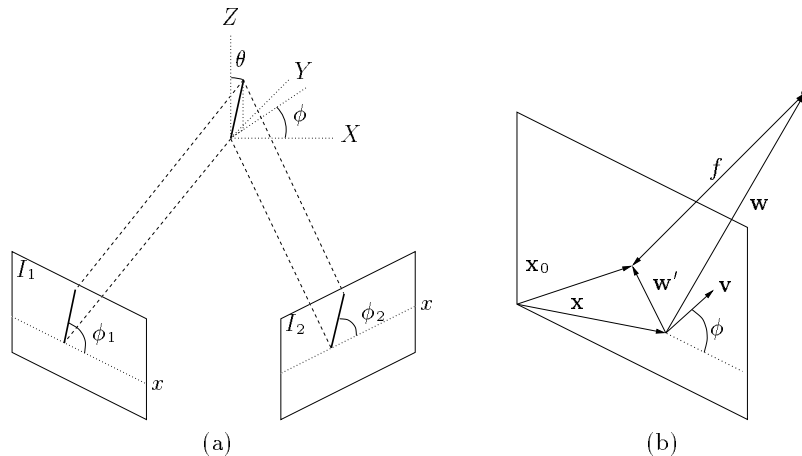


Fig. 4. (a): Setup for computing the orientation of a 3D edge; (b): The plane as spanned by the two vectors \mathbf{v} , \mathbf{w}

which is measured in meter, we correct it by the number of pixels per meter to yield the focal length in pixels. The second vector \mathbf{w} is the sum of the focal length and \mathbf{w}' . The vectors \mathbf{v} and \mathbf{w} define the plane normal vector to which the the inverse camera rotation has to be applied. These steps have to be done for both planes. Finally, the edge orientation vector \mathbf{r} is the intersection of two planes. The transformation of the resulting unit direction vector into spherical coordinates gives the angles θ and ϕ . Thus a 3D edge is given as $E = (\phi, \theta; c)$. The confidence is computed as the mean of the confidences of the corresponding 2D edges.

We already mentioned that this method does not work if both planes are parallel. This is the case if the real world edge is located in the plane that contains the real world junction and the optical centers of both involved cameras (see figure 5). We call this plane the critical plane. Due to measurement and rounding errors it can happen that the orientation vector is below the projection of the critical plane in one image and above it in the other, if the angle between the planes becomes very small. In such cases we can only give the orientation of the critical plane and not the full edge orientation as a result.

2.4 Semantic Interpretation of the Found Junction Hypotheses

The edge information of the merged junction hypotheses can be used to characterize the junction hypotheses. It may be that one or more of the edges are artifacts and do not correspond to real edges. To allow for such uncertainties, our assessment uses the following strategy:

We divide the edge set into all possible pairs of subsets. For each set of pro-edges \mathbf{E}_p (those supposed to be real edges supporting a certain assessment), we apply a minimum-like OWA operator to the confidences of the edges in the set,

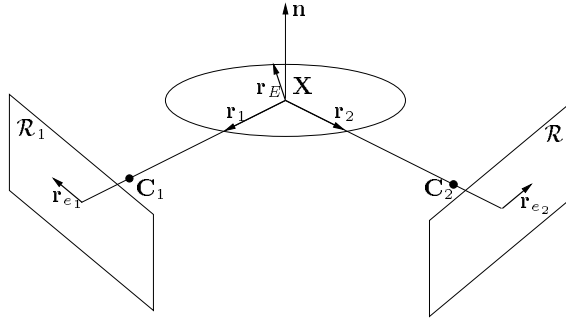


Fig. 5. Configuration of the critical plane

thus ensuring that all confidences are rather high. For each set of contra-edges \mathbf{E}_c (those supposed to be artifacts) we apply a maximum-like OWA operator. For a good assessment, the first value will be high and the second will be low. Thus the difference $c(\mathbf{E}_p, \mathbf{E}_c)$ will be a good measure of the quality of an assessment. Again, each junction hypothesis must have at least two pro-edges.

