ADVANCES IN COLOR IMAGE SEGMENTATION*

L. Lucchese[†‡] and S.K. Mitra[†]

[†]Dept. of Electrical and Computer Eng., University of California, Santa Barbara

[‡]Dept. of Electronics and Informatics, University of Padua, Italy

Abstract

Segmentation is the low-level operation in image processing which partitions an image into disjoint and homogeneous regions. In this paper, we consider segmentation of color images, a task which plays a very important role in many multimedia applications concerned with color information. In particular, we review the main contributions in this field in the past few years.

1 Introduction

The essential goal of segmentation is to decompose an image into parts which should be meaningful for a certain application. In many multimedia applications color segmentation plays an important role. For instance, in *digital libraries* large collections of images and videos need to be catalogued, ordered, and stored in order to efficiently browse and retrieve visual information [1, 2]. Color and texture are the two most important low-level attributes used for *content-based retrieval* of information in images and videos. Because of the complexity of the problem, segmentation with respect to both color and texture is often used for indexing and managing the data. Another example is in the transmission of information over the Internet. Nowadays, huge streams of multimedia data circulate over the Internet where the limited bandwidth available creates the need for data compression. Current technology provides coding schemes which try to reduce visual artifacts by imitating the functions of the human visual system [2, 3]. They seek a semantic representation of the scene by subdividing it into regions which are psycho-visually meaningful. Such a partitioning is obtained through segmentation. Compression is then achieved by allocating more bits to areas visually more important and fewer bits to less important details. A further example is in the latest wireless communication systems which allow the transmission of both speech and images. *Hand-held wireless sets* are now available which may also display color imagery with a limited resolution. Compression issues arise as in the above example with the further constraint of a limited availability of bits for displaying. In this application, segmentation is then important not only for compression but also for color quantization.

Until a few years ago, segmentation techniques were proposed mainly for gray-level images on which rather comprehensive surveys can be found in [4]-[7]. There has been a remarkable growth of algorithms for segmentation of color images in this last decade. Most of the times, these are kind of "dimensional extensions" of techniques devised for gray-level images; thus they exploit the well-established background laid down in that field. In other cases, they are *ad hoc* techniques tailored on the particular nature of the color information and on the physics of the interaction of light with colored materials. Here we present a brief survey on these techniques and we propose a classification scheme for them. Basically, we divide the segmentation algorithms into: 1) *feature-space based techniques*; 2) *image domain-based techniques*; and 3) *physics-based techniques*. Each category is then further subdivided. As far as the first two categories are concerned, the further subdivision is suggested by the analogous classification schemes proposed for gray-level images [4]-[7]. Such a classification is not always straighforward since some techniques resort to more than one strategy to achieve segmentation and thus cannot be sharply categorized. The techniques of the third category instead adopt specific models of the interaction of light with colored materials of various nature and therefore they have no counterpart in the field of gray-level image segmentation.

This work is organized as follows. Section 2 presents the feature-based techniques; Section 3 reports on the class of image domain-based techniques. The algorithms based on physical models describing the interaction of light with color are discussed in Section 4. Section 5 finally draws the conclusions.

2 Feature space-based techniques

If we assume that color is a constant property of the surface of the objects portrayed in an image and we map each pixel of the image into a certain color space, it is very likely that the different objects present in the image will manifest themselves as clusters or clouds of points. The spreading of these points within each cluster is mainly determined by color variations due to shading effects and to the noise of the acquisition device. On the other hand, if, instead of mapping pixels into color spaces, we build some *ad hoc* histograms upon color features, such as hue, for instance, it is likely that the objects will appear as peaks within these histograms. Therefore, the problem of segmenting the objects of an image can be viewed as that of finding some clusters, according to the first strategy mentioned above, or as that of finding the peaks of some histograms, according to the second strategy. These two approaches share a common property: they work in a certain feature space and they neglect the spatial relationships among colors. For this reason, we have decided to group them under the common denomination of *feature space-based techniques*; they will be however separately analized in the following sections.

*This work was supported in part by C.N.R., Italy, through a fellowship for research abroad (n. 203.15.9) and in part by a University of California MICRO grant, with matching supports from National Semiconductor Corp., Lucent Technologies, Raytheon Missile Systems, Tektronix Corp., and Xerox Corp.

0-7803-5796-5/99/$10.00 © 1999 IEEE

2.1 Clustering

A great many techniques have been proposed in the literature of cluster analysis [11, 12]. A classical technique for color image segmentation is the *k-means* (or *c-means*) *algorithm* [13], widely adopted also for vector quantization and data compression. Park *et. al* [14] apply this algorithm to a pattern space represented by RGB coordinates while Weeks and Hague [15] apply it to the HSI space. The k-means algorithm has been mostly used however in its fuzzy version (*fuzzy k-means algorithm*) [16]-[20]; a comparison between k-means and fuzzy k-means clustering is reported in [21]. The *possibilistic approach* to clustering of [22] is closely related to these fuzzy techniques. *ISODATA* (Iterative Self-Organizing Data Analysis Techniques) [12] is another algorithm often used for color space clustering [23, 24]. Comaniciu and Meer [25] resort instead to the *mean shift algorithm* which is a nonparametric procedure for estimating density gradients of pattern distributions. *Competitive learning* based on the least-square criterion is employed in [26, 27] whereas the theory of *connected components* [28] is adopted by Wang *et al.* in [29]. An original technique, proposed by Yung and Lai [30], adopts the *constrained gravitational clustering*. The RGB space is represented in a tree strucure by Uchimura in [31] and clustering is achieved by simplification of the tree. Kehtarnavaz *et al.* choose a 2D color space called *geodesic chromaticity* in which they introduce a *multi-scale clustering*; this algorithm determines the prominent color clusters through their *lifetime* [33]. Shi and Malik [34] and Shi *et al.* [35] tackle image segmentation via clustering as a graph partitioning problem. Wu and Lehay [36] originally devised an algorithm for segmentation based on the *minimum cut* of the graph representing an image in a certain feature space. In [34] and [35] the authors further develop this idea and report an interesting technique for finding a normalized version of the minimum cut. Moreover, Shah [37] formulates the analytic analog counterpart of the graph-theoretic formulation given above. Lucchese and Mitra [38] present a technique which first finds clusters in the $u'v'$ chromaticity plane and then associates them with proper luminance values, respectively, with a 2D and a 1D k-means algorithm.

