2010 IEEE International Conference on Robotics and Automation
Anchorage Convention District
May 3-8, 2010, Anchorage, Alaska, USA

Physical Road Marker Property Estimation using Monoscopic Vision

Thomas Gumpp, Dennis Nienhüser, J. Marius Zöllner

Abstract— **In this paper algorithms are presented to extract lane markers and their properties from monoscopic camera images. A filter approach that takes into account the visual appearance of the markers is presented. Another contribution constitutes the measurement of physical marker width and length from perspectively distorted images using only the calibrated camera images. Also distances between markers can be estimated. All algorithms are independent from a specific lane model.**

I. INTRODUCTION

Research in autonomous vehicles has gained traction tremendously in the last years. Due to competitions like the DARPA Grand Challenge and the DARPA Urban Challenge, a lot of effort has been spent to create more intelligent systems. One goal of this development is the creation of automated vehicles that can safely be operated within the existing road infrastructure, where automated and conventional vehicles are going to drive at the same time. Another goal that can be seen as a preliminary step is the assistance of the driver through guidance and warning signals. In all these scenarios it is crucial to build a scene representation as complete as possible. Using different sensors and obtaining information from sensor fusion is commonly accepted as a necessity. One example which will be the subject of this paper are road markings. Current research concentrates on extracting marker positions and gradients to track lanes using mostly Kalman- and Particle Filters [1][7].

To safely navigate on roads, more information than the lane course is needed. The following examples are taken from german roads, but similar concepts can be found worldwide: A solid line indicates a lane border that must not be crossed for overtaking. When leaving a motorway, deceleration lanes are used to adjust the vehicle's speed. These lanes are separated from standard lanes using wider markers painted at a higher frequency [9].

Driving on these lanes autonomously, or in an assistance scenario, mandatory traffic rules can only be adhered when information about the different lane markers is available.

II. STATE OF THE ART

Only a few publications dedicated to the subject of lane marker classification can be found. Most of them address the discrimination of dashed and solid lines. [2] presented an approach that used intensity thresholds to distinguish dashed from solid markers. Other than that, different marker types cannot be detected. [5] used horizontal scanlines to detect marker information, and represent it as symbolic information. [6] showed how to extract lane markers using a fixed set of shapes by comparing moment features. [4] used frequency analysis to determine lane marker types for straight lanes.

To extract all information embedded in the lane markers, a more flexible solution is needed. This solution must be able to directly measure marker length and width in a metric world coordinate system. Extraction of marker information must be invariant to the vehicles relative pose. Also to reliably extract lane boundary types, the length of gaps between markers is an important feature, which has to be considered.

T. Gumpp , D. Nienhüser and J. M. Zöllner are with Intelligent Systems and Production Engineering, FZI Forschungszentrum Informatik, Karlsruhe, Germany. {gumpp, nienhues, zoellner}@fzi.de

III. MARKER DETECTION

The goal of the first part of the presented algorithms is to provide a method for the lane model independent detection of marker candidates in the field of view of the camera. These are used for generating region of interests for the subsequent extraction algorithms.

A. Marker candidate extraction

Lane markers are usually painted on tarmac or similar road surfaces in a bright color such as white or yellow. Therefore orthogonally to the marker direction a typical signature of dark road color, bright marking and again dark road color is to be expected. On German roads and highways, marker width has a specified range of 12-30 cm. These properties have been used to develop the specific filter mask for lane markings now presented. It is based on the lane marker extraction algorithms presented in [8].

