PROBABILITY MODELS FOR CONNECTED OPERATORS

D. Gatica-Perez, M-T. Sun
Department of Electrical Engineering, University of Washington.
Box 352142. Seattle, WA, 98195, USA.

C. Gu
Microsoft Corporation.
One Microsoft Way. Redmond, WA 98052, USA.

Abstract Connected operators filter images by selectively altering the intensity of connected sets of pixels, without introducing new image contours. While in algebraic terms, connected operators are defined as operations in lattices of functions, and have generally been formulated on a deterministic basis, they also correspond to hierarchical agglomerative clustering algorithms, which are known to have well defined probabilistic formulations. In this paper, we argue that a probabilistic approach both generalizes existing connected operators and allows for new designs. In particular, we introduce an operator based on the development of models of visual similarity among image regions, and sequential Maximum a Posteriori (MAP) binary classification. Experimental results show that such operator is useful for content-based simplification of image collections and videos.

Keywords: Connected Operators, Hierarchical Clustering, Classification, Mixture Models.

1. Introduction

Connected operators [16] are non-linear filters that transform an image by changing the intensity of connected sets of pixels of constant color, called *flat zones*. The flat zones of an image induce a partition of its support. A connected operator coarsens such partition without introducing any new intensity discontinuities. This characteristic makes connected operators attractive for applications in which contour preservation is necessary, like image simplification and segmentation [19], [4], [8], [13], [10], [15]. In particular, content-based image and video applications, as discussed in the MPEG-7 standard, deal with visual content at the region and object levels, and would then be benefitted by

Redistribution rights reserved CSIRO Publishing. ISBN 0 643 06804 X

the development of techniques to generate compact region-based image representations [2], [17].

As an algorithm, a connected operator consists on two transformations: (i) *merging* and (ii) *re-coloring* of an image flat zones [16], [11]. However, the definition of connected operator does not specify the mechanisms by which the colors of an image and its flat-zone partition should be modified. Different classes of operators have been proposed either by imposing theoretical properties (extensiveness, idempotence, increasingness), or by designing them to perform a particular function (filtering of small, elongated, or dark/bright regions).

With a few exceptions [18], connected operators have been formulated in deterministic terms. In this paper, we argue that the introduction of probability models in the design of connected operators is useful both to generalize existing formulations and to define new designs [1], [3], [6]. In particular, we introduce an operator based on the formulation of hierarchical clustering as a sequential binary classification process, and on the development of statistical models of visual similarity among image regions. Its performance is illustrated on image collections and video sequences.

The paper is organized as follows. Section 2 reviews a general methodology to implement attribute-based connected operators. A probabilistic view of it is suggested in Section 3. Section 4 extends such concept to propose an operator based on sequential binary MAP classification. Our formulation requires the selection of a feature space, and the specification of class-conditional and prior distributions. The details are described in Section 5. Section 6 presents results on images and videos. Section 7 provides some final remarks.

2. Attribute-based connected operators

This class of operators transform images using criteria of the style "*given a certain order, merge and re-color all flat zones whose* `ListOfAttributes` $\sim$ `ListOfParameters`", where $\sim$ denotes a relation (i.e., $\leq$, $\geq$, $=$) and the attributes are measures of region size (area, or length of convex hull), shape (geodesic distance, eccentricity) [19], [4], [10], or other non-geometric features (color, motion, contrast, entropy) [8], [15]. A general merging strategy to implement attribute-based connected operators was proposed in [8], and requires three elements for its specification: a *region model*, which represents the attributes of each flat zone; a *merging order*, which defines the order in which two zones will be used as candidates for merging; and a *merging criterion* to decide whether the selected zones should be merged. The filtering process admits several variations, but it is also based on three steps: the generation of a region-based image representation, the construction of the *merging sequence* that results of the merging procedure, and the reconstruction of the image from the region-based representation. A filtered image consists of all the remaining region models, thus eliminating all components that were previously merged. The operators can be efficiently implemented with the use of adjacency graphs and hierarchical queues [8]. Two examples are shown in Fig. 1. Recent alternatives consist of the generation of the merging sequence until idempotence, followed by tree pruning [15].