2.2 Adaptive k-means clustering

A special classification has to be devoted to a class of segmentation algorithms that combine the idea of k-means clustering with the desirable properties of local adaptivity to the color regions and of spatial continuity. In this sense, this class of algorithms might be regarded as lying in between the feature space-based techniques discussed here and the image domain-based techniques to be considered next. The traditional clustering techniques considered in the previous section classify pixels into clusters based only upon their color; each cluster is then characterized by a constant color value and no spatial constraints are imposed. In [39] Pappas introduces a generalization of the k-means clustering algorithm which is adaptive and includes spatial constraints; this algorithm considers the segmentation of gray-level images as a *maximum a posteriori probability (MAP) estimation*. The extension of this technique to color images is proposed by Chang *et al.* in [40]. A *Gibbs Random Field (GRF)* [41, 42] is used as an image prior to model and enforce spatial homogeneity constraints. Saber *et al.* [43] extend the algorithm of [40] to synergically combine color image segmentation and edge linking; in particular, they apply a split-and-merge strategy (see Section 3.1) on the regions of the segmented map so as to enforce consistency with the edge returned by color edge detector (see Section 3.3). Luo *et al.* in [44] modify the algorithm of [40] to incorporate a color space called Lst (instead of the RGB) and a certain color difference that can be defined within this space; they claim that with these provisions their algorithm can return segmentations physically more coherent. The same authors in [45] extend the algorithm of [40] by introducing in it derivative priors and by combining both region-based and edge-based statistical forces in segmentation.

2.3 Histogram thresholding

Histogram thresholding is among the most popular techniques for segmenting gray-level images and several strategies have been proposed to implement it [4]-[8]. In fact, peaks and valleys of the 1D brightness histogram can be easily identified, respectively, with objects and backgrounds of gray-level images. In the case of color images, things are a little more complicated since one has to identify different parts of a scene by combining peaks and valleys of three histograms or by partitioning a 3D histogram. A common problem with the histogram-based techniques is that often, because of noise, the profiles of the histograms are rather jagged giving rise to spurious peaks and thus to segmentation ambiguities; to prevent this from happening, some smoothing provisions are usually adopted.

Celenk and Uijt de Haag [46] indipendently threshold three histograms based on RGB coordinates by maximizing within-group variance and combine the three results with a predicate logic function. Shafarenko *et al.* [47] use a *watershed algorithm* [48] to segment either the 2D or the 3D color histogram of a color image; the histograms are built from $L^*u^*v^*$ coordinates and "coarsened" through convolution with a spherical window to prevent oversegmentation. Tseng *et al.* [49] use only hue information and suggest a circular histogram thresholding of such attribute. The histogram smoothing is achieved by means of a *scale-space filter* [33]. The approaches of [50]-[53] have in common the partition of a cylindrical color space representing hue, saturation, and intensity into chromatic and achromatic regions. The former is segmented by using the hue histogram and the latter is segmented by using the intensity histogram. A scale-space filtering is adopted in [52]. A similar approach is followed also by Bobottka and Pitas [54] who single out faces from color images by defining appropriate domains corresponding to skin-like regions within the HSV space; by disregarding the value V (luminance), robustness is obtained against changes in illumination and shadows. Guo *et al.* [55] suggest an *entropy-based thresholding* which assumes that samples or patterns in the feature space $L^*u^*v^*$ are generated by two distinct sources called *mode* and *valley*; first they classify patterns in either categories by using entropy thresholding and then they determine the number of modes in the feature space with a *modified Akaike's information criterion*. Saber *et al.* [56] model the distribution of the chrominance components of the objects in a scene as *Gaussian probability density functions (PDF's)* allowing this way an adaptive setting of the object-class thresholds. Liu *et al.* [57] devise an adaptive threshold function for both RGB and HSI spaces by using *B-splines*; they can separate cell nuclei by means of this thresholding function which is obtained in a prelimiary *learning phase*. Lucchese and Mitra [58] suggest smoothing the hue histogram in $L^*u^*v^*$ coordinates by working with the low-low band of the wavelet transform of the image undergoing segmentation.

3 Image domain-based techniques

All the segmentation algorithms of the previous section exclusively operate in some feature spaces. Thus the regions (segments) they return are expected to be homogeneous with respect to the characteristics represented in these spaces; however, there is no guarantee at all that these regions have also spatial compactness, which is a second desirable property in segmentation applications besides homogeneity. In fact, cluster analysis and histogram thresholding account in no way for the spatial locations of pixels; the description they provide is global and it does not exploit the important fact that points of a same object are usually spatially close due to *surface coherence* [8]. On the other hand, if pixels are clustered exclusively on the basis of their spatial relationships, the end result is likely to be with regions spatially well connected but with no guarantee that these regions are also homogeneous in a certain feature space. In the literature of segmentation of gray-level images, a great many techniques have been suggested that try to satisfy both feature homogeneity and spatial compactness at the same time [4, 6]. The latter is ensured either by subdividing and merging or by progressively growing image regions, while the former is adopted as a criterion to direct these two processes [4, 5, 6, 8]. According to the strategy preferred for spatial grouping, these algorithms are usually divided into split-and-merge and region growing techniques; this distinction may also be extended to the corresponding algorithms for color image segmentation which will be analized in the following sections. In the class of image domain-based techniques we have considered also a family of algorithms which exploit spatial information in neural network classifiers and the group of algorithms that partition images by finding the edges between homogeneously colored regions.

