Free-space Detection using Online Disparity-supervised Color Modeling

Willem P. Sanberg, Gijs Dubbelman and Peter H.N. de With

***Abstract*— This work contributes to vision processing for intelligent vehicle applications with an emphasis on Advanced Driver Assistance Systems (ADAS). A key issue for ADAS is the robust and efficient detection of free drivable space in front of the vehicle. To this end, we propose a stixel-based probabilistic color-segmentation algorithm to distinguish the ground surface from obstacles in traffic scenes. Our system learns color appearance models for free-space and obstacle classes in an online and self-supervised fashion. To this end, it applies a disparity-based segmentation, which can run in the background of the critical system path and at a lower frame rate than the color-based algorithm. This strategy enables an algorithm without a real-time disparity estimate. As a consequence, the current road scene can be analyzed without the extra latency of disparity estimation. This feature translates into a reduced response time from data acquisition to data analysis, which is a critical property for high-speed ADAS. Our evaluation over different color modeling strategies on publicly available data shows that the color-based analysis can achieve similar (77.6% vs. 77.3% correct) or even better results (4.3% less missed obstacle-area) in difficult imaging conditions, compared to a state-of-the-art disparity-only method.**

I. INTRODUCTION

In recent years, vehicles are becoming increasingly intelligent with so-called Advanced Driver Assistance Systems (ADAS). This development is expected to significantly reduce traffic accidents, traffic congestion and fuel consumption simultaneously. To ensure traffic safety, ADAS can e.g. indicate the location of potentially hazardous obstacles to the driver and the area position of safely drivable road. On the longer term, ADAS and related technologies will allow the development of fully autonomous vehicles. In this work, we improve a state-of-the-art vision-based free-space detection system by efficiently exploiting multiple image modalities.

To robustly facilitate situational awareness at a moving platform, several complementary sensor modalities should be employed. These modalities can include RADAR, LIDAR, ultrasound, and (thermal) imaging. The benefit of using vision-based systems is that they provide dense scene information in a cost-effective way. Image data is also a rich source of information, since it comprises of several informative properties. For stereo-based video imaging, these informative aspects include not only the usual *texture, color* and *shape* features, but also *optical flow* motion analysis and *disparity* estimation. All these elements can contribute to a robust situational analysis, such as e.g. the detection of partially occluded pedestrians who are about to cross the street. Although LIDAR, RADAR or ultrasound provide valuable complementary information, in this paper we solely focus on vision-based detection systems.

Willem Sanberg, Gijs Dubbelman and Peter de With are with the Department of Electrical Engineering, Video Coding and Architectures Research Group, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands `w.p.sanberg@tue.nl`

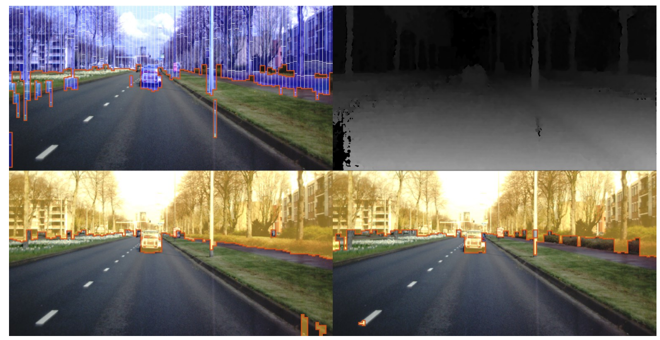

Fig. 1. Stixel segmentation results comparing a disparity-only result (top left), a RGB-only result using 10 frames for online learning (bottom left), and a new gray-only result for which just 3 frames are used for the online learning (bottom right). The orange overlay with a dark border depicts the space occupied by obstacles according to the detection algorithm. At the top right, the disparity signal is shown, which has several artifacts, due to low texture in the road region (left) and a pole reflection in the windshield (middle). These artifacts cause false obstacle detections in the original disparity-based algorithm [1]. Here, we show that it is possible to obtain similar or even better segmentation results with our color modeling, while requiring less data in the process.