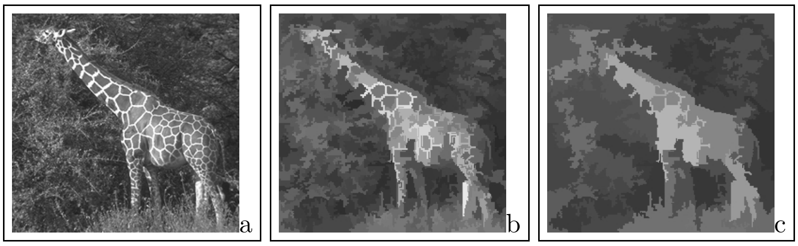

Figure 1. (a) *Giraffe.* (b-c) Filtered images with area connected operator. All components with area less than (b) 100 and (c) 1000 pixels, are removed. The merging order depends on both area (minimum size) and intensity (color difference) attributes.

In the described approach, the specification of the merging order and the merging criterion, which defines the rules for filtering are often based on experience, and can be hard to specify. An alternative to encode the a priori knowledge of the problem consists on the use of probability models [6].

3. A probabilistic view of attribute-based connected operators

Hierarchical agglomerative clustering algorithms based on probability models are increasingly used in pattern recognition [1], [3], [6]. The simplest algorithm consists of a *sequential binary classifier*, which at each step chooses a pair of regions and decides whether they should be merged according to a probability model and Bayesian decision theory [6]. Let R_i and R_j denote the i-th and j-th regions of an image, and let $\mathcal{E}$ be an r.v. that indicates whether such regions belong to the same final component,

$$\mathcal{E} = \begin{cases} 1 & \text{if } \Omega(R_i) = \Omega(R_j) \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where $\Omega(R_i)$ is the component that contains R_i. The MAP criterion establishes that given some knowledge about the world $\mathcal{I}$, and a realization x_{ij} of X (representing the features extracted from R_i and R_j), the value of $\mathcal{E}$ that must be selected is

$$\mathcal{E}^* = \arg\max_{\mathcal{E}} \Pr(\mathcal{E}|x, \mathcal{I}), \tag{2}$$

where $\Pr(\mathcal{E}|x, \mathcal{I})$ is the posterior probability mass function. By Bayes' rule,

$$p(x|\mathcal{E} = 1, \mathcal{I}) \Pr(\mathcal{E} = 1|\mathcal{I}) \underset{H_1}{\overset{H_0}{\lessgtr}} p(x|\mathcal{E} = 0, \mathcal{I}) \Pr(\mathcal{E} = 0|\mathcal{I}), \tag{3}$$

where $p(x|\mathcal{E}, \mathcal{I})$ are the class-conditional probability density functions of the observed features, $\Pr(\mathcal{E}|\mathcal{I})$ are the priors of $\mathcal{E}$, and H_1 (resp. H_0) denotes the hypothesis that the pair of regions should be merged (resp. not merged).

Attribute-based connected operators have a straightforward formulation in these terms. The area operator described in [8], which removes all regions that are smaller than an *area parameter* λ, can be used as an example. The merging criterion states that two regions have to be merged if at least one of them is smaller than λ. Let $x_{ij} = (|R_i|, |R_j|)$, denote the areas of regions R_i and R_j. If we model the class-conditional pdfs by

$$p(x|\mathcal{E} = 0, \lambda < |E|/2) = \begin{cases} \frac{1}{|\Delta_\lambda|} & \text{if } x \in \Delta_\lambda \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

$$p(x|\mathcal{E} = 0, \lambda \geq |E|/2) = \delta(|R_i| - \frac{|E|}{2}, |R_j| - \frac{|E|}{2}), \tag{5}$$

$$p(x|\mathcal{E} = 1) = \begin{cases} \frac{1}{|\Delta^c_\lambda|} & \text{if } x \in \Delta^c_\lambda \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

where Δ_λ, and Δ^c_λ denote the regions shown in Fig. 2, $|E|$ denotes the area of the entire image support, and if we further assume non-informative priors, then Eq. 3 implements the merging criterion as specified in [8]. Furthermore, the specification of a merging order and a region model are responsible for the particular re-coloring operation. Furthermore, other attribute-based connected operators admit a similar reformulation, with multivariate distributions representing multiple attributes [4].

