Mean Shift Analysis and Applications

Dorin Comaniciu Peter Meer
Department of Electrical and Computer Engineering
Rutgers University, Piscataway, NJ 08854-8058, USA
{comanici, meer}@caip.rutgers.edu

Abstract

A nonparametric estimator of density gradient, the mean shift, is employed in the joint, spatial-range (value) domain of gray level and color images for discontinuity preserving filtering and image segmentation. Properties of the mean shift are reviewed and its convergence on lattices is proven. The proposed filtering method associates with each pixel in the image the closest local mode in the density distribution of the joint domain. Segmentation into a piecewise constant structure requires only one more step, fusion of the regions associated with nearby modes. The proposed technique has two parameters controlling the resolution in the spatial and range domains. Since convergence is guaranteed, the technique *does not* require the intervention of the user to stop the filtering at the desired image quality. Several examples, for gray and color images, show the versatility of the method and compare favorably with results described in the literature for the same images.

1 Introduction

Low level computer vision tasks are misleadingly difficult and often yield unreliable results, since the employed techniques rely upon the correct choice by the user of the tuning parameter values. Today, it is an accepted fact in the vision community that the execution of low level tasks should be task driven, i.e., supported by independent high level information. To be able to successfully complement this paradigm, the low-level techniques must become more autonomous. In this paper we propose such a technique for image smoothing and for segmentation.

The *mean shift estimate* of the gradient of a density function and the associated iterative procedure of mode seeking have been developed by Fukunaga and Hostetler in [7]. Only recently, however, the nice properties of data compaction and dimensionality reduction of the mean shift have been exploited in low level computer vision tasks (e.g., color space analysis [4], face tracking [1]).

In this paper we describe a new application based on the theoretical results obtained in [5]. We show that high quality edge preserving filtering and image segmentation can be obtained by applying the mean shift in the combined spatial-range domain. The methods we developed are conceptually very simple being based on the same idea of iteratively shifting a fixed size window to the average of the data points within. Details in the image are preserved due to the nonparametric character of the analysis which does not assume a priori any particular structure for the data.

The paper is organized as follows. Section 2 discusses the estimation of the density gradient and defines the mean shift vector. The convergence of the mean shift procedure is proven in Section 3 for discrete data. Section 4 defines the processing principle in the joint spatial-range domain. Mean shift filtering is explained and filtering examples are given in Section 5. The proposed mean shift segmentation is introduced and analyzed in Section 6.

2 Density Gradient Estimation

Let $\{\mathbf{x}_i\}_{i=1\ldots n}$ be an arbitrary set of n points in the d-dimensional Euclidean space R^d. The *multivariate kernel density estimate* obtained with kernel $K(\mathbf{x})$ and window radius h, computed in the point $\mathbf{x}$ is defined as [13, p.76]

$$\hat{f}(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{\mathbf{x}-\mathbf{x}_i}{h}\right). \tag{1}$$

The optimum kernel yielding minimum mean integrated square error (MISE) is the Epanechnikov kernel

$$K_E(\mathbf{x}) = \begin{cases} \frac{1}{2}c_d^{-1}(d+2)(1-\mathbf{x}^T\mathbf{x}) & \text{if } \mathbf{x}^T\mathbf{x} < 1 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where c_d is the volume of the unit d-dimensional sphere [13, p.76].

The use of a differentiable kernel allows to define the estimate of the density gradient as the gradient of the kernel density estimate (1)

$$\hat{\nabla} f(\mathbf{x}) \equiv \nabla \hat{f}(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^{n} \nabla K\left(\frac{\mathbf{x}-\mathbf{x}_i}{h}\right). \tag{3}$$

Conditions on the kernel $K(\mathbf{x})$ and the window radius h to guarantee asymptotic unbiasedness, mean-square consistency, and uniform consistency are derived in [7].

For the Epanechnikov kernel (2) the density gradient estimate (3) becomes

$$\hat{\nabla} f(\mathbf{x}) = \frac{1}{n(h^d c_d)} \frac{d+2}{h^2} \sum_{\mathbf{x}_i \in S_h(\mathbf{x})} [\mathbf{x}_i - \mathbf{x}]$$

$$= \frac{n_{\mathbf{x}}}{n(h^d c_d)} \frac{d+2}{h^2} \left(\frac{1}{n_{\mathbf{x}}} \sum_{\mathbf{x}_i \in S_h(\mathbf{x})} [\mathbf{x}_i - \mathbf{x}] \right) \quad (4)$$

where the region $S_h(\mathbf{x})$ is a hypersphere of radius h having the volume $h^d c_d$, centered on $\mathbf{x}$, and containing $n_{\mathbf{x}}$ data points. The last term in (4)

$$M_h(\mathbf{x}) \equiv \frac{1}{n_{\mathbf{x}}} \sum_{\mathbf{x}_i \in S_h(\mathbf{x})} [\mathbf{x}_i - \mathbf{x}] = \frac{1}{n_{\mathbf{x}}} \sum_{\mathbf{x}_i \in S_h(\mathbf{x})} \mathbf{x}_i - \mathbf{x} \quad (5)$$

is called the *sample mean shift*. Using a kernel different from the Epanechnikov kernel results in a weighted mean computation in (5).

