Chapter 8

Conclusions and Future Work

8.1 Conclusions

We dedicate the next lines to putting together all the conclusions that were drawn in each Chapter that conforms this Thesis. A global conclusion for the whole Thesis can be briefly summarized by saying that our aim was to attack the problem of segmenting color images coming from a mobile robot so that a posterior object recognition procedure could be carried out. Nevertheless, because of the varying nature of this kind of images, first we had to solve the problem of changing color.

These are basically the kind of tasks that were endeavoured in this Thesis, which could be further parceled into the following four stages, namely, theoretical study of the color problem, proposal of a color constancy algorithm and a color image segmentation algorithm, as well as some measures to compare images based on the segmentations thus obtained, which can also be used to show that color variation was reduced by means of color constancy, providing as a consequence more selective ways to compare images on the basis of their content.

8.1.1 State-of-the-Art Survey on Color Const. Methods

The main concern of the large number of approaches reviewed in Chapter 2 is that of recovering an estimate for the scene illuminant which could help to remove the effect of it on color images. This is a quite difficult problem and no general method that fully solves it has been proposed yet. However, there exist many partial solutions which are worth to be considered.

The most effective methods using synthetic images are those based on image statistics, such as color–by–correlation and neural networks. Some variants of the 3D gamut–mapping approach also work pretty well with the advantage of estimating the illuminant intensity. Similar results are found using real images as in the case of synthetic images. Concerning the job of object recognition and image indexing, results indicate that 3D gamut–mapping approaches outperformed others. However, that improvement is not enough to completely cope with all the light variation existing in real images. To attain better results it would be necessary to estimate more precisely the color of the light under which images are captured in order to render their colors back to those of a canonic illuminant.

Despite a recent reevaluation of former experimental data for several common color constancy algorithms shows that the relative performance of those algorithms changes considerably depending on the criteria by which they are judged, it is concluded that gamut–mapping and color–by–correlation still perform significantly better than the other algorithms tested. Finally, it must be reiterated that the problem of color constancy has conciliated lots of efforts from many researchers but not until recently a few of effective algorithms have been achieved being capable of working with real images.

8.1.2 State-of-the-Art Survey on Color Segment. Methods

There is no universal theory on color image segmentation yet, since it is basically one of psychophysical perception. So, to the problem of what segmentation method should be utilized, there is no answer. It depends on the application and experience. Most algorithms are extensions of former gray–level techniques, such as histogram thresholding, clustering, graph–partitioning, region–growing, edge detection, and fuzzy approaches. However, color allows more reliable image segmentations than for gray–level, since more information is provided.

As a rule, most clearly specified, algorithmically efficient, and robust are methods designed for particular small applications assuming well specified knowledge about the scene. Conversely, general purpose algorithms are not robust and usually not algorithmically efficient. It seems that separating processes for region segmentation and for object recognition is the reason of failure of general purpose segmentation algorithms. Unfortunately, authors do not habitually estimate the algorithmic complexity of their methods and often ignore comparing their novel ideas with existing ones.

A capital problem of color segmentation is how to employ color information as a whole for each pixel. A particular segmentation method can be directly applied to each component of a color space independently, or on the contrary to work with the color representation as a whole. For the first option, problems arise when trying to combine the results to obtain the final segmentation, whereas for the latter, the difficulty is that of finding a useful color space and metric to compare color differences. The choice of a color space is entirely an image/application–dependent question, since there is not any color space better than others and suitable to all kinds of images yet. Moreover, the need for a *color constancy* stage comes up when combining region segmentation and color object recognition.

Despite some authors have proposed some heuristic measures for the quantitative evaluation of segmentation results, the goodness of a segmentation approach depends on so many factors such as homogeneity, spatial compactness, continuity, correspondence with psycho–visual perception, that a single measure is uncertain to capture all of them in a meaningful way. Such goodness should be evaluated in respect to the adequacy that a given segmentation provides in the particular application one is interested in.