Multi-view image processing, in particular stereo vision, has the potential to provide 3-D scene information at a more affordable price point than that of high-end laser-based systems, which are often accompanied by RTK-GPS, e.g. [2][3]. In stereo vision, the disparity, which is analogous to depth, can be estimated densely and in real-time [4]. This gives a direct description of the geometry of the scene and it facilitates, for example, a separation of flat, drivable surfaces from erect obstacles [5][6]. The Stixel World method [1] is a state-of-the-art approach to analyze such a geometry description of the scene. It is a fully probabilistic framework to distinguish free space from obstacles in the disparity signal, which can be implemented efficiently provided that several assumptions are made. This framework is generally more flexible and more robust than its predecessors.

A pitfall of the original Stixel World framework is that it requires a disparity signal of a certain quality. However, the quality of disparity estimation often degrades in cases of occlusions, reflections or image regions with too little texture information. Unfortunately, such degradations are common in traffic scenery. As a result of this degraded signal, the original Stixel World framework detects many false obstacles, rendering the results useless for a practical system under adverse conditions. An example of this is shown at the top-left image of Fig. 1. In our recent work [7], we show that the performance of such a disparity-based framework can be improved by fusing color into the algorithm. This strategy resolves many erroneous results of the disparity analysis at a low additional computational cost, in contrast to alternative solutions such as high-quality cameras or more advanced disparity estimation techniques.

A key property of any ADAS is the response time, i.e. the time delay between data acquisition and the response to the result of the analysis. Since ADAS preferably need to function at high vehicle speeds, the response time of such systems should be as fast as possible. Hence, any delay that can be removed from the critical path of the analysis is beneficial to the value and applicability of the system, provided that it does not degrade the reliability of the results. Therefore, we will explore the possibility of removing the disparity analysis from the critical system path. Although fast disparity estimation methods exist [4][8], they typically either rely on sub-optimal algorithms processing at a low resolution, or they are based on customized hardware that is not commonly available. To illustrate this, even in the state-of-the-art system presented in [1], the dedicated FPGA disparity estimation takes 40 ms per frame, whereas the stixel analysis of the data takes 30 ms, when executed on a general, high-quality multi-core CPU.

For these reasons, we will not rely on a *strong fusion* of disparity and color in this work, even though the result presented in [7] clearly shows the qualitative benefits of that. In contrast, we propose here to process the most recent camera frame using an efficient color-only stixel segmentation. The disparity estimation and analysis, which is only required for our online color modeling, can be processed in parallel and at a lower frame rate. Two examples of our novel color-only stixel segmentation are shown at the bottom-left and bottom-right image of Fig. 1, illustrating that we can achieve better results than the state-of-the-art disparity approaches, even with color modeling of a low complexity.

An alternative to online color modeling is offline color modeling [9], which would completely remove the need for online disparity estimation. However, we have a strong preference for an online learning approach, given the challenging nature of traffic environments, which is full of varying weather conditions, complex scenery, varying geographical settings and highly dependent on the time of the day. For instance, in low-light situations, urban traffic scenes tend to contain predominantly gray-tones. We consider it more feasible to build a robust, yet discriminating color model that is tuned to that specific time and place, rather than building a generic model that holds for every environment and weather condition.

The remainder of this paper is structured as follows. First, we will describe the probabilistic Stixel World framework in Section II and explain briefly how it can be used with disparity, color or both data signals. Then, in Section III, we present the aspects of the system that will be evaluated in this paper. We then describe our validation method and the corresponding results in Sections IV and V, respectively. Lastly, conclusions are provided in Section VI.

II. THE STIXEL WORLD

Let us now give a short overview of the general Stixel World framework from [1], which we have used as a basis of our work. The main goal of stixel segmentation is to find the optimal labeling L^* of vertically stacked, piecewise planar ground or obstacle segments for input data $\mathbb{D}$, which can be any signal modality. Finding L^* can be formulated as a MAP estimation problem, as in

$$L^* = \arg\max_{L \in \mathbb{L}} P(L|\mathbb{D}), \tag{1}$$

which can be solved efficiently using Dynamic Programming. Using Bayes' theorem and assuming, among others, independence between columns and between data measurements at individual pixels, the posterior probability can be written as a chain of conditional probabilities by

$$P(L|\mathbb{D}) \sim \prod_{u=0}^{w-1} P(D_u|L_u) \cdot P(L_u). \tag{2}$$

Here, u is the column index and w the image width. The probability $P(L_u)$ models *a-priori* world knowledge constraining the labeling, to avoid dispensable segments and physically unlikely situations. This world model offers a way to regularize the results for image-column optimality, whereas the methods of [5] and [6] potentially lead to sub-optimal results, since they mostly analyze data locally. The details concerning $P(L)$ are presented in [1]. Finally, the likelihood of the data given a certain labeling, can be written as

$$P(D_u|L_u) \sim \prod_{n=1}^{N_u} \prod_{v=v_n^b}^{v_n^t} P(d_v|s_n, v), \tag{3}$$

where n is the segment index, N_u the number of segments in L_u, and v_n^b and v_n^t the bottom and top row-index of segment s_n, respectively. This segment has a label $l_n \in \{g, o\}$, representing the ground and obstacle classes, respectively.

