A Compact Planar-patch Descriptor based on Color

Eduardo Fernández-Moral, Javier González-Jiménez and Vicente Arévalo
Mapir Group, Universidad de Málaga, E.T.S. Ingeniería de Informática-Telecomunicación, 29071, Málaga, Spain

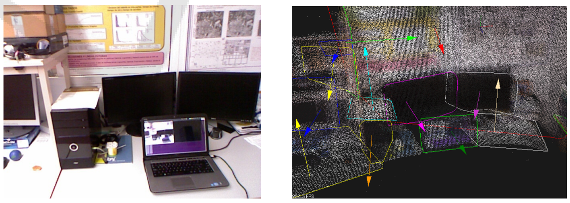
Keywords: Planar-patches Matching, Color Descriptor, RGBD Sensors, Plane-based Mapping, Place Recognition, Loop Closure.

Abstract: The representation of the world upon planar patches has proven to be simple, robust and useful for a variety of robotic tasks, including SLAM, autonomous navigation, or scene recognition. In this work we investigate how to incorporate color information into such representation to improve the matching of planar patches while maintaining the model compactness, which is essential for real-time applications. We propose a descriptor based on the dominant color of the patch, which is defined as the center of the biggest cluster in the patch histogram. In the paper, different color spaces and methods for extracting the dominant color are analyzed. We compare this descriptor with a recent proposal (saturated hue based histogram) and provide some conclusions on the trade-off between their descriptiveness and compactness. Finally, we present experimental results showing how our color descriptor can be exploited to increase the efficiency of both: plane-based place recognition and planar patch categorization.

1 INTRODUCTION

The combination of color and depth information provided by RGBD sensors (like Kinect) has demonstrated to be useful for a wide range of problems as 3D modeling, object recognition or Simultaneous Localization and Mapping (SLAM), among others. One of the main challenges when using RGBD sensors in real-time robotics applications is how to effectively process and represent the overwhelming flow of data they deliver. A suitable way of accomplishing that is by extracting a structure of planar patches from it. This strategy has already been applied in problems such as visual odometry (Martinez-Carranza and Calway, 2012), augmented reality (Chekhlov et al., 2007), SLAM (Weingarten and Siegwart, 2006) or place recognition (Fernández-Moral et al., 2013).

Planar patches, or planes for short, can be efficiently computed from depth images (Poppinga et al., 2008), (Holz and Behnke, 2013), and can be described by very simple geometric features, like the normal vector and the convex hull (see fig. 1). The question that we pose here is how to represent the radiometric information of the planar patches while maintaining the compactness of the representation. The answer to this question depends clearly on the application. In this work, the context of our research is that of matching planar patches for real-time place recognition, what involves extensive search for patch correspondences. Thus, selecting a color descriptor involves the non-trivial issue of maintaining a trade-off between distinctiveness, compactness, and computational cost.



a) b)

Figure 1: Plane-based representation. a) RGB image of the scene. b) Point cloud representation with the segmented planar patches superimposed.

This problem of finding a color descriptor for planar patches was posed recently in (Pathak et al., 2012) in the context of registering 3D scans, where the authors adopted a hue based histogram to improve the efficiency of registration. In this paper, in contrast, we explore the idea of finding a descriptor based on the dominant color of the plane. This insight, though intuitive, has not been sufficiently studied in the literature. This problem has interest for the research community for several reasons: first, most planes in indoor environments do have a dominant color; second, the dominant color is more robust to the partial observation of planes than histograms; and finally, the lack

Fernández-Moral E., González-Jiménez J. and Arévalo V..
A Compact Planar-patch Descriptor based on Color.
DOI: 10.5220/0005015102960302
In *Proceedings of the 11th International Conference on Informatics in Control, Automation and Robotics* (ICINCO-2014), pages 296-302
ISBN: 978-989-758-040-6

of distintiveness is largely compensated by the high efficiency required by online back-end processes for classification or place recognition which benefit from a more compact, fast to compare descriptor.