3.1 Split-and-merge techniques

A common characteristic of these methods is that they start with an initial inhomogeneous partition of the image (usually the initial segment is the image itself) and they keep performing splitting until homogeneous partitions are obtained. A common data structure used to implement this procedure is the *quadtree representation* [8, 12] which is a multiresolution scheme. After the splitting phase, usually there exist many small and fragmented regions which have to be somehow connected. The merging phase accomplishes this by associating neighboring regions and guaranteeing that homogeneity requirements are met until maximally connected segments can be produced. The *region adjacency graph (RAG)* is the data structure commonly adopted in the merging phase [8, 12]. In many algorithms, smoothness and continuity of color regions are enforced with the adoption of a *Markov Random Field (MRF)* [41, 42] which basically is a stochastic process characterized by the following property: the conditional probability of a particular pixel taking in a certain value is only a function of the neighboring pixels, not of the entire image. Besides, the *Hammersley-Clifford theorem* [42] establishes the equivalence between MRF's and *Gibbs distributions*.

Panjwani and Healey [59] model color texture in RGB components by means of a *Gaussian Markov Random Field (GMRF)* which embeds the spatial interaction within each of the three color planes as well as the interaction between different color planes. In the splitting phase, the image is recursively partitioned into square regions until each of them contains a single texture described by a color GMRF model. This phase is followed by an agglomerative clustering phase which consists of a conservative merging and of a stepwise optimal merging process. Liu and Yang [60] define instead an MRF on the quadtree structure representing a color image and use the above mentioned equivalence with a Gibbs distribution. With a *relaxation process* [5] they control both splitting and merging of blocks in order to minimize the energy in the Gibbs distribution; this is shown to converge to a *MAP estimate* of the segmentation. Numerous are the variations in the split-and-merge strategies. In [61] a k-means algorithm is used for both classifying pixels in the splitting phase and grouping pattern classes in the merging phase. In [62] the splitting is initially performed by segmenting the luminance and then refined by checking the chrominance homogeneity of the regions obtained; the merging is based on an *ad hoc* cost function. In [63] the splitting is operated with the *watershed transform* [48] of the gradient image of the luminance component simplified by a *morphological grayscale opening* [8, 9, 10]; the merging step is realized with a *Kohonen's self-organizing map (SOM)* [12]. Shafarenko *et al.* [64] apply instead the watershed transform to the $L^*u^*v^*$ gradient of images and merge the patches of the watershed mosaic according to their color contrast until a termination criterion is met. A similar splitting approach is adopted in [65] whereas the merging phase is performed by iteratively processing the RAG constructed upon the resulting oversegmented regions. Barni *et al.* [70] implicitly implement a split-and-merge strategy with a fuzzy expert system. Gevers *et al.* [66, 67] believe that split-and-merge algorithms based on a quadtree structure are not able to adjust their tessellation to the underlying structure of the image data because of the rigid rectilinear nature of the quadtree structure; therefore, they suggest replacing it with an *incremental Delaunay triangulation* [12]. A further alternative possibility is to use *Voronoi diagrams* [12] as proposed by Schettini *et al.* [68] and by Itoh and Matsuda [69].

Broadly speaking, we can fit to the class of split-and-merge techniques also some algorithms based upon differential equations and pyramidal data structures. At first glance, they do not appear to belong to this category since the strategies they adopt to achieve segmentation are rather different from those reviewed so far; but a more careful look into them will bring to light an underlying split-and-merge idea.

Pollak *et al.* [71] and Gao *et al.* [72] apply *stabilized inverse diffusion equations (SIDE's)* [73] to segmentation of vector-valued images. The finest possible segmentation is initially assumed: each pixel represents a separate region. During an evolution process driven by diffusion equations, two neighboring regions are merged whenever a certain color difference equals zero.

The usefulness of pyramidal representation of images for segmentation was pointed out by Burt *et. al* [74] about two decades ago and ever since a number of methods to segment images by working with pyramids have appeared. It is well-known that pyramids are data structures in which images can be represented at different resolutions (fine-to-coarse) by means of tapering layers recursively obtained by averaging and downsampling their respective underlying layers (the finest layer at the bottom of a pyramid is the image itself) [8]. Thus *father-son relationships* can be naturally introduced between adjacent layers of pyramids; segmentation can be achieved with a *pyramid-linking process* [74] based on a tree structure where the values of the fathers at a certain high layer are propagated down to the sons of the lowest level. The construction of a pyramid can be regarded as a splitting phase while the subsequent linking process can be seen as a merging phase. Recently, Lozano and Laget [75] have suggested *fractional pyramids* for segmentation of color images and Ziliani and Jensen [76] have proposed a modified version of the linking approach of [74].

3.2 Region growing techniques

An homogeneous region of an image may be obtained through a growth process which, starting from a preselected seed, progressively agglomerates points around it satisfying a certain homogeneity criterion; the growth process stops when no more points can be added to the region. The region growing techniques are mainly aimed at processing single regions; nevertheless, by combining different and subsequent growth processes, one may agglomerate in regions all the points of an image, obtaining this way a segmentation of it. After a region growing procedure, there might exist some very small regions or there could be two or more neighboring regions grown at different times exhibiting similar attributes. A common post-processing provision consists therefore in a merging phase that eliminates such instances by generating broader regions. The region growing can be considered a sequential clustering or classification process [5]; thus the dependence of the results on the order according to which the image points are processed has to be accounted for. The main advantage offered by this kind of techniques is that the regions obtained are certainly spatially connected and rather compact. As for the clustering techniques of Section 2.1, where a similar problem arises in the feature space, also for the region growing techniques one is faced with the problem of choosing suitable seed points and an adequate homogeneity criterion.