The distribution $P(d_v|s_n, v)$ in Eq. (3) represents the probability of a single valid data measurement d_v at a certain row v, assuming that it would belong to a potential segment s_n. The model for $P(d_v|s_n, v)$ should reflect the nature of the employed signal modalities. There are several relevant approaches in literature. The authors of [1] employed a dense stereo-disparity signal as the sole data modality. They proposed to model $P(d_v|s_n, v)$ as a mixture model, containing a uniform distribution that models outliers and a Gaussian distribution that models inliers, to assess how well the measurement fits the potential segment for each class. For ground segments, the expected disparity is a linear planar surface and for obstacle segments a fronto-parallel surface.

A different approach is presented in [7], where the Stixel framework is extended, such that it exploits both the disparity signal and the color data. This strategy increases the robustness of the system against adverse conditions such as low light, bad weather, or a low-quality sensing system. To this end, the authors redefine the likelihood term of Eq. (3) to be $P(D_u, C_u|L_u)$ with the additional term $P(c_v|l_n)$ in the right-hand side of the equation, thereby treating color

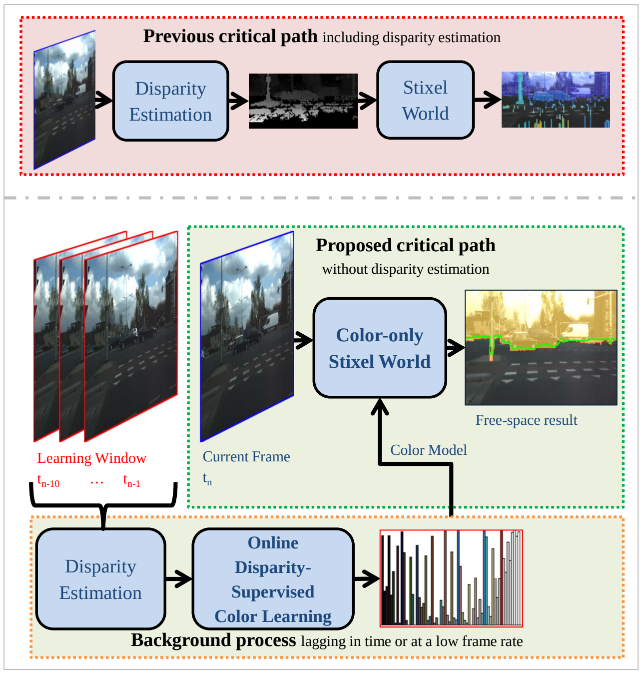

Fig. 2. A comparison between the existing and our proposed framework. Previous systems [1][7] require disparity estimation in their critical path, in contrast to our proposed system. Note that our disparity-supervised color modeling in the lower part of the scheme can be lagging or run at a lower frame rate than the critical path, by varying the range of the learning window.

and disparity as independent signals. Additionally, note that this color posterior is assumed to be independent of the segment location, since it only considers the class label l_n in contrast with disparity modeling. The color posterior $P(c_v|l_n)$ is learned in an online fashion using the labeling of several preceding frames as training masks for both the ground and obstacle classes. With these masks, normalized color histograms are calculated, which are then transformed to posteriors using Bayes' rule.

In our subsequent research, we explore the feasibility of segmenting the traffic scene images using a color-only version of the Stixel World algorithm [10]. The benefit of this approach is that the disparity estimation can be removed from the critical system path, as illustrated in Fig. 2. In [10], we present color modeling that is more suited for stand-alone processing compared to [7], by making it distance-aware. To this end, we have specified the color-based likelihood as

$$P(C_u|L_u) \sim \prod_{n=1}^{N_u} \prod_{v=v_n^b}^{v_n^t} P(c_v|s_n, v). \tag{4}$$

The distance-aware color processing consists of two aspects: (a) weighing each pixel with its corresponding real-world surface during the process of calculating the color histograms, and (b) leveraging the regular and the distance-weighted color posteriors based on v while evaluating Eq. (4). This approach leads to a more balanced color analysis of far-away and close-by image regions to cope with the inherent geometric distortion of cameras in a robust way [10].