8.1.3 Continuous Color Formation Model

In this Chapter a continuous framework is suggested analytically tackling the color change issue as a natural continuous generalization of the usual color formation equations. This model is able to describe the formation of multispectral color signals by means of a unique expression which is particularized as needed in order to attain handier discrete formulations. As a result, the continuous color formation equation is identified as a Fredholm's integral equation of the first kind (IFK). Based on the description of the IFK equations in Appendix A, several numerical schemes are provided so as to undergo practical computations related with the recovery of spectral functions and the calculation of color maps.

Some conclusions can be derived from the work accomplished hitherto. First, the IFK model is a general model which contributes to the analytical study of the solutions to the color change problem, its existence, uniqueness, and behavior. Second, constraints to be fulfilled by the solution and the data subspaces in order a solution to the problem exists are established. This is new since that issue has always been approached from an *ad hoc* point of view, materialized into a discrete model, as one of those described above. Moreover, we explicitly show the structure of the solution function, as well as translate the continuous general expression into a numerical approach. Third, we relate several former well–established color constancy models to the one suggested by the discretization tools in Appendix A. In addition, it is clearly shown how those models are derived from an IFK and which of them are allegedly likelier to furnish better results.

After the theoretical study, some results are provided to support the statements done before. First, the problem of generalizing a discrete sensor is solved. Afterwards, it is shown that, in the case of spectral recovery, the IFK model in general provides better results than the Bilinear model, both in continuous color and in RGB coordinates, despite the fact that in the latter the improvement is not as great as in the former. This makes the IFK model an optimum candidate to be used in deeper recovery schemes. Furthermore, computing the combination of basis that provides the best results, we state that, by the addition of only one more sensor, we could recover pretty good approximations of the reflectance functions and the amount of such improvement is display.

Thereafter, the explicit computation of color mappings using physical models is attempted. We apply all the models described so far, as well as a linear fitting scheme to determine two different things. First, which linear model is the best at explaining the color change and how much the error of each of them is. Secondly, we compare the physical models with the results obtained by regression in order to determine which models do better. Results confirm that almost there is no difference between a general linear (affine) transformation and a homogeneous one, principally when the coefficients of the applications are constrained to be positive. Additionally, the error made by a diagonal application is taken into account, which turns out to be of the same order as the error made by the rest of models. These facts confirm that a linear transformation and a diagonal map suffice with only a marginal loss in precision. Despite those are not new results, they are proven here for a greater number of surfaces and illuminations.

Finally, it is shown that the IFK and the Bilinear models almost provide as good results as regression, whereas the Quadrature and the Diagonal models perform poorly. Hence, it seems clear that both the IFK and the Bilinear models closely predict the variation of colors due to light changes, while the Quadrature and the Diagonal discretizations are too weak for the same purpose, especially the latter. In particular, both the IFK and the Bilinear models present very similar error values, which suggests that in case a transformation in RGB coordinates is computed, both schemes basically converge. This is because both sides of those mappings are equally discretized to be RGB coordinates and no further information can be comparatively drawn from the continuous color description of the IFK model.

8.1.4 A Color Constancy Algorithm

Chapter 5 is devoted to the description of our color constancy algorithm which is able to render the image colors back to those that would been seen under the conditions of a canonic light. This algorithm is based on the computation of the histogram for all feasible maps arising between two sets of colors, namely, the one belonging to the actual image and the canonic set. In order to have a tool to select the most suitable mapping from all feasible transformations, a measure of likelihood for each of them is suggested. The likelihood function depends both on the number of times a particular transform is recovered from these sets of colors and its effectiveness in the task of recovering the most similar colors to the canonic set. This is done by computing a measure of similarity between transformed and canonic histograms.