To extract the road markings, the grayscale image is convoluted horizontally and vertically using the central, symmetrical segment of a gaussian kernel $g(x)$, which is defined by the interval $[-l_g, l_g]$. An offset o_g is subtracted from the gaussian so that negative and positive areas enclosed by the graph and the x-axis are equal, i.e. the integral of the kernel equals 0. To calculate said offset, the gaussian cumulative distribution function $c_g(x)$ is used to define the roots of the desired function, $-x_0$ and x_0:

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (2)$$

$$c_g(x) = \frac{1}{2}(1 + \mathrm{erf}(\frac{x-\mu}{\sigma\sqrt{2}})) \quad (3)$$

$$0 = (l_g - x_0) \cdot g(x_0) - (c_g(-x_0) - c_g(-l_g)) - \quad (4)$$
$$(c_g(0) - c_g(x_0) - (x_0 \cdot g(x_0))) \quad (5)$$
$$o_g = g(x_0) \quad (6)$$

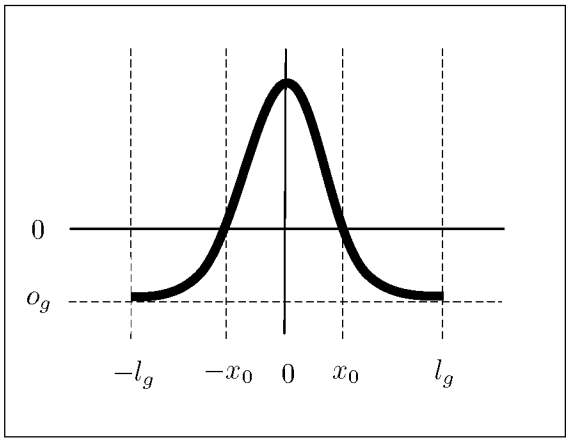

Fig. 1. Plot of $g(x) - o_g$ used as underlying function to sample convolution masks for extracting lane markers [8].

The variance σ of the gaussian kernel influences its performance on lane markings of variable width and has been experimentally chosen.

The gaussian kernel mask is sampled in a way that the distance d_m of the roots $2 \cdot x_0$ is equal to the width of lane markings. For each image line this width in pixels has to be calculated using extrinsic and intrinsic camera parameters (17).

As marker width is increasing within the lower parts of the image, very long filter masks up to 90 pixels for the used camera configuration would be necessary. Folding the image using such large masks requires a high computational effort that contradicts near real-time operation of the algorithm. Instead a pyramidal approach is carried out by scaling image regions so that the marker width in pixels remains equal to an acceptable width. To create the complete filter result, the filter response of these regions has to be scaled up again.

Figure 2 shows the resulting marker confidence image. To actually gain distinct marker candidates, scanlines are used to extract maximum peaks in the image. This is resulting in a list of 2d marker candidates in the image plane. These are transformed to metric positions using the following approach.

B. Inverse Perspective Transformation

1) Coordinate Systems: Two coordinate systems are relevant in the context of this paper: The camera coordinate system (CCS) and the word coordinate system (WCS). Pixel positions on the image plane are expressed in the CCS using u and v for horizontal and vertical coordinates, the origin being the top-left corner. The WCS is a coordinate system following the right-hand rule modeled after ISO 8855 which has x forward, z up and y pointing to the left-hand side of the vehicle.

Fig. 2. Result of filtering a greyscale camera image with the marker extraction mask.

$$X_{CCS} = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}, \qquad X_{WCS} = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (7)$$

2) Projection: The camera image is a result of the perspective projection of the world near the vehicle onto the image plane. The projection of a point X_{WCS} given in the WCS into the CCS can be expressed using homogeneous matrix multiplication. Necessary are the projection matrix P and the matrix R, which contains a homogeneous transformation matrix with the camera's inverse orientation and translation in WCS. The camera's focal lengths are f_x, f_y; its principal point is u_0, v_0.