III. ONLINE COLOR MODELING EXPERIMENTS

The key contribution here is to perform an elaborate analysis on the critical design choices of the online, distance-aware, self-supervised learning framework, as presented in [10]. The framework processes preceding stereo frames and generates a free-space vs. obstacle labeling based on disparity. Consecutively, this labeling is exploited as self-supervised training masks for the color representation of these two classes. The relevant design choices concern the color representation, consisting of preprocessing and color space selection, and the selection of the frames in the learning window.

A. Color Representation

Several aspects of the color representation are kept constant throughout this paper. First, we employ the median-cut algorithm on the frames in the learning window [11]. This ensures that we have an adaptive color representation that has both a sufficiently low complexity for fast processing and is still suitable for the current traffic scene, as the color reduction is performed online. Second, we perform a further reduction of the data by employing stixels that span 11 image columns. This increases the robustness and decreases the computational load at the cost of horizontal resolution in the labeling. To condense the image data into a single stixel-data vector, we calculate the first and the second mode of an 11×11 pixel window in the color data, horizontally centered at the central image column of the corresponding stixel. These aspects are recommended approaches as presented in [7] and [10]. Let us now briefly describe the evaluated color settings.

1) HEQ: We test the added value of performing Histogram Equalization on the raw RGB images (separately on each color plane) prior to converting it to a different color space.

2) RGB: We employ RGB as the full-color reference color space.

3) HS: To increase the robustness against varying lighting conditions, we test the strength of the Hue and Saturation dimensions of the HSV color space in our proposed framework.

4) IllumInv: The Illuminant Invariant color space presented in [12] is a more elaborate method for robust handling of changing lighting conditions and even shadows. It requires an automated offline camera-calibration method to find a parameter θ, which can then be used to transform each new image into an illuminant-invariant gray-scale image. We have adopted the proposed robust entropy-based calibration method and found that $\theta = 90 \pm 0.5°$ for our camera, but refer explicitly to [12] for more details on this color space and calibration method.

5) Gray: We also execute our segmentation on a gray-scale representation as a baseline for extreme cases of monochrome lighting conditions. Moreover, it would significantly reduce the constraints on the camera hardware and the corresponding data bandwidth when the gray-scale analysis is successful.

B. Learning Window

As described earlier, our method exploits preceding frames (which are analyzed based on their disparity signal) as Learning Window (LW) to construct color models of the obstacle and ground classes. The settings of a LW are the oldest frame, counting backwards from the current one, the step size and the final frame to be considered, annotated 'LW start:step:end'. The recommended approach in [7] relies on the 10 most recent preceding frames (LW10:1:1), which we will use as a reference setting. Since the aim of this work is to reduce the computational complexity and, most importantly, the system latency, we experiment with two alternative frame selections. First, we test a learning window without the two most recent frames (LW10:1:3). This way, the color model lags by two frames but it allows a longer data investigation interval before it is required for the analysis. Second, we test the more extreme case that has a lower frame rate and a lag, by only considering frames $t-9$, $t-6$ and $t-3$ (LW9:3:3).

Analogous to [10], we generate class-posterior color distributions from the labeled pixels in the LW frames and apply distance-aware weighting to correct the geometric distortion in the imaging process.

IV. EXPERIMENTS

To evaluate the different design settings, we employ two publicly available stereo RGB data sets with 188 annotated test frames in total [7][10]. The data consists of a large variety of relevant traffic situations under both good and adverse imaging conditions, such as dark roads with cyclists and cars, road repair sites, highway scenes etc. Both images with bright weather and under dim, clouded or even rainy conditions are present, leading to many low-contrast regions that are especially difficult for disparity-based methods. All frames are captured with a BumbleBee2 stereo camera (baseline: 12 cm; resolution: 1024×768 pixels; frame rate: 20 Hz), which is a relatively basic, low-cost camera when compared to several high-end or custom models used in other set-ups [1][13]. The details of our employed SGBM disparity estimation [14] and several improvements that we made for the baseline system are provided in [7] and [10]. Note that we cannot execute our algorithm on benchmarks such as the KITTI dataset [15], since those, unfortunately, do not contain the required preceding frames of annotated road images.