As far as gray-level image segmentation is concerned, several region growing strategies can be found in the literature [4, 5, 6]. For color images some new interesting strategies for region growing-based segmentation have been recently proposed. Tremeau and Borel [77] suggest several different homogeneity criteria operating in RGB coordinates. In a first phase, they generate a certain number of connected regions with a growing process and, in a second phase, they merge all the regions having similar color distributions; after the second phase, the regions are therefore colorically similar but they may be disconnected. Kanai [78] develops a segmentation algorithm which resorts to both color and intensity information. The markers (seeds) are extracted from intensity via *morphological open-close operations* and from color through quantization of the HSV space; *joint markers* are defined as the sets comprising both kinds of markers. A region growing process based on a watershed algorithm starts from these *joint markers*. A region merging process eventually reduces the number of segmented regions. In [79], the inital seeds are generated by retaining the significant local minima of the magnitude of the color image gradient; however, with this algorithm the two following situations might arise: 1) there is more than one seed per region; 2) small objects do not have any seed. The authors devise a procedure for obtaining markers having a one-to-one correspondence with the image regions. The region growing is performed with a watershed-like algorithm proposed by the authors and working on the original color image instead of on a gradient image. Deng *et al.* [80] determine a limited number of color classes within an image through *color quantization* and propose a criterion for "good" segmentation based on them. The application of this criterion within local windows and at multiple scales generates *J-images* in which high and low values respectively correspond to possible region boundaries and to region centers. A region growing method is adopted where the seeds are the valleys of the *J-images*; the resulting oversegmentation is finally removed with a merging phase. Rehrmann and Priese [81] suggest using a special hexagonal topology in a hierarchical region growing algorithm which results indipendent of the starting point and of the order of processing. Ikonomakis *et al.* [82] develop an algorithm to segment both gray-scale and videophone-type color images; the procedure is a standard region growing process followed by region merging. Color homogeneity is tested with measurements in the HSI space.

If one define a cluster as a "collection of touching pixels that have almost the same color while the change in color is gradual," the fuzzy nature of the segmentation problem can be emphasized. Moghaddamzadeh and Bourbakis [83, 84] have adopted this vision of the problem and advanced two algorithms working in RGB coordinates to implement a region growing strategy for both fine and coarse segmentation of color images. A fuzzy approach for region growing segmentation is adopted also in [85] whose algorithm is based upon several linguistic rules defining relationships among hue, chroma, and intensity. Colantoni and Laget [86] compare the results of four different algorithms obtained by the various combinations of region growing and watershed transform in a presegmentation step and in the segmentation algorithm. Images are represented in $L^*a^*b^*$ coordinates and handled by means of RAG's and *contour graphs*.

3.3 Edge-based techniques

Segmentation can also be obtained by detecting the edges among regions. The literature of edge-based segmentation for gray-level images is rich of techniques [4, 5, 7] and several are also the algorithms proposed for detection of discontinuities within color images. It is well-known that edges can be found in gray-level images by resorting to functions approximating gradients or Laplacians of images, which are of course scalar functions. Gradient functions for color images may be basically defined in two ways: 1) by embedding in a single measure the variations of all three color channels or 2) by computing the gradients of the single channels and by combining them according to certain criteria.

The first approach is based on the *first fundamental form*, which in differential geometry constitutes the multidimensional extension of the single-valued gradient [87]. Upon this metric for vector-valued funtions are based the chromatic edge detectors of [88] and [89], both operating in RGB coordinates. Examples of the second approach are given instead by [90] and [91]. Carron and Lambert [90] propose three different combinations of gradients of hue, saturation and intensity computed in HSI coordinates. Tao and Huang [91] find clusters in the RGB space and compute egdes as the transitions from one cluster to another; the gradient information in each color channel is computed with a Sobel operator.

A truly original algorithm for boundary detection is proposed by Ma and Manjunath [92]: they use a kind of *predictive coding* model to identify the direction of change in color and texture at any point and at a given scale; this give rise to an *edge flow* which, through propagation, converges to the image boundaries. Perez and Koch [93] gather several arguments in favor of hue as the most important color attribute for segmentation; in particular, they demonstrate that, if the *integrated white condition* holds, hue is invariant to certain kinds of highlights, shading, and shadows. Egde detection is achieved by finding the zero crossings of the convolution of the *hue image* with a suitable Laplacian function. *Neural networks* in the form of Kohonen's SOM's [12] are used for contour segmentation in [94] and [95].

In the context of the edge-based techniques we can fit also the framework for object segmentation based on *color snakes*. Snakes or *active contours* were originally proposed by Kass and Witkin [96] and have received great attention ever since. The classical snakes approach consists in deforming an initial contour towards the boundary of an object to be detected; the deformation is obtained by minimizing a global energy designed in such a way that its local minimum is attained in correspondence of the boundary of the object [97]. The formulation of active contours for vector-valued images (and therefore for color images) is due to Sapiro [97, 98] who defines a new Riemannian metric based on the first fundamental form. Moreover, Sapiro shows the close relationships existing beween the active contours for color images, which he calls *color snakes*, and other algorithms based on *partial differential equations (PDE's)*, *anisotropic diffusion*, and *variational approaches* for image segmentation [33]. Gevers *et al.* propose instead *color invariant snakes* [99] that use color-invariant gradient information to drive the deformation process; in this way the snakes return region boundaries rather insensitive to disturbances due to shadowing, shadows and highlights.