The full assessment a contains the location \mathbf{x} of the junction hypothesis, the set of pro-edges \mathbf{E}_p , and a confidence c derived from both pro-edges \mathbf{E}_p and contra-edges \mathbf{E}_c : $a = (\mathbf{x}, \mathbf{E}_p, c(\mathbf{E}_p, \mathbf{E}_c))$. We discard assessments with very low confidences. In most cases there is only one assessment with a high confidence, however ambiguous cases may yield two or more assessments with significant confidences.

3 Experimental Results

We took stereo pairs of many objects, among them our calibration object (a cube with rectangular markers). The matching methods were evaluated on this object. We compared the found correspondences with the hand-determined correspondences. Our feature-based matching found the correct match in 89.4% of the cases, and in every case the correct match was among the first 8 proposed matches. The feature- and gray-scale-based matching was correct in 86.4% of the cases; again the correct match was among the first 8 proposed matches. There were 265 junctions evaluated.

Since the sides of the rectangles on our calibration cube are parallel to the axes of the world coordinate system, we can easily determine the orientational errors by evaluating the value of the largest component of the unit vector in the direction of an edge. We noted that the orientational error of the edges depends on the locational error of the junction: If the location was determined with an error of less than 5 mm, the angular error was 10.1 degree. The angular error of the edges rose to about 30 degree for junctions where the location was determined with an error of 65 to 70 mm.

4 Conclusion

In this paper we presented a method for extracting and modelling junctions in 3D. The method is independent of the 2D junction preprocessing as long as it allows for an interpretation of 2D junctions as a center \mathbf{x} with edges of orientations ϕ_i and confidences c_i .

Potential improvements include active systems to do more thorough surveys of single junctions and algorithms that establish a global description of a given scene out of the found junctions. To find correspondences of 2D junctions in stereo images, we use semantic information and gray scale information in separate stages of the algorithm. It would be more convenient to apply both kinds of information in one scheme. Structured multivectors [3] could be a suitable framework for this kind of integration.

4.1 Acknowledgements

We would like to thank Marcus Ackermann and Bodo Rosenhahn whose work at the software library KiViGraP was very helpful for our experiments. For technical support we would like to thank Henrik Schmidt.

References

1. Marcelo Alonso and Edward J. Finn. *Physics*. Addison-Wesley, 1970.
2. Olivier D. Faugeras. *Three-Dimensional Computer Vision*. MIT Press, 1993.
3. Michael Felsberg and Gerald Sommer. Structure multivector for local analysis of images. Technical Report 2001, Christian-Albrechts-Universität, Kiel, 2000.
4. Marco Hahn. Semiglobale Verfahren zur Generierung von Eckpunkthypothesen in 2D und 3D. Master's thesis, Institut für Informatik und Praktische Mathematik, Christian-Albrechts-Universität, Kiel, 1999.
5. Marco Hahn and Norbert Krüger. Junction detection and semantic interpretation using hough lines. In *EIS*, 2000.
6. L. Kitchen and A. Rosenfeld. Gray-level corner detection. *Pattern Recognition Letters*, pages 95–102, 1982.
7. H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, 1981.
8. Markus Michaelis and Gerald Sommer. Junction classification by multiple orientation detection. In Jan-Olof Eklundh, editor, *European Conference on Computer Vision*, number 801 in Lecture Notes in Computer Science, pages 101–108. Springer-Verlag, 1994.
9. Hans P. Moravec. Towards automatic visual obstacle avoidance. In *Proceedings 5th International Joint Conference on Artificial Intelligence*, page 584, 1977.
10. Laxmi Parida, Davi Geiger, and Robert Hummel. Junctions: Detection, classification, and reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(7):687–698, 1998.
11. Karl Rohr. Recognizing corners by fitting parametric models. *International Journal of Computer Vision*, 9(3):213–230, 1992.
12. Ronald R. Yager. Ordered weighted averaging aggregation operators in multi-criteria decision making. *IEEE Transactions on Systems, Man and Cybernetics*, 18:183–190, 1988.