V. RESULTS

We have tested all combinations of the selected color and learning window settings, resulting in 24 different executions (runs). The effect of the individual color and learning window settings is shown in Fig. 3, by means of a box plot. Using a paired t-test, applying *HEQ* provides a significant improvement over not using equalization ($p = 3.04 \times 10^{-8}$). Likewise, the *RGB* color space outperforms the *HS*, *IllumInv* and *Gray* representations ($p = 1.12 \times 10^{-18}$, $p = 1.93 \times 10^{-4}$ and $p = 1.85 \times 10^{-29}$, respectively), and a full learning window (LW10:1:1) is better than a shorter, lagging one (LW10:1:3) with $p = 3.25 \times 10^{-3}$. No significant difference

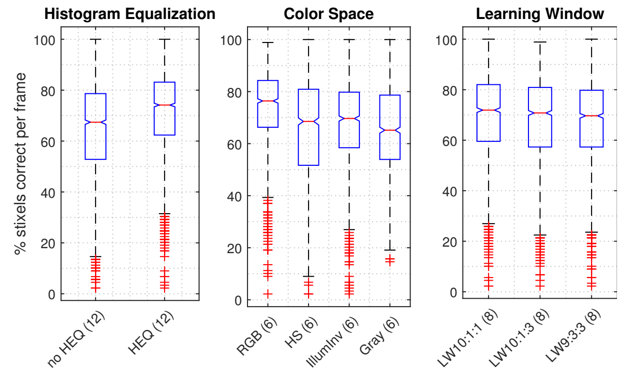

Fig. 3. Box plots comparing the different settings over all runs. For each frame in each run, the percentage of stixels with a correct free-space estimate is calculated, which is visualized as a box plot per setting. Hence, each box contains 188 data points per run. The number of runs per box is denoted in brackets in each label.

was found between the results of using the full or low frame-rate learning window (LW9:3:3) ($p = 8.33 \times 10^{-1}$).

Additional quantitative results are provided in Fig. 5. In this figure, all stixels over all frames are evaluated together for each run individually. For each stixel, a free-space evaluation is performed, by comparing the detected free space by the true free space, generated from the ground-truth annotations. We calculate the deviation as a percentage of the true free space. For robustness reasons, free-space detections are counted as correct when they are within the range of 30% too short or 15% too long. This asymmetrical range reflects the fact that missing an obstacle is more dangerous than detecting one too close. For the same reason, we distinguish the incorrect stixels into obstacle misses (free space is too long) and false obstacle detections (free space too short). Although a deviation of 30% may seem a large fraction, it corresponds to only a few pixels after several meters and only some centimeters before that. The results are shown on the combined data as well as on the individual datasets. The rightmost graph in Fig. 5 clearly shows that the added value of color processing is more pronounced for the EHV-road-ITSC15 data. This can be explained by the fact that the EHV-road-ITSC14 contains both frames with bright and dim lighting conditions, whereas EHV-road-ITSC15 is solely focused on dark, clouded, low-light and rainy frames. These situations are specifically difficult for disparity-based methods, rendering color data more advantageous. Of all color settings, Run f results in the highest percentage of correctly detected free-space (77.64%, averaged over all data), which is similar to the disparity-only method (77.25%). For the EHV-road-ITSC15 data, the improvement is higher: 78.01% compared to 74.39%. When specifically focusing on reducing the number of missed obstacles in difficult imaging conditions, Run h reduces the percentage of erroneous stixels from 17.18% to 12.85%, compared to the disparity-only method. On the combined data, the stixel-error fraction reduces from 13.81% to 11.52%.