In order to find such compact descriptor, we study different color spaces and radiometric features, looking for invariance to illumination, point of view and partial occlusion (section 2). We conclude that a color descriptor based on the patch dominant color in normalized *RGB* space provides the best balance between distinctiveness and compactness. In section 3, this descriptor is compared with the hue based histogram reported in (Pathak et al., 2012), which was previously proposed for a similar problem, showing that though the distinctiveness of both are simmilar, our alternative is advantageous in terms of computation time.

We provide experimental results for place recognition with plane-based maps using the color descriptor proposed in this work. We demonstrate that the efficiency of the previous geometry-based solution is significantly improved (around 6 times faster) by using color information. In all the experiments, we compare our results with with the hue-based histogram.

2 SELECTING A COMPACT COLOR DESCRIPTOR FOR PLANAR PATCHES

In this section we address the problem of finding the simplest color descriptor for a planar patch focused on the problem of patch matching. This descriptor must be highly invariant to viewpoint, lighting conditions and partial occlusion, and also, it must be efficiently calculated. Note that the utility of this descriptor is not to unequivocally identify planar patches, but to prune the search space by adding a very compact radiometric information to the geometric features of the planar model.

In the context of matching planar patches, a common solution is that of maximizing the photoconsistency between them (Argiles et al., 2011). The main limitation of this strategy, which comes as a consequence of the lack of compactness and invariance of the descriptor, is that maximizing the photoconsistency is prohibitively expensive for many applications, especially when there is not a good initial estimation of the registration (e.g. loop closure detection). Closer to our work are those that describe the patch radiometric information through its histogram (Hafner et al., 1995), (Swain and Ballard, 1991). In this line, (Pathak et al., 2012) posed recently the problem that we address in this paper: showing how color information can be exploited to increase the efficiency of 3D scan registration. A well illuminated scene is assumed in that work, where the authors adopt a hue based histogram with 2 extra bins to keep intensity saturated values (black and white), and test different measures for histogram distance. However, they do not take into account the fact that many planar patches have a single color, so that the histogram contains redundant information. Also, this descriptor is not robust to partial occlusion, which is rather common when doing exploration and mapping.

In this paper, we propose to describe the patch with its dominant color. A similar strategy is used in video compression (Manjunath et al., 2001) to define blobs having the same color. In this way the descriptor storage and the computation of distances are reduced to a minimum. This is important in a number of problems where many match combinations have to be checked in real-time. In order to select such a descriptor we need to address some issues: first, the selection of the color space which offers the best suitability to obtain an invariant and distinctive dominant color (subsection 2.1); second, to define the way this dominant color is extracted (subsection 2.2); and third, to adapt the descriptor for cases where the dominant color is not reliable enough (subsection 2.3).

2.1 Selection of the Color Space

In order to obtain a distinctive dominant color, the histograms of the patches must be invariant to illumination conditions, shading and viewpoint. These characteristics are highly dependent on the color space used to represent the radiometric information, as we show in the analysis below. Note also that the fact of selecting the dominant color makes the descriptor inherently robust to partial occlusion when the physical plane has a clearly defined dominant color, which is the most common situation. If this is not the case, e.g. a textured plane with different colors, the dominant color is not a good descriptor and it should not be used for matching.

Different color spaces have been studied in the context of object recognition in (Gevers and Smeulders, 1999). This work concludes that normalized *RGB* (*rgb*), saturation and hue (*HS*), and the color models $c_1c_2c_3$ and $l_1l_2l_3$ are highly invariant to changes in viewing direction and illumination (see table 1 for the formulation of these color spaces). Below, we analyze these color spaces for a dataset containing 1000 observations of plane surfaces from different scenarios, spanning diverse viewing conditions (changing viewpoint and illumination, partial occlusion, etc.). Below we study some relevant properties

Table 1: Formulation of several color spaces from the RGB data.