3.4 Neural network-based classification techniques

A class *per se* is constituted by segmentation techniques adopting classification techniques based on *neural networks*. It is well-known that neural networks are structures made up of large numbers of elementary processors (cells) massively interconnected and performing simple functions [10, 12]. Their design try to imitate the information processing of biological neural cells. Despite the complexity that in some cases they require to be implemented, they offer two important properties in pattern recognition tasks: high degree of parallelism, which allows for very fast computational times and makes them suitable for real time applications, and good robustness to disturbances, which allows for reliable estimates. Another interesting feature is that, in the case of image segmentation, neural networks permit to account for spatial information; on the other hand, one has to know beforehand the final number of segments within an image and to perform a preliminary *learning phase* during which the network is trained to recognize patterns. Usually the number of classes is derived with some *a priori* knowledge on the problem or in a preprocessing stage.

A number of algorithms have been proposed for segmenting gray-level images with neural networks [7]. We will discuss next some of the neural network-based techniques offered for color image segmentation. Campadelli *et al.* [100] present two segmentation algorithms based on the idea of [101] of regarding the segmentation problem as the problem of minimizing a suitable energy function for a *Hopfield network* [12]. A similar approach based on the minimization of an energy function associated with a Hopfield neural network is undertaken in [102], where a preclassification algorithm spots out some *regions of interests (ROI's)* in a biomedical RGB image, and in [103], where an active-region segmentation algorithm is presented. Okii *et al.* [104] present an algorithm based on a three-layered neural network for segmentation of medical stained images, where three are the possible classes, nuclear cell, interstitium, and background represented by three different colors. In this regard, we point out that the adoption of three layers is very common in neural networks since this structure is capable of implementing arbitrarily complex decision surfaces composed of intersecting hyperplanes in the pattern space [10, 12]. Classical is also the *learning phase* adopted in [104] which is obtained with a *back-propagation algorithm* [10, 12]. Similarly, Funakubo [105] uses two three-layered neural networks with learning through back-propagation to separate cells from background in medical images. The aim of [106] and [107] is slightly different from the usual one of segmentation and it consists in determining the colors of inks used to generate a multi-colored picture created by printing dots of cyan, magenta, yelow, and black. Nine possible combinations arise which constitute the output of a hierarchical modular neural network. Other examples of neural networks used for color segmentation are [108], where a neural network is trained to identify the color of a desired object for automated tracking purposes, and [109], where a *neural gas network* is employed.

4 Physics-based techniques

All the algorithms examined so far are certainly prone to segmentation errors if the objects portrayed in the color images are affected by highlights, shadowing and shadows. All these phenomena cause the appearance of color of uniformly colored surfaces to change more or less drastically, whence those algorithms are very likely to return oversegmented regions. The only way to overcome this drawback is to analyse how light interact with colored materials and to introduce models of this physical interaction in the segmentation algorithms. This motivated the name of *physics-based techniques* given to them. The mathematical tools they use do not significantly differ from those adopted by the algorithms of the previous section; the major difference with respect to those is the underlying physical model accounting for the reflections properties of colored matter.

Colored materials may be divided into three main categories: optically inhomogeneous dielectrics, optically homogeneous dielectrics, and metals. A milestone in the field of physics-based segmentation was laid by Shafer in [110] where he introduces the *dichromatic reflection model* for inhomogeneous dielectrics; it states that the total radiance of the light reflected by an inhomogeneous dielectric may be split into two contributions, one stemming from the object's surface, and the other from the underlying object's bulk. Besides, each of these terms can be factored into a geometric contribution and a spectral power distribution. This model may effectively explain some particular shapes of clusters in color spaces. Based upon this, Klinker *et al.* [111] set up an algorithm (using either a split or a region-growing strategy) which makes some optical hypotheses relating objects' colors, shading and highlights and try to justify with them the shapes of clusters in the RGB space. The main limitation of this technique is that it can be applied only to inhomogeneous dielectrics. Simplicity and effectiveness of representation have made the dichromatic reflection model very popular and many physics-based techniques for segmentation resort to it [112]-[115]. Tsang and Tsang [116] use the dichromatic reflection model in the HSV space to detect edges.

A very major contribution related with the model proposed by Shafer is represented by the work of Bajcsy *et al.* [117]. They propose a color reflection model based on the dichromatic reflection model for dielectric materials and on a particular color space, called S space, built upon three orthogonal basis functions. They prove that it is possible to separate specular and diffuse interface reflections and some inter-reflections from body reflections since they produce clusters with very peculiar shapes in the S space. The algorithm suggested in [117] allows segmentation of uniformly colored dielectric surfaces under singly colored scene illumination.

Healey [118] proposes the *unichromatic reflection model* for metals by supporting it with extensive experimental results. This model states that metals give rise to a reflectance function which stems only from their surfaces and which, analogously to the dichromatic reflection model, can be separated into a geometric factor and a purely spectral component. This independence of wavelength and geometry in the reflectance function hints that geometric effects in a scene can be factored out of color pixel values in an image (*normalized colors*). In [118] and [119] Healey comes up with two segmentation algorithms based on such a normalization of color which can handle inhomogeneous dielectrics and metals at the same time.

The methods discussed above are able to handle one or two classes of materials (inhomogeneous dielectrics and metals) in the presence of a single illumination source. A more general and more complicated algorithm which also takes into account for multiple illuminations is presented by Maxwell and Shafer in [120]. They introduce a general framework for segmentation of complex scenes which formulates multiple physical hypotheses about image formation. These hypotheses define broad classes for shape, illumination, and material properties of simple image regions obtained through an initial rough segmentation. A ranked set of possible segmentations is generated by analyzing, merging, and filtering the hypotheses; the pruning of such set finally yields a restricted number of plausible segmentations (interpretations) of the scene.