We provide additional analysis by means of five theoretical

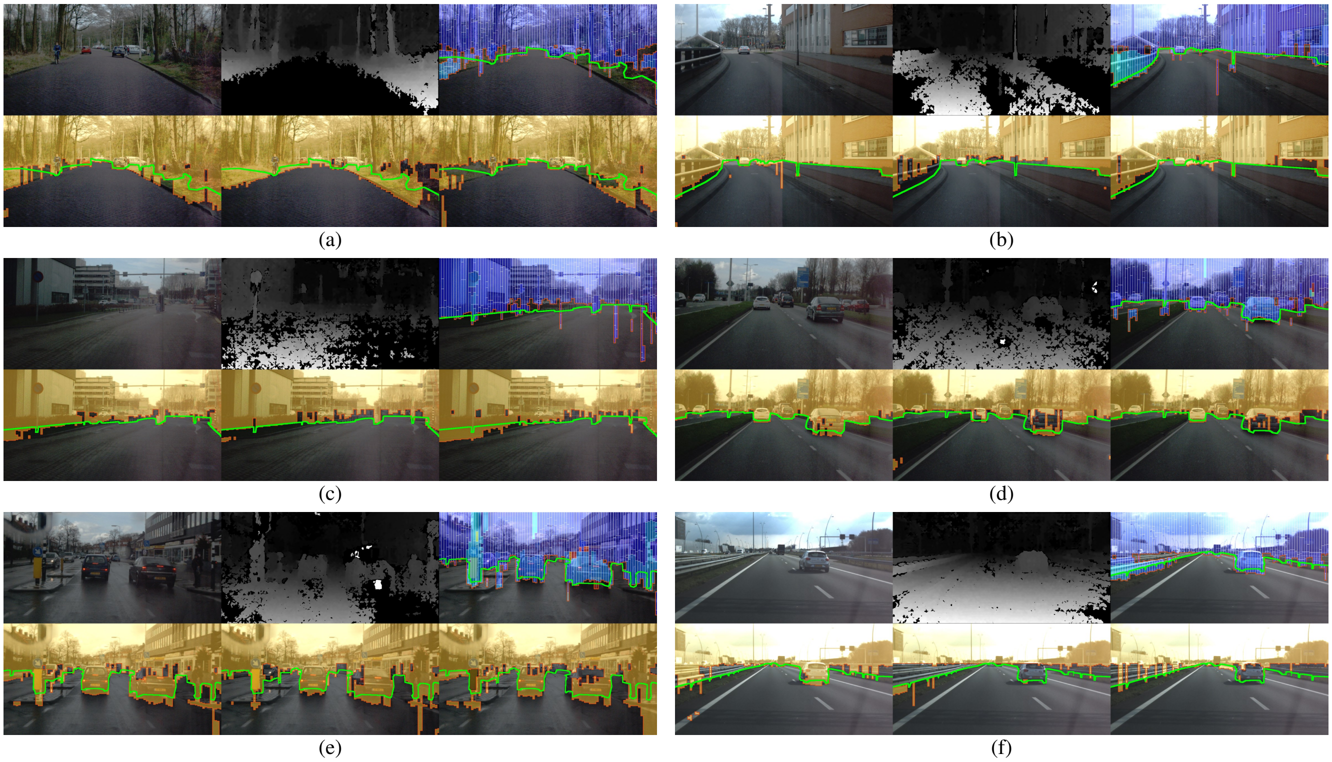

Fig. 4. Qualitative results of the disparity method and three of our runs on six stereo frames. Each subfigure contains two input and four result images. The top left and middle images show the rectified and cropped left camera image and the corresponding disparity image. Next to that, the disparity baseline result is shown, where the stixel segments are colored by their distance (red (close) to blue (far)). The three bottom result images illustrate different color settings, from left to right: RGB+HEQ with LW10:1:1, IllumInv+HEQ with LW10:1:1 and Gray+HEQ with LW9:3:3. In the color-only results, a homogeneous overlay of the detected obstacle region is visualized. The bright green line indicates the border of the ground-truth annotation of the drivable surface. Subfigures a, c and e show that our color-only results provide similar or better results in various situations. The right column (b, d and f) shows examples of scenes where not all color settings provide equally acceptable results.

runs at the bottom of Fig. 5. These scores are generated by selecting the optimal setting for each frame out of a (subset of) the available runs, to assess the added value of the processing choices and to provide insights in where the most gain is to be expected in future research. First of all, it is noteworthy that for *every* setting, there are frames in the data set on which it performs best. If the optimal score is selected from all possible runs (including disparity), the highest theoretical score can be achieved (86% correct), as could be expected. However, also with the color data alone there is room for improvement, compared to using the same color space and preprocessing step for every frame. So, even with our adaptive median-cut color indexing, the system can extract more information from different color representations in different situations (fourth bar from below in Fig. 5; 83% correct). Also, note that even with the simplest learning window (LW9:3:3), the color-only Stixel World can outperform the disparity one with a more sophisticated color representation (the bottom bar in Fig. 5; 80% correct), even though using more frames is still better (third bar from below in Fig. 5; 82% correct).