$$P = \begin{pmatrix} u_0 & -f_x & 0 & 0 \\ v_0 & 0 & -f_y & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad (8)$$

$$X_{CCS} = P \cdot R \cdot X_{WCS} \quad (9)$$

3) Inverse Projection: To reverse the projection from CCS to WCS, the inverse of the projection matrix is needed. Due to the unknown distance of the projected point, the inverse P^{-1} is defined as

$$P^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ -\frac{1}{f_x} & 0 & \frac{u_0}{f_x} \\ 0 & -\frac{1}{f_y} & \frac{v_0}{f_y} \\ 0 & 0 & 1 \end{pmatrix} \quad (10)$$

$$X_{WCS}^{IP} = R^{-1} \cdot P^{-1} \cdot X_{CCS} \quad (11)$$

To create a useable inverse transformation from the CCS to the WCS, the well-known flat earth assumption is used. It allows us to represent the detected marker candidates in the world coordinate system by limiting possible world positions of the measurements to the floor plane. The coordinate z of the WCS point X_{WCS}^{FP}, that is projected onto the floor plane, is assumed to be equal to 0. This leads to a bijective relation

between pixels and points on the floor plane. X^C_{WCS} contains the position of the camera.

$$X^C_{WCS} + r \cdot (X^{IP}_{WCS} - X^C_{WCS}) \stackrel{!}{=} 0 \tag{12}$$

$$\Rightarrow \qquad r = \frac{z^C}{z^C - z^{IP}} \tag{13}$$

$$X^{FP}_{WCS} = X^C_{WCS} + r \cdot (X^{IP}_{WCS} - X^C_{WCS}) \tag{14}$$

In the representation of the WCS, equidistant sampling of the most confident points is possible .

C. Orientation Projection / Inverse Projection

The orientation of objects in the WCS can be projected by transforming start- and endpoints X^S_{WCS}, X^E_{WCS} of a short directional vector from the WCS to the CCS. The orientation in the image plane can be calculated using

$$\alpha_{CCS} = \text{atan2}(u^E_{CCS} - u^S_{CCS}, v^E_{CCS} - v^S_{CCS}) \tag{15}$$

To inversely project the orientation of gradients extracted from the image plane onto the floor plane, basically the same technique is used. Start- and endpoints of the gradient X^S_{CCS}, X^E_{CCS} in the image plane are projected to the floor plane, therefore their z coordinate is 0. The gradient can then be calculated as

$$\alpha_{WCS} = \text{atan2}(x^E_{WCS} - x^S_{WCS}, y^E_{WCS} - y^S_{WCS}) \tag{16}$$

D. Length Projection / Inverse Projection

Length projection is also needed later in the presented algorithms. To project a length from the WCS to the CCS, the length L_{WCS} needs to be given as a relative vector to its base position X_{WCS} in the world. The projected length l_{CCS} is then defined as

$$l_{CCS} = ||X_{CCS} - P \cdot R \cdot (X_{WCS} + L_{WCS})|| \tag{17}$$

To project a length L_{CCS} measured on the image plane onto the floor plain, the same principle of projecting a distance is used. Base position X_{CCS} and the offset-vector $O_{CCS} = X_{CCS} + L_{CCS}$ are inversely projected to the floor plane. The distance is then calculated as

$$l_{WCS} = ||O_{WCS} - X_{WCS}|| \tag{18}$$

IV. MARKER GRADIENT AND WIDTH

After initially detecting lane marker candidates, the algorithms presented in this section provide additional information about the markers.

A. Orientation

Detecting orientation of the marker candidate is vital for the following steps. Marker width and length can only be measured reliably along and orthogonal to the axis of the marker, which in turn is determined by its orientation. To create a robust estimation, a ROI in the grayscale image around the marker candidate is created. This region is meant to be the equivalent of a 40 cm square, so this distance is perspectively projected at the position of the detected marker to calculate the equivalent size in the image plane in pixels using equation (17). An example for projected regions can be seen in figure 3.

Fig. 3. Edge intensity in ROIs used for marker gradient and marker width estimation.

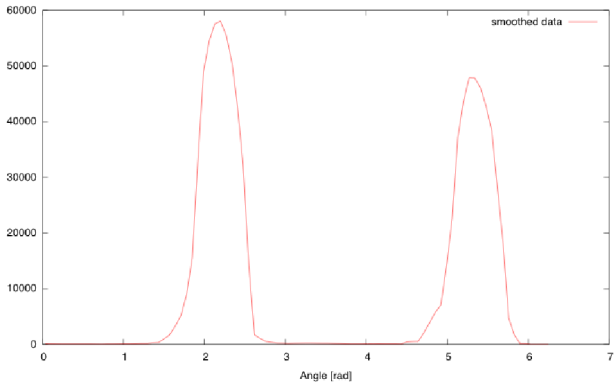

Fig. 4. Intensity-weighted histogram of edge orientation gained from a ROI patch.