5 Conclusions

In this paper we have presented an overview of algorithms for color image segmentation, which represents an important issue for many multimedia applications. A universal algorithm for segmenting images certainly does not exist and, on the contrary, most techniques are tailored on particular applications and may work only under certain hypotheses. Some authors [60, 121] have proposed heuristic measures for quantitative evaluation of segmentation results. However, the goodness of a segmentation result depends on so many factor such as homogeneity, spatial compactness, continuity, correspondence with psycho-visual perception [7], *etc.*, that a single measure is unlikely to capture all of them in a meaningful way. Such goodness should be evaluated by the usefulness that segmentation can provide in the particular application one is interested in.

References

[1] Special Session on Digital Libraries: Representation and Retrieval, *IEEE Trans. on PAMI*, Vol. 18, No. 8, Aug. 1996.

[2] Special Issue on Segmentation, Description, and Retrieval of Video Content, *IEEE Trans. on CASVT*, Vol. 8, No. 5, Sept. 1998.

[3] M. Kunt *et al.*, Proc. of the IEEE, Vol. 73, No. 4, pp. 549-575, Apr. 1985.

[4] K.S. Fu *et al.*, *Patt. Rec.*, Vol. 13, pp. 3-16, 1981.

[5] A. Rosenfeld and A. Kak, *Digital Picture Processing*, Vol. 2, 2*nd* Ed., Academic Press, New York, NY, 1982.

[6] R.M. Haralick *et al.*, *Comp. Vis., Graph., and Im. Proc.*, Vol. 29 , No. 1, pp. 100-132, Jan. 1985.

[7] N.R. Pal *et al.*, *Patt. Rec.*, Vol. 26, No. 9, pp. 1277-1294, 1993.

[8] R. Jain *et al.*, *Machine Vision*, McGraw-Hill, Inc., New York, NY, 1995.

[9] A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[10] R.C. Gonzales and R.C. Woods, *Digital Image Processing*, Addison-Wesley, Reading, MA, 1992.

[11] M.R. Anderberg, *Cluster Analysis for Applications*, Academic Press, New York, NY, 1973.

[12] S.-T. Bow, *Pattern Recognition and Image Preprocessing*, Marcel Dekker, Inc., New York, NY, 1992.

[13] J. McQueen, *Proc. of the 5th Berkeley Symp. on Math. Stat. and Prob.*, Vol. 1, pp. 281-296, 1967.

[14] S.H. Park *et al.*, *Patt. Rec.*, Vol. 31, No. 8, pp. 1061-1076, Aug. 1998.

[15] A.R. Weeks *et al.*, *Proc. of the SPIE'97*, pp. 143-154.

[16] J. Wu *et al.*, *Jour. of Electr. Imaging*, pp. 397-403, Oct. 1994.

[17] Y. Kamei-Harada *et al.*, *Jap. Journal of Fuzzy Theory and Syst.*, Vol. 6, No. 5, 1994, pp. 651-672.

[18] A. Brun-Buisson *et al.*, *Proc. of 1995 Int'l Conf. on Quality Control by Artif. Vision*, pp. 104-109.

[19] Y. Rui *et al.*, *Proc. of ICIP'96*, Vol. 1, pp. 53-56.

[20] P. Schmid *et al.*, *Proc. of the 6th Int'l Conf. on Image Proc. and its Appl.*, Vol. 2, pp. 688-692.

[21] S. Ray *et al.*, *Proc. of Dig. Image Comp.: Tech. and Appl.*, Dec. 1995, pp. 86-92.

[22] K.-B. Eum *et al.*, *Proc. of 1996 IEEE Int'l Conf. on Syst., Man and Cyb.*, Vol. 2, pp. 1150-1155.

[23] N. Liu *et al.*, *Proc. of the 2nd Australian and New Zealand Conf. on Int. Inf. Syst.*, Dec. 1994, pp. 293-296.

[24] K. Takahashi *et al.*, *Proc. of ACCV'95*, Vol. 1, pp. 523-527.

[25] D. Comaniciu *et al.*, *Proc. of CVPR'97*, pp. 750-755.

[26] T. Uchiyama *et al.*, *IEEE Trans. on PAMI*, Vol. 16, No. 12, pp. 1197-1206, Dec. 1994.

[27] A. Verikas *et al.*, *Proc. of 1995 Int'l Conf. on Neural Net.*, Vol. 4, pp. 1713-1718.

[28] A. Rosenfeld, *Patt. Rec.*, Vol. 16, No. 1, pp. 47-50, 1983.

[29] W. Wang *et al.*, *Proc. of 1997 IEEE Int'l Conf. on Syst., Man, and Cyb.*, Vol. 2, pp. 1089-1093.

[30] H.C. Yung *et al.*, *Opt. Eng.*, Vol. 37, No. 3, pp. 989-1000, Mar. 1998.

[31] K. Uchimura, *Trans. of the Inst. of Elec. Eng. of Japan - C*, Vol. 114-C, No. 12, pp. 1320-1321, Dec. 1994.

[32] N. Kehtarnavaz *et al.*, *Proc. of 1998 IEEE Southwest Symp. on Image Anal. and Interp.*, pp. 142-147.

[33] B.M. ter Haar Romeny, Editor, *Geometry-Driven Diffusion in Computer Vision*, Kluwer Acad. Publ., Dordrecht, The Netherlands, 1994.

[34] J. Shi *et al.*, *Proc. of CVPR'97*, pp. 731-737.

[35] J. Shi *et al.*, *Proc. of ICIP'98*, Vol. I, pp. 943-947.

[36] Z. Wu *et al.*, *IEEE Trans. on PAMI*, Vol. 15, No. 11, Nov. 1993.