The aforementioned observations are illustrated with the qualitative results in Fig. 4, where the disparity-only results are compared to three of our color-only strategies. We show the setting that performed best (RGB+HEQ, LW10:1:1), one of the runs that relied on the color space that was specifically designed for this context (IllumInv+HEQ, LW10:1:1), and the results with the lowest computational complexity, since it uses gray-scale images and only three LW frames (GRAY+HEQ, LW9:3:3). The left three images show that our methods are all capable of delivering similar or better results than the disparity-only framework. The images in the right column of Fig. 4 illustrate that different settings perform best in different situations, so that the system performance could be increased by adapting the color modeling in even more ways than we currently do. For example, color spaces may be combined or selected online, or the most informative frames within the learning window could be selected adaptively. Metrics and methods guiding this online decision-process will be investigated in future research.

VI. CONCLUSIONS

We have explored a stixel-based probabilistic framework for color-based free-space vs. obstacle segmentation. Our system learns color appearance models for free-space and obstacle classes in an online and self-supervised fashion. To this end, it applies a disparity-based segmentation, which can run in the background of the critical system path and at a lower frame rate than the color-based algorithm. As a bonus, this approach enables operation without a real-time

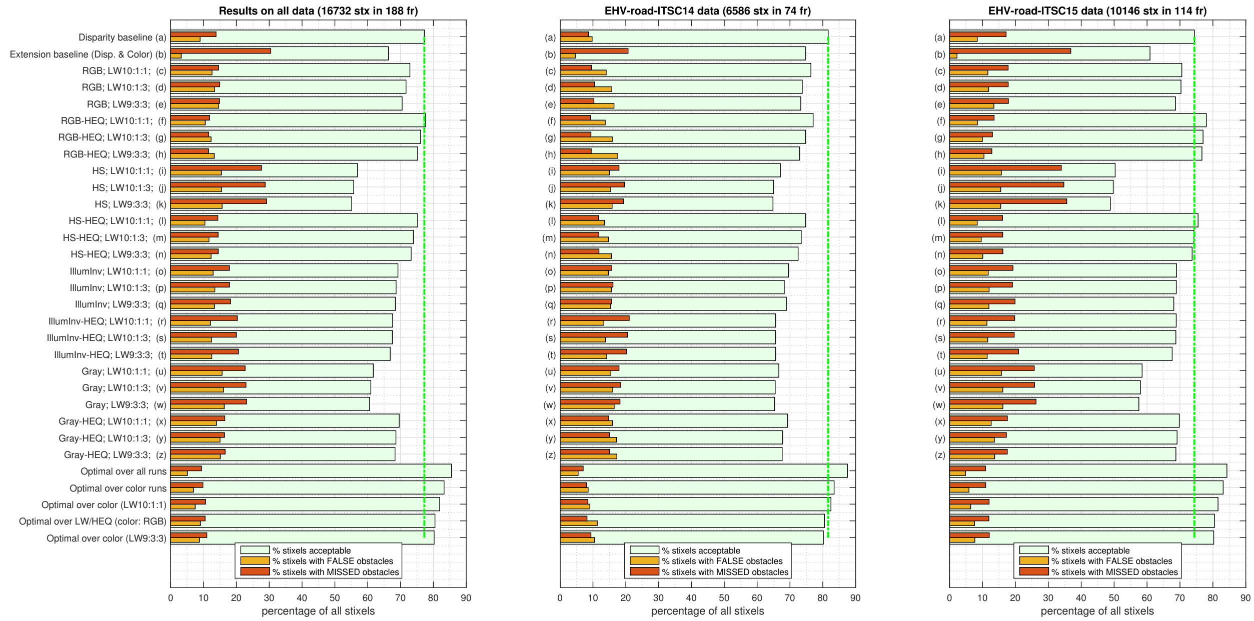

Fig. 5. Quantitative results of free-space segmentation for the baseline that use disparity alone (a) [1] or strongly fused color and disparity (b) [7], and all of our runs. The labels show if histogram equalization (HEQ) is applied, which color space is used and which frames are in the learning window (LW). The learning window parameters are indicated with LW start:step:end as before. The final four entries are theoretical optima, generated by selecting the optimal setting for each frame out of all runs (a-z, including disparity), out of all color-only runs alone (c-z), out of all runs with LW10:1:1 (c, f, ..., x), out of all runs with RGB (c-h) and out of all runs with LW9:3:3 (e, h, ..., z). They give an approximate upper bound for the current processing framework, which could be achieved by, e.g., online selection of color spaces.

disparity estimate. Consequently, the current road scene can be analyzed without the extra latency of disparity estimation. This feature results into a reduced response time from data acquisition to data analysis, which is a critical property for high-speed ADAS.