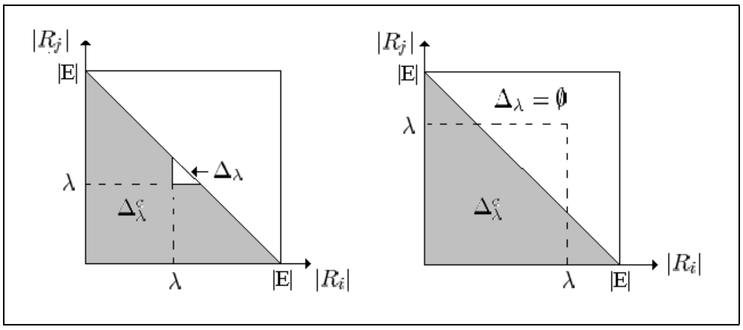

Figure 2. Support for class-conditional pdfs for area connected operator. The two cases correspond to $\lambda \leq \frac{|E|}{2}$, and $\lambda > \frac{|E|}{2}$. In the former case, $\Delta_\lambda = \{x \in \mathcal{R}^2$ such that $\min(|R_i|, |R_j|) > \lambda$ and $|R_i| + |R_j| \leq |E|\}$, so $|\Delta_\lambda| = (|E| - 2\lambda)^2/2$. In the latter case, $\Delta_\lambda = \emptyset$.

4. Connected operators based on sequential MAP binary classification

We propose to build operators for which both the merging criterion and the merging order are model-based, by sequentially evaluating the pair of connected regions that yield the *largest binary classification likelihood ratio*,

$$L = \frac{p(x|\mathcal{E} = 1, \mathcal{I}) \Pr(\mathcal{E} = 1|\mathcal{I})}{p(x|\mathcal{E} = 0, \mathcal{I}) \Pr(\mathcal{E} = 0|\mathcal{I})}, \tag{7}$$

merging whenever $L \geq 1$, updating the region-based representation, and iterating the process until H_1 in Eq. 3 is not longer valid for any pair of connected regions. This greedy, sub-optimal merging strategy does not require heuristic parameter determination, and bears similarity to the Highest Confidence First (HCF) method used in Bayesian image analysis [5]. Implementation details of the methodology can be found in [9].

Some properties of this class of operators can be established.

Property 1 *Let $\phi : Fun(\cdot) \rightarrow Fun(\cdot)$ denote a connected operator based on sequential MAP binary classification, and $Fun(\cdot)$ denotes a complete lattice of multi-valued functions. Then ϕ is (i) idempotent, (ii) not-invertible, (iii) not injective, (iv) neither increasing nor decreasing in general, and (v) neither extensive nor anti-extensive in general.*

5. Probability Models

The reformulation of filtering as a sequential binary classification problem requires the determination of a useful feature space, and the selection of models for the distributions. In the first place, multivariate features defined on pairs of regions could represent regional differences of color, texture, and other attributes [9]. In the second place, a parametric representation is often accurate for feature spaces of low dimension [6].

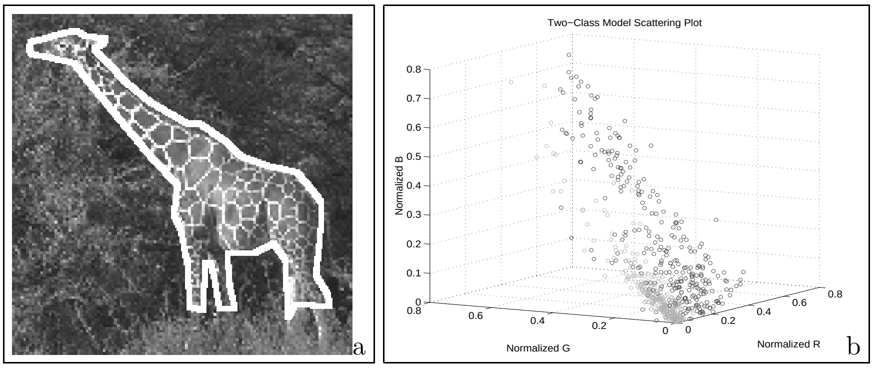

Figure 3. (a) *Giraffe.* (b) Scatter plot for the 3-D color two-class model. A random sample of 300 vectors from each class is displayed. Intra-object and inter-object feature vectors are represented by light and dark gray circles, respectively.