After using sobel filters to extract horizontal and vertical gradient images of the ROI, these are used to calculate intensity and phase edge images. To create a representation of the gradient distribution in the ROI, the phase values are aggregated in a histogram. The phase value determines the histogram bin. The corresponding intensity is interpreted as a confidence measure for detected edges, and therefore mapped to the increment value. The resulting histogram is smoothed by gaussian filtering.

An example of such a histogram can be seen in figure 4. Both peaks for the opposite edges of the marker are clearly visible. As the sign of the extracted orientation is irrelevant, the histogram bins corresponding to the same absolute angle are combined and the bin with the peak in the resulting histogram is taken as dominant gradient for the ROI.

B. Inverse Orientation Projection

The gradient extracted from the edge image is also based on the perspectively disturbed view of the lane markers. To gain orientation information useable in the context of the WCS, 2d orientation has to be inversely projected onto the floor plane using equation (16). Figure 5 shows the resulting marker orientation from figure 3 in a birdview representation.

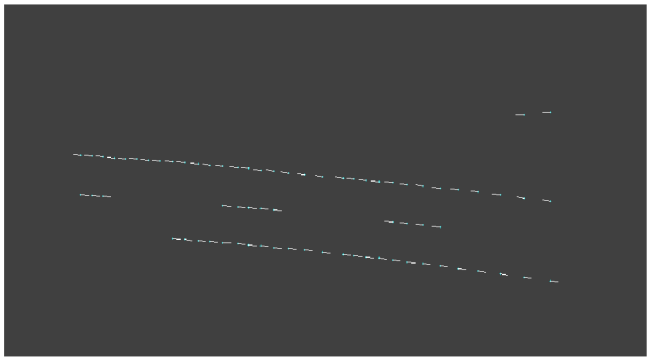

Fig. 5. Bird view representation of extracted gradients in the WCS.

C. Width Extraction

The goal of the following set of algorithms is the visual estimation of marker width. The relevant visual features are the edges between marker and road surface. Orthogonal to the 2d gradient in CCS extracted in section IV-A, a projection base line is constructed in the CCS as a weighted histogram. The histogram's bin for the value 0 is aligned with the position of the marker. The corresponding bin for each ROI point is then calculated by projecting it onto this base line, and the increment is again determined by edge intensity. Figure 6 shows a result.

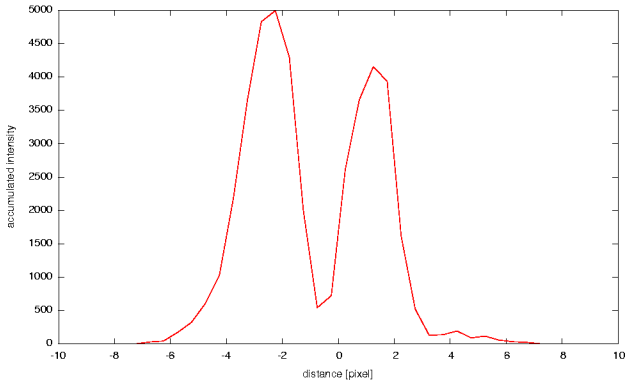

Fig. 6. Intensity weighted histogram of projected edge pixels. Left and right marker edges at -2.25 and 1.25.

Splitting the histogram at bin number 0, representations for the left and right edge distribution are separated. After smoothing each of these histograms, the bin with the highest values represents the position of the left and right marker edge. Adding the absolute values of these positions leads to the 2d orthogonal distance d_{CCS} between the marker edges in the CCS. For the example visualized in figure 6, d_{CCS} would be 3.5 pixels wide.