A set of experiments using a pretty nourished number of images of objects under different illuminations were undergone in order to evaluate the performance of the proposed algorithm. These experiments consist in measuring the similarity there exists between the images in the set and those taken as the canonic ones in two specific situations, i.e., before and after the color constancy stage is carried out on them. Two types of similarity measures are employed to compute the difference between images. Analogous results are also obtained from the most similar color constancy method we found, namely, Finlayson's 2D gamut–mapping, and are compared to those obtained by our algorithm, which outperformed Finlayson's as well as attaining similar results to the color–by–correlation approach, the best color constancy algorithm to our knowledge.

The advantage of our algorithm with respect to the others is that ours is able to work with the information provided by only a canonic image as the reference to colors, rather than using an *a priori* correlation matrix that must span for any color and for any light. Besides, it is not restricted just to a discrete group of illuminants, but rather to a piece of a continuous set. Constraints on the set of feasible lights can easily be incorporated to the selection heuristic, improving the performance of the algorithm as a consequence. The behaviour of our color constancy algorithm is similar to the best known color constancy algorithms but, in contrast, it requires less information and is thus more adequate in tasks where there exists little control and knowledge on light sources, such as mobile robotics.

8.1.5 A Color Image Segmentation Algorithm

In Chapter 6, we claim that the problem of segmenting color images is faced, no matter their origin is static or from a video sequence, in a way that both coherent and stable segmentations are sought. For us, coherence means that components in a segmentation must correspond as close as possible to actual regions of the segmented scene, whereas stability has to do with the existence of components through time in a sequence, meaning that two consecutive frames must generate similar segmentations where corresponding components encompass similar areas in the scene.

To that purpose we suggest a *greedy* algorithm based on the computation of the minimum spanning tree which grows components attending to local properties of pixels. The process is fully controlled by an energy function that estimates the probability whether two components may be put together or not. Spurious regions that are helplessly generated during the growing process are removed accordingly to a quality index identifying such class of regions. Hence, a fast algorithm is achieved providing image segmentations that are good enough for identification purposes.

The segmentation algorithm is additionally extended to handle sequences in order to get stabler segmentations through time. For each new frame, this job is done by propagating forward the segmentation in the previous image, i.e, regions which get joined in a frame forming a bigger component are matched to other segments in the posterior frame by way of a distance that weights both *position* and *color appearance*, and then, these segments are grouped into a new component. Thus, it is granted that a pair of corresponding components in two consecutive frames of the sequence look similar.

Results show that segmentations using the Felzenszwalb&Huttenlocher's algorithm [FH98a], from which our method is inspired, have been improved and are similar in coherence and stability to those achieved by Figueiredo's EM in [FJ02], though being far faster. Furthermore, our segmentation algorithm will be used in the next Chapter to obtain the segmentations needed to carry out a set of experiments related with image retrieval and object recognition.

8.1.6 Comparing Images by Content

Chapter 7 illustrates the problem of comparing images by means of the content obtained from their color segmentations. A substantial group of measures of similarity is proposed within the frame of the IRM distance and the employ of multivariate Gaussian distributions for the color description of image regions. The performance of each and every one of these distances is examined in tasks such as images retrieval and object recognition. The best overall results are obtained by using the graph–partitioning algorithm described in Chapter 6 along with Lab color space and Fréchet distance, being as high as 95.21% of correct object identifications, which outperforms the other approaches assayed in previous Sections of the Chapter.

Two of these distances (Fréchet and Euclidean) are selected afterwards in order to study the effects of the color constancy algorithm in Chapter 5 in the segmentation of images from an autonomous robot, focusing on how this process reduces the mean distances defined amid images within a category compared to those involving images in distinct sets in two different situations, namely, before and after carrying out the color constancy step. Thus, we want to answer the question whether or not the reduction in color dissimilarity may help these sets of images to be more easily distinguished each other by way of their segmentations.