*C*olor space	*F*ormulation
rgb	$r(R,G,B)=\frac{R}{R+G+B}$ $g(R,G,B)=\frac{G}{R+G+B}$ $b(R,G,B)=\frac{B}{R+G+B}$
HS	$H(R,G,B)=\arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)+(R-B)}\right)$
	$S(R,G,B)=1-\frac{\min(R,G,B)}{R+G+B}$
$c_1c_2c_3$	$c_1(R,G,B)=\arctan\left(\frac{R}{\max(G,B)}\right)$ $c_2(R,G,B)=\arctan\left(\frac{G}{\max(R,B)}\right)$ $c_3(R,G,B)=\arctan\left(\frac{B}{\max(R,G)}\right)$
$l_1l_2l_3$	$l_1(R,G,B)=\frac{(R-G)^2}{(R-G)^2+(R-B)^2+(G-B)^2}$ $l_2(R,G,B)=\frac{(R-B)^2}{(R-G)^2+(R-B)^2+(G-B)^2}$ $l_3(R,G,B)=\frac{(G-B)^2}{(R-G)^2+(R-B)^2+(G-B)^2}$

of such color spaces:

2.1.1 Histogram Invariance

To extract a dominant color descriptor invariant to viewpoint (including the effects of partial occlusion and shades), the histograms main peak must be stable along different views of the same plane. To measure the histogram stability in a given color space, we check the similarity of all histograms corresponding to the same plane by means of a chi-squared (χ^2) distance measure (Pele and Werman, 2010). This measure is used to compute the histogram distances of all pairs of views of the same plane. Then, the mean distance of all analyzed pairs is averaged for all tested planes to obtain a global measure of the color space stability, see table 2.

2.1.2 Histogram Dispersion

The histograms of planes with a well defined dominant color are generally unimodal and with little dispersion. However, such characteristic does not apply to all the planes in the environment, and also, it varies depending on the color space. To accept that a plane has a dominant color we make use of a simple heuristics which requires that at least 50% of the patch pixels are contained in a bandwidth of $\pm 5\%$ of the histogram range, centered at such dominant color. Thus, we define the concentration rate C as the number of planes that fulfills this condition in all color components between the total of planes. We have found that the above condition is fulfilled in 97.5% for planes represented with *rgb* and 92.8% for planes represented with $c_1c_2c_3$, while the other color spaces

Table 2: Suitability of different color spaces to represent planar patches according to: histogram stability, histogram dispersion and computation time. The values shown correspond to the average of 100 different planes, with 10 observations each. For all properties, less means better.

	rgb	$c_1c_2c_3$	$l_1l_2l_3$	*HS*
Stability χ^2	0.10	0.11	0.13	0.14
Disp $(1-C)$	0.03	0.07	0.74	0.77
C. time (μs)	10.7	104.9	23.0	11.3

present much lower rates. Table 2 shows the dispersion rate in this experiment, defined as $(1-C)$.

2.1.3 Computation Time

Another important criterion to consider is the computation time required to transform the original color space to the target one. This is less critical because this cost is small in comparison with the whole process of segmenting the planes, whichever the chosen color space is. The average of this time for this dataset is also indicated in table 2.

Taking into account the criteria studied above, we notice that *rgb* is the one with the best properties, and therefore, it is the one adopted in the rest of this paper.

2.2 Computing the Dominant Color of a Plane

There exist several ways to define the dominant color for a planar patch. In this work we have tested the mode of the histograms, and the centroid of the largest cluster extracted with two variants of the mean shift algorithm: with fixed (*FMS*) and variable bandwidth (*VMS*), respectively. Mean shift has been broadly used for color segmentation (Comaniciu and Meer, 1997). Though it has limitations for real-time applications due to its computational cost, in our case the cost of the mean shift is affordable since most histograms present unimodal distributions and we only extract one cluster, so that it converges in very few iterations.

We compare the distinctiveness of the dominant color obtained with the above techniques using a binary classifier based on the color difference of two patches, expressed as $|\mathbf{c}_i - \mathbf{c}_j|$. Thus, when this difference is larger than a threshold the patches are considered to belong to different physical planes. This classifier is tested, for a range of thresholds, with the previous dataset in which we know beforehand which observation corresponds to each plane (i.e. the ground truth).