To calculate the accurate width of the marker, the angle difference γ between the normal of the 2d gradient in the CCS , and the projection of the normal of the 3d orientation in WCS (see (15)) has to be considered. Therefore, the width of the marker w_{CCS} is calculated as:

$$w_{CCS} = d_{CCS}/cos(\gamma); \tag{19}$$

Figure 9 shows the result of the presented algorithm. Marker width is extracted for all visible lane markers.

Fig. 7. Extracted marker widths for multiple lanes in the image plane.

To express the extracted marker width in meters, it has to be inversely projected back onto the ground plane using equation (18).

The estimated position of the marker coordinates is meant to be in the middle of the lane markers. Due to sampling and scaling errors, imperfections can accumulate to a degraded position estimation. By setting the coordinates exactly between the left and right marker edge, the marker position can be estimated more accurately in the CCS. The adjusted 2d coordinates are inversely projected to also correct the 3d position in the WCS.

V. MARKER LENGTH

Markers can occur as a solid line, or dashed in different lengths and frequencies. Solid lines must not be crossed by a car. The different kinds of dashed lines carry information about the lanes that they are separating. One example on the german Autobahn are the markers that separate regular lanes from exit lanes, as these are wider and the gaps between the markers are narrower than the markers, that separate regular lanes. So to safely navigate on these types of lanes, it is

crucial to extract the correct length of markers and the gaps in between.

A. Point Linking

First the lane marker measurement points need to be linked to marker strips. Using an agglomerative clustering approach, strips of marker points are created.

B. Lane Marker Patch Generation

A suitable representation of the marker strip image region is a necessary prerequisite to reliable extract the length of the markers. In the given perspectively distorted camera image, lengths are not directly measurable. Also in curves, the markers are not on a straight line, which also causes problems for extracting lane marker length.

So ideally the marker image region would be in a straight line, without perspective distortions, in which the marker's dimensions are unaffected by its position in the image. It should have a width that is slightly wider than the maximum marker width of 30 cm.

To create this representation first a suitable result image has to be created. Experiments showed that a resolution of 50 pixels per meter yields good results, so the resulting image height is $0.4m \cdot 50 = 20$ and its width is $l \cdot 50$, where l is the length of the concatenated segments, which are created by connecting neighboring lane marker points.

For each segment, a ROI is projected into the source image. These ROIs are undistorted and stitched together into the result image by using a warp transform. The resulting image is stitched from the chained ROIs creating an image strip. In curves, the segments gently follow the bent markers. As these segments are then mapped to the straight result image, the markers are also straightened. An example for a resulting marker strip image is the top row of figure 8.

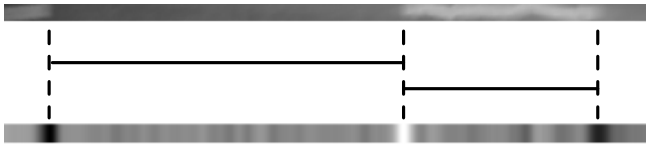

Fig. 8. Top row shows a marker strip image generated from lane marker candidate positions. Bottom row shows the result of applying a Prewitt filter to the vertically averaged image. Lines indicate gap- and marker length.

C. Length extraction

To perform length extraction using the undistorted, straightened marker strip image, first an average of the intensity in each column is calculated. This results in a vector in which higher values indicate a marker, while lower values are to be expected for the gaps in between. Binarizing this data using a threshold between the min and max values seems to be straightforward, but unfortunately the intensity levels even within one frame vary greatly. This can be to the extend that the values that are measured from a marker near the car are lower than those measured in a gap between markers at a distance of 40m. This can also be observed in figure 8.

As using the absolute values in the vector is not feasible, the marker and gap ranges have been extracted using gradients. Resolution in the original image is low for markers further away from the vehicle, so the interpolated intensity image shows no sharp edges, but a rather shallow slopes (compare top row of figure 8). Therefore the edge extraction convolution kernel needs to be long to cover these smooth transitions between markers and road surface. A empirically determined one-dimensional Prewitt edge filter with a length of 30 pixels is sufficient.