Results exhibit that, if colors in those sets are originally variable enough, the reduction in the mean distance within a category is greater than that between different categories, which also means that the relative distance between image sets has neatly grown therefore. Otherwise, the reduction is more balanced, since there is no big differences to reduce. This way, the applied distance measures based on segmentations are more selective at distinguishing images from dissimilar scenes as a consequence of the color constancy step performed on them.

In brief, this Chapter shows that color constancy reduces the amount of color variation within sets of images picturing the same scene by producing more similar segmentations. As claimed in Chapter 3, this is necessary if a identification process based on the segmentation of color images must be independent from the conditions of illumination. Moreover, as a by–product of the aforementioned goal, a significant number of content–based similarity measures are also studied in tasks that involve either object recognition or image retrieval. Those measures are thus more selective at distinguishing images from unlike scenes due to color constancy.

8.2 Future Work

Despite the amount of work done hitherto in this Thesis, both in color constancy and color segmentation, there are probably a myriad of new paths that could be followed or even should have been followed in advance to attain the objectives pursued for this Thesis. Among them, we can think of at least the next list with aspects which should be further worked out.

8.2.1 Color Constancy

- Investigating effective and inexpensive ways to get a linearized signal from an off–the–shelf camera in case its response appears not to be linear.

- Suggesting algorithms that automatically remove highlights and specularities from real images of unknown composition. These processes would furnish a *matte world* pictures which would be easy to coherently segment into more accurate homogeneous regions.

- Spectral recovery of reflectance functions from a set of color measures is an interesting issue to which more attention should be devoted. Also, determining ways to get better approximations as well as those of using the minimum number of sensors without losing essential information. This could derive in new multispectral cameras with few sensors that would hugely increase the quality of the recovered color information.

- Despite it is shown that the diagonal model for color change works fairly well, it would be interesting to think in new strategies to decrease the number of combinations that rises when extending the color constancy algorithm to computing liner maps other than the diagonal ones, since the algorithm might present here a bottleneck in time of process and precision due to the dimensional increase.

- Furthermore, it could be interesting to investigate new termination schemes for the color constancy algorithm to reduce the number of computations and the possibility of taking into account several peaks, which could correspond to combinations of different illuminants.

- Another extension of that research should cope with color constancy in video sequences. The goal then would be to maintain colors in a given object, that is been tracked, as constant as possible despite the uncontrolled illumination changes that it might experience.

8.2.2 Color Image Segmentation

- A good direction would be to include other cues to the segmentation algorithm such as texture, shape, or movement, bearing always in mind the time restriction that any segmentation process for mobile robotics suffers.
- In order to maintain the segmentations as stable as possible through time, it would be advisable keeping track of the segmented regions not only by their appearance, as done in our approach, but also by other cues such as shape.
- Along with the tracking, another enhancement would be the inclusion of models for color adaptation. These models should be used to predict the color change as objects relatively change their position from the mobile platform. Some experiments were already done on this subject by the author [VLTS02], though they have not been explicitly included in this Thesis.
- An important routine, most of the time completely forgotten, is that of foreground–background segmentation, which would be extremely helpful in order to focus the attention onto some object in the whole scene depending on the tasks scheduled to be carried out.
- It is also capital for a mobile robot to be capable of self–localizing by the only information provided by its visual system, comparing these images to its database of knowledge historically recorded during an initial phase of world exploration. This knowledge of the world could also be combined with other information provided by different means to build a more accurate description of the environment.
- To attain that purpose at least three basic capabilities would be of relevance, namely, generating a comprehensive description of a whole scene from a set of different views at different times, retrieving in a fast manner the likeliest images from a database using a single image as a query, and finding targets within an image of the scene.
- These capacities also need more research on content–based descriptors, distance measures, and partial image scanning techniques. It would be ideal to browse through huge image databases to find the set of likeliest images presenting a given detail observed in the scene, and *vice versa*, i.e., to find in the scene an object assigned to be the target in a robotic task.

As seen, we have only touched the top of the heap and raised a bit of dust!