5.1 Generating data

The described methodology is useful whenever training data are available. This is usually the case of video sequences or image databases, where a subset of images could be used to learn the models, and subsequently used to filter an entire video or a set of still images of similar content [17]. An example is shown in Fig. 3(a). The scene has been manually partitioned into two objects (foreground/background). To investigate the behavior of visual similarity among pairs of regions, a fine image partition was generated, and the mean RGB value of each region was computed. Then, the RGB difference between every pair of regions belonging to the same object and located within a spatial window was computed. The same procedure was repeated for pairs of regions that do not belong to the same object. Fig. 3(b) shows a random sample of intra- and inter-object feature vectors, labeled with light and dark gray, respectively. More elaborate features would provide higher discrimination.

5.2 Modeling Likelihood Functions with Mixture Models

The full joint class-conditional pdfs of the observed features are represented by a multivariate Gaussian mixture model,

$$p(x|\mathcal{E},\Theta,\mathcal{I}) = \sum_{i=1}^{N_{\mathcal{E}}} \omega_i p(x|\mathcal{E},\theta_i,\mathcal{I}), \tag{8}$$

where $N_{\mathcal{E}}$ is the number of components in each mixture, ω_i denotes the prior probability of the i-th component, $p(x|\mathcal{E},\theta_i,\mathcal{I}) = \mathcal{N}(\mu_i,\Sigma_i)$ is the i-th d-dimensional Gaussian with full covariance matrix, parameterized by $\theta_i = \{\mu_i,\Sigma_i\}$,

$$p(x|\mathcal{E},\theta_i,\mathcal{I}) = \frac{1}{(2\pi)^{d/2}|\Sigma_i|^{1/2}} e^{-\frac{1}{2}(x-\mu_i)^T\Sigma_i^{-1}(x-\mu_i)}, \tag{9}$$

and $\Theta = \{\{\omega_i\},\{\theta_i\}\}$ represents the set of all parameters. The Expectation-Maximization (EM) algorithm constitutes the standard procedure for Maximum Likelihood estimation (ML) of a set of parameters [12]. EM is useful for a broad range of problems where the observed data is in some sense incomplete. In the case of a Gaussian mixture, the incomplete data are the unobserved mixture components, with prior distribution $\{\omega_i\}$. EM is based on increasing the conditional expectation of the log-likelihood of the complete data given the observed data by using an iterative hill-climbing procedure.

Additionally, model selection can be automatically estimated using the Minimum Description Length (MDL) principle [14], by choosing

$$N_{\mathcal{E}}^* = \arg\max_{N_{\mathcal{E}}} (\log L(\Theta|\tilde{X}) - \frac{n_{N_{\mathcal{E}}}}{2}\log N), \tag{10}$$

where $L(\cdot)$ denotes the likelihood of the training set, N is the number of training vectors, $\tilde{X}$ is the training set, and $n_{N_{\mathcal{E}}}$ is the number of model parameters,

$$n_{N_{\mathcal{E}}} = (N_{\mathcal{E}}-1) + N_{\mathcal{E}}d + N_{\mathcal{E}}\frac{d(d+1)}{2}. \tag{11}$$

When two models fit the data in a similar way, the simpler model is chosen.

5.3 Modeling of Priors

In the Bayesian approach, the prior probability mass function $\Pr(\mathcal{E}|\mathcal{I})$ encodes the previous knowledge or belief about the merging process characteristics [6]. The simplest assumption is a uniform prior, $\Pr(\mathcal{E}=0|\mathcal{I}) = \Pr(\mathcal{E}=1|\mathcal{I})$, and it has been used in the results presented in the next section. Note that priors themselves could also be ML-estimated from data, although in strict terms, this technique does not conform the Bayesian principle [6].

6. Results

Fig. 4 illustrates the performance of the proposed methodology for the *Giraffe* image. Using the regional mean in RGB space as region model, and training data as described in Section 5, the EM-estimated mixture pdfs consist of six components for each class. After a post-processing stage that sieves small isolated components, the filtered image has been nicely simplified while preserving the major object segments. Of course, the generalization of the learned model to a different set of images is more interesting. Fig. 4 shows the results obtained on two images of somewhat similar color content, using the color model generated for *Giraffe*. The operator has simplified the images, significantly reducing the number of flat zones, but capturing the main components.