[37] J. Shah, *Proc. of ICIP'98*, Vol. III, pp. 766-769.

[38] L. Lucchese *et al.*, *Proc. of CBAIVL'99*, Fort Collins, CO, June 1999.

[39] T.N. Pappas, *IEEE Trans. on SP*, Vol. 40, No. 4, pp. 901-914.

[40] M.M. Chang *et al.*, *Journal of Elec. Imag.*, Vol. 3, No. 4, pp. 404-414, Oct. 1994.

[41] G.R. Cross *et al.*, *IEEE Trans. on PAMI*, Vol. 5, No. 1, pp. 25-39, Jan. 1983.

[42] S. Geman *et al.*, *IEEE Trans. on PAMI*, Vol. 6, No. 6, pp. 721-741, Nov. 1984.

[43] E. Saber *et al.*, *Proc of ICASSP'96*, Vol. 4, pp. 2176-2179.

[44] J. Luo *et al.*, *Proc. of ICIP'97*, Vol. III, pp. 58-61.

[45] J. Luo *et al.*, *Proc. of ICIP'98*, Vol. III, pp. 780-784.

[46] M. Celenk *et al.*, *Proc. of the SPIE'98*, pp. 250-259.

[47] L. Shafranenko *et al.*, *IEEE Trans. on IP*, Vol. 7, No. 9, pp. 1354-1358, Sept. 1998.

[48] H. Digabel *et al.*, *Proc. of the 2nd European Symp. on Quant. Anal. of Microstruc. in Material Sci. , Biol. and Med.*.

[49] D.-C. Tseng *et al.*, *Proc. of the 3rd Int'l Conf. on Doc. Anal. and Rec.*, Aug. 1995, Vol. 2, pp. 673-676.

[50] D.-C. Tseng *et al.*, *Proc. of the Nat'l Sci. Counc., Rep. of China - A*, Vol. 18, No. 3, pp. 305-313, May 1994.

[51] P. Pujas *et al.*, *Proc. of the 3rd Int'l Conf. Interface to Real and Virtual Worlds*, Feb. 1994, pp. 285-295.

[52] N. Herodotou *et al.*, *Proc. of 1998 IEEE Symp. on Adv. in Dig. Filt. and Dig. Proc.*, pp. 25-29.

[53] D. Androutsos *et al.*, *Proc. of ICIP'98*, Vol. II, pp. 770-774.

[54] K. Sobottka *et al.*, *Proc. of the 2nd Int'l Conf. on Aut. Face and Gesture Rec.*, Oct. 1996, pp. 236-241.

[55] G. Guo *et al.*, *Proc. of ICIP'98*, Vol. III, pp. 299-302.

[56] E. Saber *et al.*, *Proc. of the SPIE'95*, pp. 72-80.

[57] L. J. Liu *et al.*, *Proc. of the SPIE'94*, pp. 773-778.

[58] L. Lucchese *et al.*, *Proc. of MMSP'98*, pp. 33-38.

[59] D.K. Panjwani *et al.*, *IEEE Trans. on PAMI*, Vol. 17, No. 10, pp. 939-954, Oct. 1995.

[60] J. Liu *et al.*, *IEEE Trans. on PAMI*, Vol. 16, No. 7, pp. 689-700, July 1994.

[61] M. Celenk *et al.*, *Proc. of the 29th Southeastern Symp. on Syst. Theory*, Mar. 1997, pp. 483-487.

[62] V.A. Christopoulos *et al.*, *Proc. of the 13th Int'l Conf. on Dig. Sig. Proc.*, July 1997, Vol. 2, pp. 861-864.

[63] S. Ji *et al.*, *Proc. of ICIP'98*, Vol. I, pp. 80-83.

[64] L. Shafarenko *et al.*, *IEEE Trans. on IP*, Vol. 6, No. 11, pp. 1530-1544, Nov. 1997.

[65] K. Saarinen, *Proc. of ICIP'94*, Vol. III, pp. 1021-1025.

[66] T. Gevers *et al.*, *Proc. of the 12th Int'l Conf. on Patt. Rec.*, Oct. 1994, Vol. 1, pp. 342-346.

[67] T. Gevers *et al.*, *Proc. of CVPR'97* pp. 1021-1026.

[68] R. Schettini *et al.*, *Proc. of EUSIPCO'94*, Vol. 1, pp. 26-29.

[69] S. Itoh *et al.*, *Proc. of EUSIPCO'96*, Vol. III, pp. 1869-1872.

[70] M. Barni *et al.*, *Proc. of EUSIPCO'96*, Vol. III, pp. 1725-1728.

[71] I. Pollak *et al.*, *Proc. of ICIP'98*, Vol. III, pp. 246-248.

[72] J. Gao *et al.*, *Proc. of ICIP'98*, Vol. III, pp. 823-827.

[73] I. Pollak *et al.*, *Scale-Space Theory in Computer Vision*, B. ter Haar Romeny *et al.*Eds., Springer, Berlin, 1997.

[74] P.J. Burt *et al.*, *IEEE Trans. on SMC*, Vol. 11, No. 12, pp. 802-809, Dec. 1981.

[75] V. Lozano *et al.*, *Proc. of Southwest Symp. on Im. Anal. and Interp.*, Apr. 1996, pp. 13-17.

[76] F. Ziliani *et al.*, *Proc. of ICIP'98*, Vol. III, pp. 303-307.

[77] A. Tremeau *et al.*, *Patt. Rec.*, Vol. 30, No. 7, pp. 1191-1204, July 1997.

[78] Y. Kanai, *Proc. of the SPIE'98*, pp. 709-720.

[79] B. Cramariuc *et al.*, *Proc. of the 13th Int'l Conf. on Dig. Sig. Proc.*, July 1997, Vol. 2, pp. 857-860.