The bottom row in figure 8 shows a typical filter response. Positive/negative peaks are shown as white/black respectively. To get the borders between marker and road region, the maximum absolute values are searched within a sliding window of 1m width. Positive peaks indicate a transition from road to marker, while negative peaks are present at the end of a marker. The distance between following positive and negative peaks is stored in a marker length histogram, while the distance between negative and positive peaks is stored in a gap length histogram (compare figure 8). Distances of peaks with identical signs are ignored. The histograms are used as a mean of temporal fusion, to gracefully handle measurement errors. Figure 10 shows a resulting histogram. After smoothing, the peak of these histograms represent the extracted lengths.

VI. EVALUATION

The following tables contain standard values for marker width and length, and the distance between markers. These values are valid for german roads, and taken from *Richtlinien für die Markierung von Straßen*(RMS)[9], which also contains additional information.

	Autobahn	other roads
narrow marker	15 cm	12 cm
wide marker	30 cm	25 cm

This table shows commonly used marker widths on german roads.

	motorway	country	city
standard	6 m (12 m)	4 m (8 m)	3 m (6 m)
warning	6 m (3 m)	4 m (2 m)	3 m (1.5 m)

This table shows the lengths used for markers (gaps). Standard markers are half the length of their distances, while warning markers are painted double the length.

For the evalutation a motor way sequence has been chosen to validate extracted marker dimensions. According to RMS, on a german Autobahn standard dashed lane markers are 15 cm wide and have a length of 6 m. The gap between the markers should be 12 m. These dimensions are evaluated against the extracted results of the presented algorithms.

Figure 9 shows a histogram of marker widths extracted from the video sequences before smoothing. The peak clearly has its center at about 15 cm. According to RMS[9], the other possible marker width for motorways is 30 cm, which could be easily distinguished.

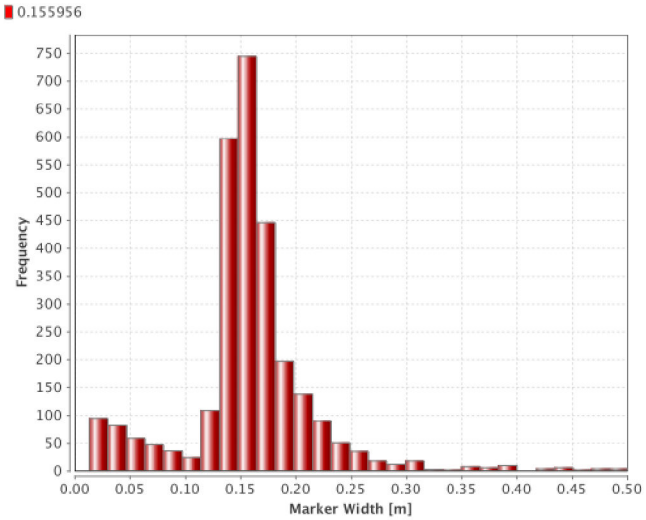

Fig. 9. Histogram of marker width

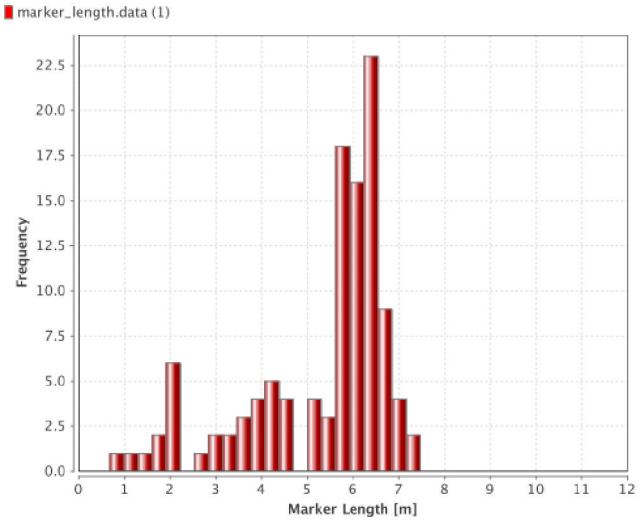

Fig. 10. Histogram of marker length

The raw histogram resulting from marker length extraction can be seen in figure 10. It shows some shorter measurements, but the bulk of measured lengths can be seen at the 6m mark. This is also the only valid marker length for motorways. It is accurately enough to distinguish it from 4m, which is used on country roads[9].