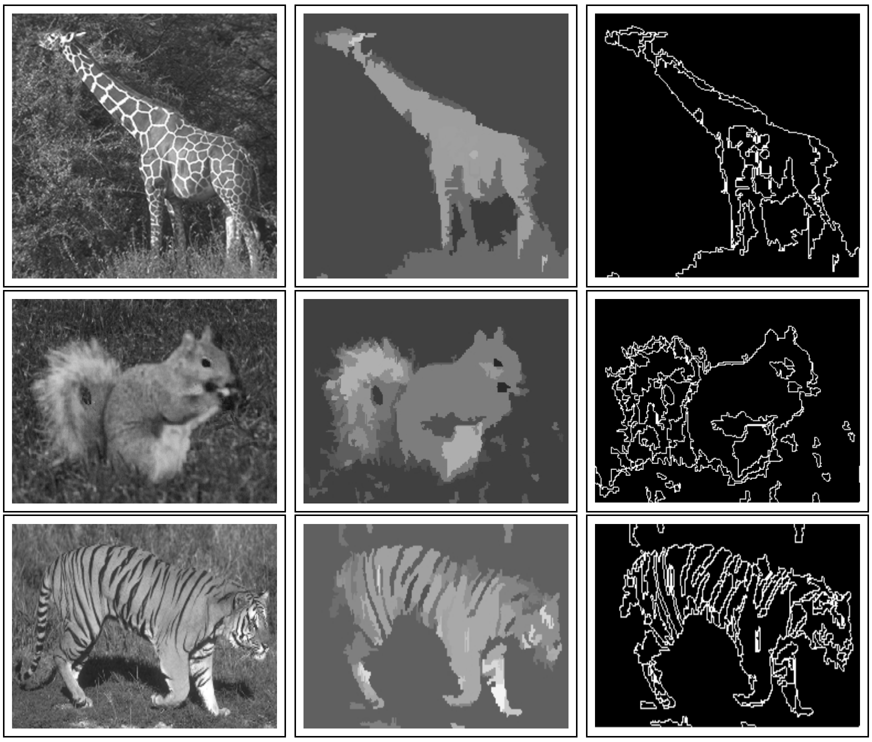

Figure 4. Left column: *Giraffe*, *Squirrel*, and *Tiger* images. Center column: Filtered images with model-based connected operator, based on the models estimated for *Giraffe* (24, 52, and 70 flat zones, respectively). Right column: Flat-zone partitions of filtered images.

Another example is shown in Fig. 5 for the *Zebras* image set. One single image was used for training, defining the median in RGB space as region model. The operator has simplified the images based on the characteristics of the probability models, and produced a small number of perceptually meaningful regions.

The methodology is also suitable for video filtering. Fig. 6 shows results for the *Bream* sequence. The mixture models (eight and six Gaussian components for each class, respectively) were estimated from the first frame, and then used to filter the entire sequence. Similar results were obtained. A final example on the *Kid* sequence, extracted from a home video database, is shown in Fig. 7.