From this experiment we obtain the distinctiveness of this classifier in terms of its sensitivity (ratio of actual positives which are correctly identified) and the

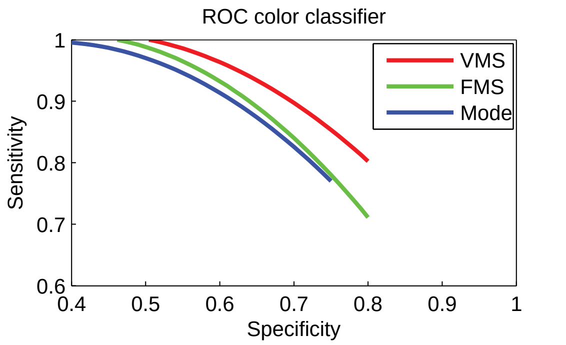


Figure 2: ROC curves (sensitivity *vs.* specificity) of the color constraints as binary classifiers.

specificity (ratio of negatives which are correctly rejected) for the different techniques to obtain the dominant color. These results are depicted as ROC (Receiver Operating Characteristic) curves in fig. 2. Every point of each curve represents a different threshold for the classifier, thus, more restrictive thresholds result in higher sensitivity and lower specificity. Note that the nearer the curve is to the optimum point (1,1) the better the classifier is. From this test we conclude that *VMS* provides the most distinctive dominant color since both, sensitivity and specificity, are higher than for the mode and *FMS* for any threshold.

2.3 Dealing with Non-distinctive Dominant Color

An important issue when describing patches with their dominant color is dealing with those cases where this description is not applicable. For example, we can not use the dominant color as descriptor for planes without a prevalent color (e.g. a checkerboard). To take this situation into account we add a boolean to our color descriptor to specify whether the distribution of the plane histogram in *rgb* has a low dispersion, as explained in the previous subsection.

Another case of interest is that of recognizing very distinct intensity levels (e.g. black and white planes have the same values $r = g = b = 0.33$). Despite this problem depends on illumination, we reckon that a minimum illumination is required to distinguish different colors, and such minimum can be sufficient to distinguish gray levels far apart when enough illumination is available. For this case, we propose to include in the descriptor the average intensity (Av.Int), which is calculated as the average $(R+G+B/3)$ of the inliers supporting the dominant color given by the previous mean shift segmentation. Thus, this part of the descriptor can be used for well illuminated scenes (what is common to many environments).

To sum up, the resulting descriptor contains 4 elements that are stored in a word of 4 bytes: 2 bytes for normalized color r and g (note that b depends on these two as $r+g+b=1$), 1 byte for the plane intensity and 1 bit for to specify the existence of a dominant color.

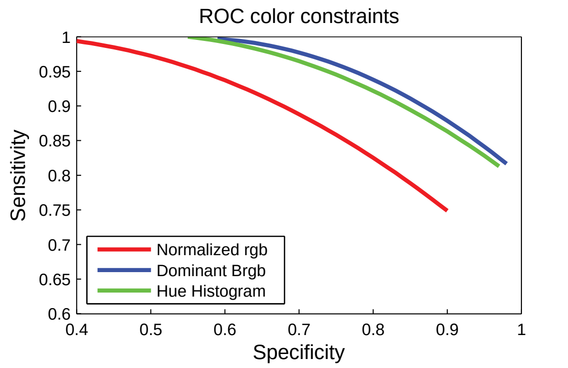


Figure 3: ROC curves (sensitivity *vs.* specificity) of different color descriptors: dominant *rgb*, dominant *rgb* including intensity and hue histogram.