To achieve reliable color-only free-space detection, we have experimented with several color spaces and different online learning settings. Our evaluation on publicly available data shows that the color-based analysis can achieve similar or even better results in difficult imaging conditions, compared to the state-of-the-art disparity-only method. As an illustrative example, our color-processing detects the correct free-space for 77.6% of all stixels, compared to the disparity-only score of 77.3%. Furthermore, our color-only method results in 4.3% less stixels with missed obstacles on the most challenging data set.

Besides the previous system aspects, the provided meta-analysis of the results shows that our approach of online color modeling is beneficial and can be extended for further improvements, with potential scores of up to 82% within the currently assessed parameter-setting space.

REFERENCES

[1] D. Pfeiffer, "The Stixel World," Ph.D. dissertation, Humboldt-Universitat Berlin, 2011.

[2] S. Thrun and M. Montemerlo, "Stanley: The robot that won the DARPA Grand Challenge," *J. of Field Robotics*, vol. 23, pp. 661–692, 2006.

[3] C. Urmson, C. Baker, J. Dolan, P. Rybski, B. Salesky, W. Whittaker, D. Ferguson, and M. Darms, "Autonomous driving in urban environments: Boss and the Urban Challenge," *AI magazine*, vol. 30, pp. 17–28, 2008.

[4] W. Van Der Mark and D. M. Gavrila, "Real-time dense stereo for intelligent vehicles," *IEEE Trans. on Intelligent Transportation Systems (TITS)*, vol. 7, no. 1, pp. 38–50, 2006.

[5] G. Dubbelman, W. van der Mark, J. C. van den Heuvel, and F. C. A. Groen, "Obstacle detection during day and night conditions using stereo vision," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*. IEEE, Oct. 2007, pp. 109–116.

[6] R. Labayrade, D. Aubert, and J.-P. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through "v-disparity" representation," *IEEE Intelligent Vehicle Symp. (IV)*, vol. 2, pp. 646–651, 2002.

[7] W. P. Sanberg, G. Dubbelman, and P. H. de With, "Extending the stixel world with online self-supervised color modeling for road-versus-obstacle segmentation," in *IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2014.

[8] S. K. Gehrig, F. Eberli, and T. Meyer, "A Real-Time Low-Power Stereo Vision Engine Using Semi-Global Matching," in *Int. Conf. on Computer Vision Systems (ICVS)*, vol. 5815 LNCS, 2009, pp. 134–143.

[9] J. M. Álvarez, M. Salzmann, and N. Barnes, "Learning appearance models for road detection," *IEEE Intelligent Vehicles Symp. (IV)*, pp. 423–429, June 2013.

[10] W. P. Sanberg, G. Dubbelman, and P. H. de With, "Color-based free-space segmentation using online disparity-supervised learning," in *IEEE Con. on Intelligent Transportation Systems (ITSC) [Accepted for Publication]*, 2015.

[11] P. Heckbert, "Color image quantization for frame buffer display," *Computer Graphics*, vol. 16, no. 3, pp. 297–307, 1982.

[12] J. M. Álvarez and A. M. López, "Road detection based on illuminant invariance," *IEEE Trans. on Intelligent Transportation Systems (TITS)*, vol. 12, no. 1, pp. 184–193, 2011.

[13] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets Robotics: The KITTI Dataset," *Int. J. of Robotics Research (IJRR)*, no. October, pp. 1–6, 2013.

[14] H. Hirschmüller, "Stereo processing by semiglobal matching and mutual information," *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 30, no. 2, pp. 328–41, Feb. 2008.

[15] J. Fritsch, K. Tobias, and A. Geiger, "A New Performance Measure and Evaluation Benchmark for Road Detection Algorithms," in *IEEE Conf. on Intelligent Transportation Systems (ITSC)*. IEEE, 2013.