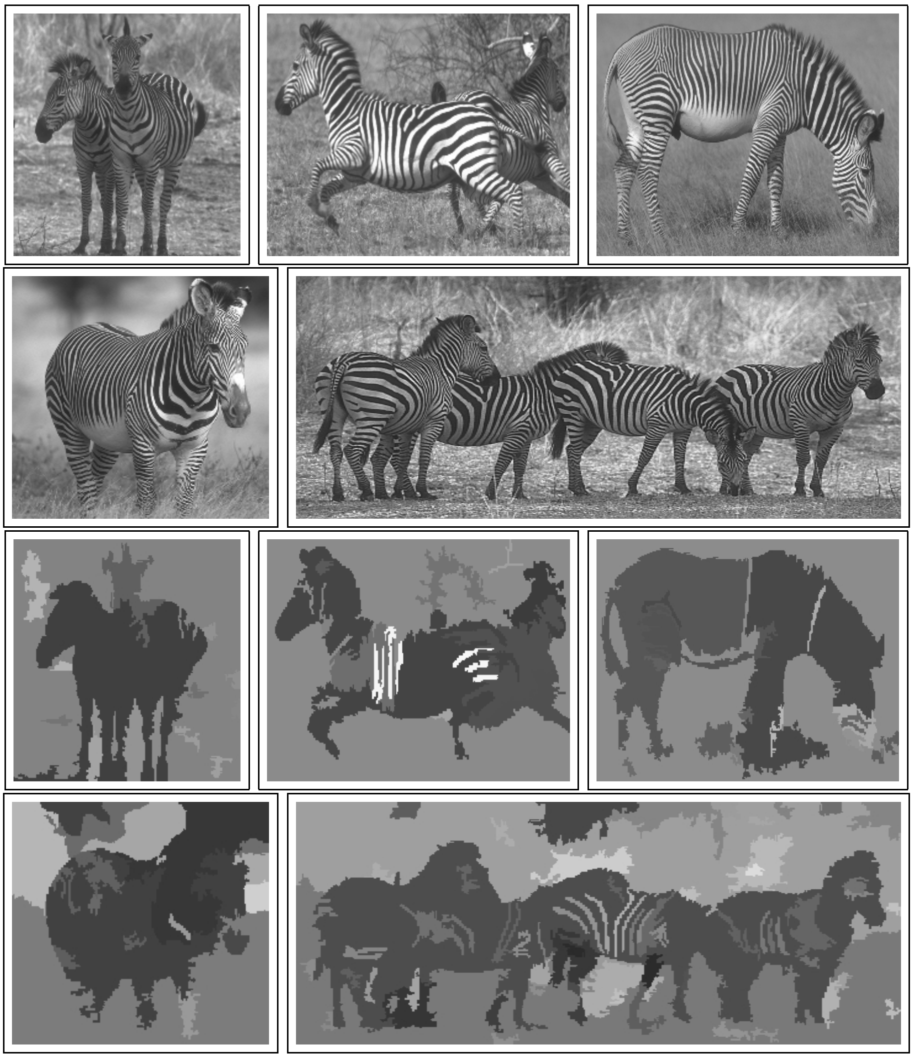

Figure 5. First and second row: *Zebras.* The top-left frame was partitioned into foreground/background, and used to estimate model parameters. Third and fourth row: Filtered images with model-based connected operator (18, 19, 19, 19, and 70 flat zones, respectively).

Although imperfect in terms of highly accurate segmentation, we believe that the filtered images, composed of a small number of visually coherent regions, could be used as a region-oriented image representation in a content-based system for indexing and retrieval of visual information [2], [17]. Furthermore, we expect that the use of better features would improve the performance.

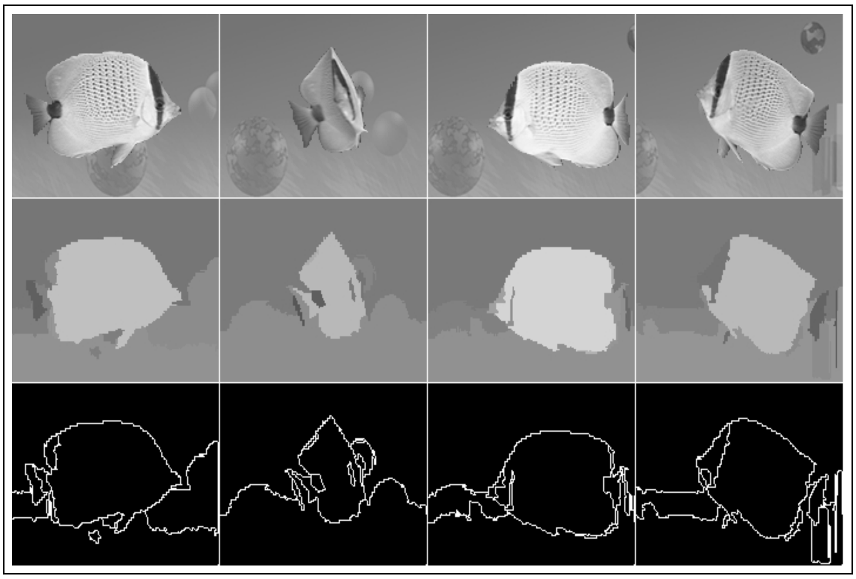

Figure 6. Top row: *Bream* image sequence, frames 0, 115, 134, and 193. The first frame was partitioned into foreground/background, and used to estimate model parameters. Center row: Filtered images with model-based connected operator (11, 9, 10, and 11 flat zones, respectively). Bottom row: Flat-zone partitions of filtered images.

7. Concluding remarks

This paper presented a methodology to design attribute-based connected operators using probability models. Based on a reformulation of previous approaches, we introduced an operator based on sequential Maximum a Posteriori (MAP) binary classification. The results suggest that, when a scene contains distinctive objects whose variability can be captured by the models, the proposed operator is effective to simplify images and sequences on a content-based fashion.