3 DOMINANT COLOR DESCRIPTOR *vs.* HUE HISTOGRAM

In this section we evaluate the distinctiveness of the proposed dominant color based descriptor and compare it with the normalized, saturated hue histogram proposed in (Pathak et al., 2012). To compare both in the same conditions such evaluation is performed in well illuminated scenes, and so, the intensity part of our descriptor is also employed. Following the work of (Pathak et al., 2012), we implement the paper method with 74 bins, 72 bins for hue values and two more bins for saturated black and white, and compare the patch histograms h_i with the Bhattacharyya distance (Bhattacharyya, 1946):

$$B(\mathbf{h_1},\mathbf{h_2}) = \sqrt{1 - \sum_{k=1}^{N} \sqrt{\mathbf{h_1}[k] \cdot \mathbf{h_2}[k]}} \qquad (1)$$

The sensitivity and specificity of a binary classifier based on the compared descriptors are evaluated using different thresholds as we did in the previous section (see fig. 3). As expected, we observe that the proposed descriptor is significantly more distinctive than the *rgb* dominant color, since the latter lacks information about the reliability of the dominant color, and also it cannot distinguish between different grayscale levels.

By comparing our descriptor with the hue based histogram we observe that their distinctiveness are similar despite the richer information of the latter (see

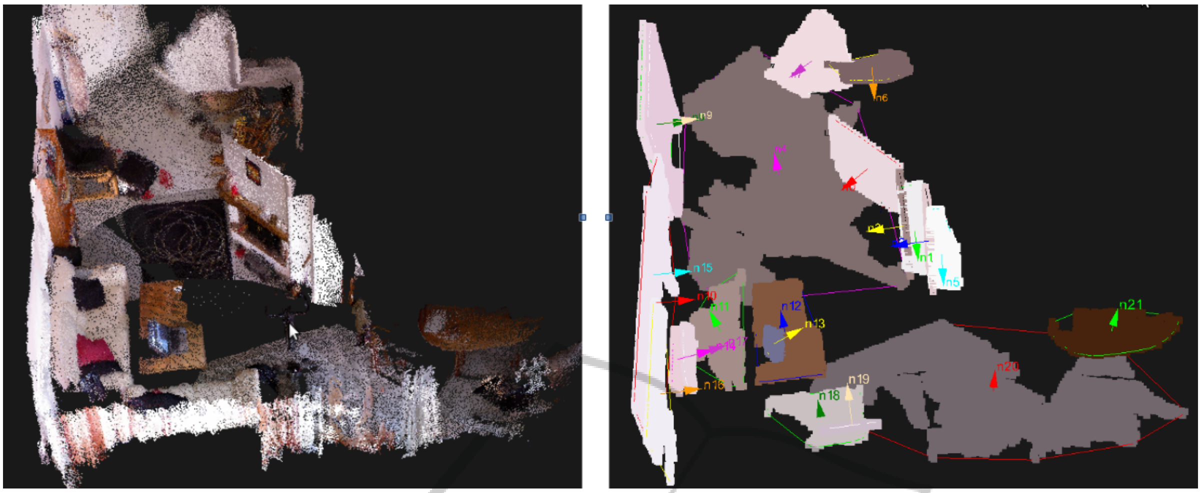


Figure 4: Plane based representation of a living room. The colored planes at the right have been extracted from the point cloud at the left.

fig. 3). The reason for this is that most planes have a prevalent color in our test environments. The fact that the sensitivity of the hue histogram is slightly lower is explained because the histogram is less robust to partial observation of planes. Contrarily, this descriptor should perform better for textured surfaces and when the patches present no occlusion, however, such cases are rare in the home and office environments we are working in, where our dominant color descriptor is more suitable.

Besides the distinctiveness of the descriptor, the compactness and the computational cost of measuring distances are two relevant issues in this work. The color descriptor we propose is the most compact representation for a number of applications, including patch matching or visualization (Pathak et al., 2012). Regarding the computation of distances between descriptors, in our case it is reduced to a subtraction in each element, while the computing distance between histograms (e.g. with the Bhattacharyya distance) requires more computation. This is an important advantage when we require to perform this operation intensively in real time as we will see in next section.