Figure 11 shows a histogram of extracted length values for the gaps between the markers. The values are not as clearly aligned, as it is the case with lane marker length and width values. Nevertheless according to [9], valid lengths on motorways are 12m gaps for a standard lane marking, and 3m gaps for special lane markings, which could be easily distinguished.

VII. CONCLUSION

In this paper algorithms have been presented to extract lane markers and their properties from camera images. A robust filtering approach has been presented to extract markers from

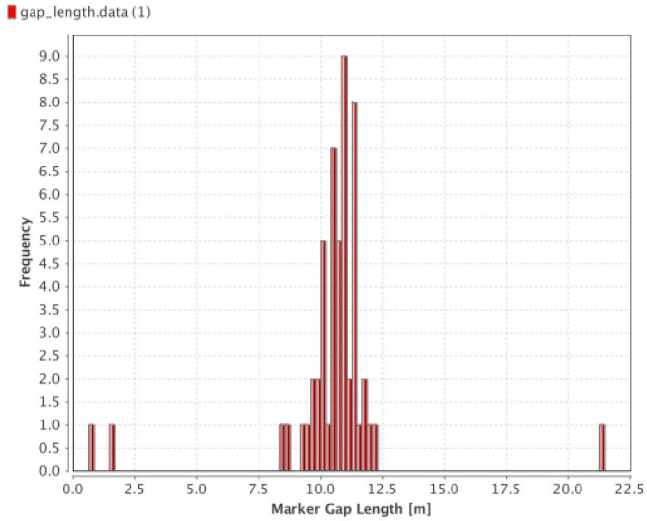

Fig. 11. Histogram of the distance between markers

greyscale images. Also their orientation has been determined in the WCS. A major contribution constitutes the measurement of physical marker width and length from perspectively distorted images using only a calibrated camera image. Also distances between markers can be estimated. This work is a necessary base for a more complete scene interpretation and representation.

VIII. ACKNOWLEDGEMENT

This work was supported by Harman/Becker Automotive Systems GmbH.

REFERENCES

[1] E. D. Dickmanns and B. D. Mysliwetz, *Recursive 3-d road and relative ego-state recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992.

[2] K. P. Wershofen, Real-time road-scene classification based on a multiple-lanetracker, *Industrial Electronics, Control, Instrumentation, and Automation*, 1992.

[3] Y. Otsuka, Multitype Lane Markers Recognition Using Local Edge Direction, *IEEE Intelligent Vehicles Symposium*, 2002.

[4] J. M. Coltado et. al, Detection and Classification of Road Lanes with a Frequency Analysis, *IEEE Intelligent Vehicles Symposium*, 2005.

[5] S. Vacek et. al, Road-marking analysis for autonomous vehicle guidance, *European Conference on Mobile Robots*, 2007.

[6] Y. Li et. al, Road Markers Recognition Based on Shape Information, *IEEE Intelligent Vehicles Symposium*, 2007

[7] N. Apostoloff, Vision Based Lane Tracking using Multiple Cues and Particle Filtering, Phd Thesis, 2005.

[8] T. Gumpp et. al, Recognition and Tracking of Temporary Lanes in Motorway Construction Sites, *IEEE Intelligent Vehicles Symposium*, 2009.

[9] Forschungsgesellschaft für Straßen- und Verkehrswesen. Richtlinien für die Markierung von Straßen (RMS) Teil 1: Abmessungen und geometrische Anordnung von Markierungszeichen (RMS-1) (in german), 1993.