[80] Y. Deng *et al.*, "Color Image Segmentation," *Proc. of CVPR'99*.

[81] V. Rehrmann *et al.*, *Proc. of ACCV'98*, Vol. 1, pp. 598-606.

[82] N. Ikonomakis *et al.*, *Canadian Jour. of Elec. and Comp. Eng.*, Vol. 23, No. 1-2, IEEE Canada, Jan.-Apr. 1998, pp. 43-47.

[83] A. Moghaddamzadeh *et al.*, *Proc. of the 3rd Int'l Fuzzy Syst. Conf.*, June 1994, Vol. 1, pp. 83-88.

[84] A. Moghaddamzadeh *et al.*, *Patt. Rec.*, Vol. 30, No. 6, pp. 867-881, June 1997.

[85] T. Carron *et al.*, *Proc. of EUSIPCO'96*, Vol. 3, pp. 1729-1732.

[86] P. Colantoni *et al.*, *Proc. of the 6th Int'l Conf. on Im. Proc. and Its Appl.*, July 1997, Vol. 2, pp. 698-702.

[87] E. Kreyszig, *Differential Geometry*, Univ. of Toronto Press, Toronto, 1959.

[88] M. Chapron, *Proc. of ICIP'97*, Vol. III, pp. 18-21.

[89] D. Zugaj *et al.*, *Patt. Rec.*, Vol. 31, No. 2, pp. 105-113, Feb. 1998.

[90] T. Carron *et al.*, *Proc. of ICIP'94*, Vol. III, pp. 977-981.

[91] H. Tao *et al.*, *Proc. of ICIP'97*, Vol. 1, pp. 834-836.

[92] W.Y. Ma *et al.*, *Proc. of CVPR'97*, pp. 744-749.

[93] F. Perez *et al.*, *Int'l Jour. of Comp. Vision*, Vol. 12, No. 1, pp. 17-42, Feb. 1994.

[94] R. Natowicz *et al.*, *Proc. of Int'l Conf. on Artif. Neural Net. and Gen. Alg.*, April 1995, pp. 360-363.

[95] R. Natowicz, *Proc. of Int'l Work. on Artif. Neural Net.*, June 1995, pp. 890-897.

[96] M. Kass *et al.*, *Int'l Jour. of Comp. Vision*, Vol. 1, pp. 321-331, 1988.

[97] G. Sapiro, *Proc. of ICIP'96*, Vol. I, pp. 817-820.

[98] G. Sapiro, *Comp. Vis. and Image Underst.*, Vol. 68, No. 2, pp. 247-253, Nov. 1997.

[99] T. Gevers *et al.*, *Proc. of the 9th British Mach. Vision Conf.*, Sept. 1998, Vol. 2, pp. 578-588.

[100] P. Campadelli *et al.*, *Image and Vision Computing Image*, Vol. 15, No. 3, pp. 161-166, Mar. 1997.

[101] C.-L. Huang, *Patt. Rec. Lett.*, Vol. 13, pp. 345-353, 1993.

[102] S. Rachid *et al.*, *Proc. of ICIP'97*, Vol. I, pp. 243-246.

[103] H. Iwata *et al.*, *Syst. and Comp. in Japan*, Vol. 29, No. 4, pp. 1-10, Apr. 1998.

[104] H. Okii *et al.*, *IEICE Trans. on Inf. and Syst. (Japan)*, Vol. E77-D, No. 3, pp. 343-350, Mar. 1994.

[105] N. Funakubo, *Proc. of IECON'94*, Vol. 2, pp. 852-856.

[106] A. Verikas *et al.*, *Elektronika ir Elektrotechnika (Lithuania)*, No. 4, pp. 7-13, 1996.

[107] A. Verikas *et al.*, *Patt. Rec. Lett.*, Vol. 18, No. 2, pp. 173-185, Feb. 1997.

[108] K.A. McCrae *et al.*, *Proc. of the SPIE'94*, pp. 306-315.

[109] E. Ardizzone *et al.*, *Proc. of Int'l Conf. on Artif. Neural Net.*, May 1994, Vol. 2, pp. 1161-1164.

[110] S.A. Shafer, *Color Res. and Appl.*, Vol. 10, No. 4, pp. 210-218, 1985.

[111] G.J. Klinker *et al.*, *Int'l Jour. of Comp. Vision*, Vol. 4, pp. 7-38, 1990.

[112] A. Mogahddamzadeh *et al.*, *Proc. of the SPIE'95*, pp. 300-309.

[113] M. Aizu *et al.*, *Proc. of DCC'95*, pp. 439.

[114] N.H. Kroupnova, *Proc. of the SPIE'96*, pp. 444-453.

[115] C.-K. Ong *et al.*, *Proc. of the 14th Int'l Conf. on Patt. Rec.*, Aug. 1998, Vol. 1, pp. 780-784.

[116] P.W.M. Tsang *et al.*, *Proc. of ICIP'96*, Vol. III, pp. 1049-1052.

[117] R. Bajcsy *et al.*, *Int'l Jour. of Comp. Vision*, Vol. 17, No. 3, pp. 241-272, Mar. 1996.

[118] G. Healey, *Jour. of the Opt. Soc. of Am. - A*, Vol. 6, No. 6, pp. 920-937, June 1989.

[119] G. Healey, *IEEE Trans. on SMC*, Vol. 22, No. 1, pp. 64-73, Jan.-Feb. 1992.

[120] B.A. Maxwell *et al.*, *Comp. Vis. and Image Underst.*, Vol. 65, No. 2, pp. 269-295, Feb. 1997.

[121] M. Borsotti *et al.*, *Patt. Rec. Lett.*, Vol. 19, pp. 741-747, 1998.