4 EXPERIMENTAL RESULTS

The experimental validation of the proposed color descriptor is presented next for two different problems. First we evaluate the improvement for recognizing previous visited places based on a planar description of the scene. Second, we analyze the advantages for planar patch classification using a random forest.

4.1 Improving Place Recognition

In this experiment, the proposed descriptor is tested in a plane-based place recognition application (Fernández-Moral et al., 2013) (available in http://www.mrpt.org/pbmap). This application works by matching sets of neighboring planar patches based on their geometric attributes and their relations. Here, we incorporate our dominant color based descriptor to speed up the search process. Figure 4 shows an example of such planar representation where each patch is colored with its dominant color. We demonstrate that the little expense of including the color descriptor improves significantly the matching of planes, and so, the place recognition performance.

In this application the planar patches are efficiently extracted at frame rate from RGBD images using a region growing technique (Holz and Behnke, 2013). Such planes are organized in a graph which stores the connection between close-by planar patches, so that a subgraph of connected nodes represents a local neighborhood of planes. Place recognition is addressed as a problem of matching neighborhoods of planes, thus, the poor distinctive information of a single plane is compensated by the strong relations between neighbor planes.

In order to match two sets of neighbor patches, an interpretation tree is utilized to apply geometric restrictions in the form of unary and binary constraints (Grimson, 1990). These constraints make use of thresholds that have been determined experimentally from training carried out in diverse indoor scenarios: office and home places. An important advantage of this approach is its flexibility to recognize places when the planes are partially observed or miss-

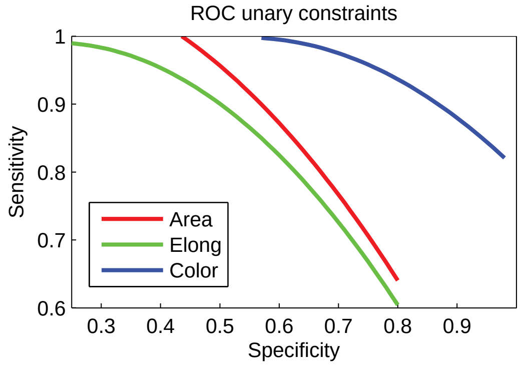


Figure 5: Comparison of the different unary constraints by their ROC curves (sensitivity *vs.* specificity).

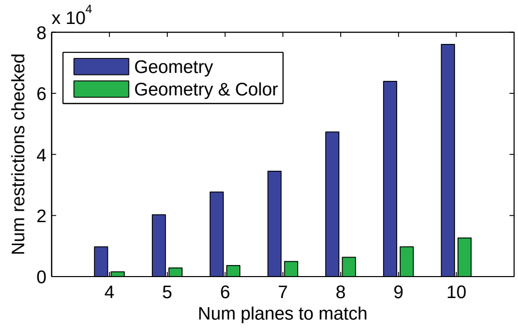


Figure 6: Performance of the place recognition process using both: only geometry and color and geometry in PbMaps.

ing. Yet, a planar-based scene representation has the benefit of being more robust to changes of viewpoint and tolerates reasonably well changes in the scene.

This solution performs intensive computation to search for previous places (loop closure) every time the map is updated at 30 Hz. By adding the color descriptor to this geometric description we can prune branches of the interpretation tree to find the solution more efficiently. Thus, the proposed color descriptor introduces a new unary constraint to avoid the matching of planes with different color. To arrange the order in which the new restriction will be evaluated, we compare its distinctiveness with the other two unary restrictions: area and elongation of the planar patches. Similarly to the previous section, we estimate the ROC curves showing the balance between sensitivity and specificity of the unary restrictions for a range of thresholds, see fig. 5. From this graph we can see that the color is the most discriminative feature and, since all unary restrictions require similar computation, we arrange their application order consequently: color, area and elongation.

To illustrate the performance improvement, we have carried out an experiment where we measure the time required for searching a place with and without the color descriptor. Figure 6 shows the average time of such search with respect to the number of planes being evaluated. We observe that performing the search using the proposed color descriptor is around 6 times faster. Such rate varies from 2 to 10 depending on the radiometric characteristics of the planar surfaces of the environment. This presents a significant increase of efficiency over the previous pure-geometric solution.