The operator presented here is one out of a whole class of connected operators that could be designed with the use of probability models [3], [6]. On one hand, approximations to global classification represent alternatives to the local classification concept described in this paper [7]. On the other hand, a comparative study between these techniques and parallel classification strategies [2] is worth it to be performed. Both issues constitute current lines of work [9].

Figure 7. Top row: *Kid* home video sequence, frames 0, 140, 180, and 200. The first frame was partitioned into foreground/background, and used to estimate model parameters. Bottom row: Filtered images with model-based connected operator (27, 32, 40, and 21 flat zones, respectively).

Acknowledgments

The wildlife images belong to the Corel stock photo collection ©. The *Kid* sequence belongs to the Kodak Home Video Database ©.

References

[1] J. D. Banfield and A. Raftery, "Model-based Gaussian and Non-Gaussian Clustering," *Biometrics*, Vol. 49, No. 3, pp. 803-821, Sept. 1993.

[2] S. Belongie, C. Carson, H. Greenspan, and J. Malik, "Color and Texture Image Segmentation Using the Expectation-Maximization Algorithm and Its Application to Content-Based Image Retrieval,", In *Proc. IEEE Int. Conf. Comp. Vis.*, Bombay, Jan. 1998.

[3] H. H. Bock, "Probabilistic Models in Cluster Analysis," *Computational Statistics and Data Analysis*, Vol. 23, pp. 5-28.

[4] E.J. Breen and R. Jones, "Attribute Openings, Thinnings, and Granulometries," *Computer Vision and Image Understanding*, Vol. 64, No. 3, Nov. 1996. pp. 377-389.

[5] P. Chou and C. Brown, "The Theory and Practice of Bayesian Image Labeling," *International Journal of Computer Vision*, Vol. 4, pp. 185-210, 1990.

[6] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd. Ed.. Wiley and Sons, 2000.

[7] C. Fraley and A. Raftery, "Model-Based Clustering, Discriminant Analysis, and Density Estimation," University of Washington, Dept. of Stats., Tech. Rep. No. 380, Oct. 2000.

[8] L. Garrido, P. Salembier and D. Garcia, "Extensive Operators in Partition Lattices for Image Sequence Analysis," *Signal Processing*, 66(2), pp. 157-180, Apr. 1998.

[9] D. Gatica-Perez, M.-T. Sun, and C. Gu, "Model-based Connected Operators," University of Washington, Dept. of Electrical Engineering, Technical Report, in preparation.

[10] R. Jones, “Connected Filtering and Segmentation Using Component Trees,” *Computer Vision and Image Understanding*, Vol. 75, No. 3, pp. 215-228, Sept. 1999.

[11] H.J.A.M. Heijmans, “Connected Morphological Operators for Binary Images,” *Computer Vision and Image Understanding*, Vol. 73, No. 1, p.99-120,1999.

[12] G.J. MacLachlan and D. Peel. *Finite Mixture Models.* Wiley and Sons, 2000.

[13] F. Meyer. From connected operators to levelings. in *Proc. Int. Symp. on Math. Morphology,* Dordrecht, pp. 191-198, 1998.

[14] J. Rissanen, “Modeling by shortest data description,” *Automatica*, 14: 465-471, 1978.

[15] P. Salembier, and L. Garrido, “Connected Operators Based on Region-Tree Pruning,” in *Proc. Int. Symp. on Math. Morphology,* Palo Alto, pp. 169-178, 2000.

[16] J. Serra and P. Salembier, “Connected Operators and Pyramids,” In *Proc. of SPIE Image Algebra and Morphological Image Processing IV*, San Diego, 1993, pp. 65-76.

[17] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, “Content-Based Image Retrieval at the End of the Early Years,” *IEEE Trans. on Pattern Analysis And Machine Intelligence*, Vol. 22, No. 12, pp. 1349-1380, Dec. 2000.

[18] V. Vilaplana, F. Marques, P. Salembier, and L. Garrido, “Region-based segmentation and tracking of human faces,” in *Proc. EUSIPCO*, Rhodes, pp. 311-314, Sept. 1998.

[19] L. Vincent, “Grayscale Area Openings and Closings, Their Efficient Implementation and Applications,” in *Proc. Int. Symp. on Math. Morphol.,* Barcelona, pp. 22-27, 1993.