5 CONCLUSIONS

This paper presents a simple, highly compact color descriptor for planar patches which can be used to improve the performance of matching and classification algorithms based on planar models. In essence, we propose to approximate the patch color by the most representative color in it in normalized *rgb* space. Though this idea may look simplistic and naive, we notice that it has not been employed before, probably due to the difficulty of finding an invariant descriptor through such little information, and only tuned histograms have been employed in similar problems. To address the problem of finding an invariant color attribute we have evaluated different color spaces and different dominant color selection strategies, and have included some extra information about its reliability and its saturation. The proposed descriptor contains the dominant *rgb* color of the plane, extracted using a mean-shift algorithm, the average intensity of this dominant color and a boolean to indicate whether such dominant color is representative enough. We have verified in common office and home environments that this descriptor is as distinctive as the hue based histogram proposed previously, while it is more compact and faster to compare. We demonstrate that the performance of a place recognition application based on matching planar patches is significantly improved (about 6 times speed-up) by using the compact color descriptor proposed here. This descriptor has also been tested to improve planar patch classification by adding such information to a compact geometric descriptor, showing that the recognition rate improves at the expense of a small increase in computation.

REFERENCES

Argiles, A., Civera, J., and Montesano, L. (2011). Dense multi-planar scene estimation from a sparse set of images. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pages 4448–4454.

Bhattacharyya, A. (1946). On a measure of divergence between two multinomial populations. *Sankhyā: The Indian Journal of Statistics (1933-1960)*, 7(4):401–406.

Chekhlov, D., Gee, A., Calway, A., and Mayol-Cuevas, W. (2007). Ninja on a plane: Automatic discovery of physical planes for augmented reality using visual slam. In *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 1–4. IEEE Computer Society.

Comaniciu, D. and Meer, P. (1997). Robust analysis of feature spaces: color image segmentation. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pages 750 –755.

Fernández-Moral, E., Mayol-Cuevas, W., Arévalo, V., and González-Jiménez, J. (2013). Fast place recognition with plane-based maps. In *Robotics and Automation (ICRA), in 2013 IEEE International Conference on*. IEEE.

Gevers, T. and Smeulders, W. (1999). Color based object recognition. *Pattern recognition*, 32(3):453–464.

Grimson, W. E. L. (1990). *Object Recognition by Computer - The role of Geometric Constraints*. MIT Press, Cambridge, MA.

Hafner, J., Sawhney, H., Equitz, W., Flickner, M., and Niblack, W. (1995). Efficient color histogram indexing for quadratic form distance functions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(7):729 –736.

Holz, D. and Behnke, S. (2013). Fast range image segmentation and smoothing using approximate surface reconstruction and region growing. In *Intelligent Autonomous Systems 12*, pages 61–73. Springer.

Manjunath, B., Ohm, J.-R., Vasudevan, V., and Yamada, A. (2001). Color and texture descriptors. *Circuits and Systems for Video Technology, IEEE Transactions on*, 11(6):703 –715.

Martinez-Carranza, J. and Calway, A. (2012). Efficient visual odometry using a structure-driven temporal map. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 5210–5215.

Pathak, K., Vaskevicius, N., Bungiu, F., and Birk, A. (2012). Utilizing color information in 3d scan-registration using planar-patches matching. In *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2012 IEEE Conference on*, pages 371–376.

Pele, O. and Werman, M. (2010). The quadratic-chi histogram distance family. *Computer Vision–ECCV 2010*, pages 749–762.

Poppinga, J., Vaskevicius, N., Birk, A., and Pathak, K. (2008). Fast plane detection and polygonalization in noisy 3d range images. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 3378–3383. IEEE.

Swain, M. J. and Ballard, D. H. (1991). Color indexing. *International journal of computer vision*, 7(1):11–32.

Weingarten, J. and Siegwart, R. (2006). 3d slam using planar segments. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pages 3062–3067.