The quantity $\frac{n_{\mathbf{x}}}{n(h^d c_d)}$ is the kernel density estimate $\hat{f}(\mathbf{x})$ computed with the hypersphere $S_h(\mathbf{x})$ (the uniform kernel), and thus we can write (4) as

$$\hat{\nabla} f(\mathbf{x}) = \hat{f}(\mathbf{x}) \frac{d+2}{h^2} M_h(\mathbf{x}), \quad (6)$$

which yields

$$M_h(\mathbf{x}) = \frac{h^2}{d+2} \frac{\hat{\nabla} f(\mathbf{x})}{\hat{f}(\mathbf{x})}. \quad (7)$$

The expression (7) was first derived in [7] and shows that an estimate of the normalized gradient can be obtained by computing the sample mean shift in a uniform kernel centered on $\mathbf{x}$. The mean shift vector has the direction of the gradient of the density estimate at $\mathbf{x}$ when this estimate is obtained with the Epanechnikov kernel.

Since the mean shift vector always points towards the direction of the maximum increase in the density, it can define a path leading to a local density maximum, i.e., to a mode of the density (Figure 1).

The *mean shift procedure*, obtained by successive

- computation of the mean shift vector $M_h(\mathbf{x})$
- translation of the window $S_h(\mathbf{x})$ by $M_h(\mathbf{x})$,

is guaranteed to converge, as it will be shown in the next section.

3 Convergence

Let $\{\mathbf{y}_k\}_{k=1,2\ldots}$ denote the sequence of successive locations of the mean shift procedure. By definition we have for each k=1,2...

$$\mathbf{y}_{k+1} = \frac{1}{n_k} \sum_{\mathbf{x}_i \in S_h(\mathbf{y}_k)} \mathbf{x}_i, \quad (8)$$

where $\mathbf{y}_1$ is the center of the initial window and n_k is the number of points falling in the window $S_h(\mathbf{y}_k)$ centered on $\mathbf{y}_k$.

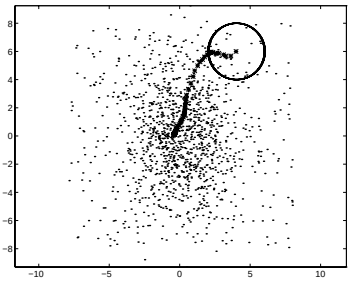


Figure 1: Successive computations of the mean shift define a path leading to a local density maximum.

The convergence of the mean shift has been justified as a consequence of relation (7), (see [2]). However, while it is true that the mean shift vector $M_h(\mathbf{x})$ has the direction of the gradient of the density estimate at $\mathbf{x}$, it is not apparent that the density estimate at locations $\{\mathbf{y}_k\}_{k=1,2\ldots}$ is a monotonic increasing sequence. Moving in the direction of the gradient guarantees hill climbing only for infinitesimal steps. The following theorem asserts the convergence for discrete data.

Theorem 1 *Let $\hat{f}_E = \left\{ \hat{f}_k(\mathbf{y}_k, K_E) \right\}_{k=1,2\ldots}$ be the sequence of density estimates obtained using Epanechnikov kernel and computed in the points $\{\mathbf{y}_k\}_{k=1,2\ldots}$ defined by the successive locations of the mean shift procedure with uniform kernel. The sequence is convergent.*

Proof Since the data set $\{\mathbf{x}_i\}_{i=1\ldots n}$ has finite cardinality n, the sequence $\hat{f}_E$ is bounded. Moreover, we will show that $\hat{f}_E$ is strictly monotonic increasing, i.e., if $\mathbf{y}_k \neq \mathbf{y}_{k+1}$ then $\hat{f}_E(k) < \hat{f}_E(k+1)$, for all $k = 1, 2 \ldots$.

Let n_k, $n_k^{'}$, and $n_k^{''}$ with $n_k = n_k^{'} + n_k^{''}$ be the number of data points falling in the d-dimensional windows (Figure 2) $S_h(\mathbf{y}_k)$, $S_h{}'(\mathbf{y}_k) = S_h(\mathbf{y}_k) - S_h{}''(\mathbf{y}_k)$, and $S_h{}''(\mathbf{y}_k) = S_h(\mathbf{y}_k) \bigcap S_h(\mathbf{y}_{k+1})$.

Without loss of generality we can assume the origin located at $\mathbf{y}_k$. Using the definition of the density estimate (1) with the Epanechnikov kernel (2) and noting that $\|\mathbf{y}_k - \mathbf{x}_i\|^2 = \|\mathbf{x}_i\|^2$ we have

$$\begin{aligned} \hat{f}_E(k) &= \hat{f}_k(\mathbf{y}_k, K_E) \\ &= \frac{1}{nh^d} \sum_{\mathbf{x}_i \in S_h(\mathbf{y}_k)} K_E \left(\frac{\mathbf{y}_k - \mathbf{x}_i}{h} \right) \\ &= \frac{d+2}{2n(h^d c_d)} \sum_{\mathbf{x}_i \in S_h(\mathbf{y}_k)} \left(1 - \frac{\|\mathbf{x}_i\|^2}{h^2} \right). \end{aligned} \quad (9)$$

Since the kernel K_E is nonnegative we also have

$$\hat{f}_E(k+1) = \hat{f}_{k+1}(\mathbf{y}_{k+1}, K_E) \geq$$

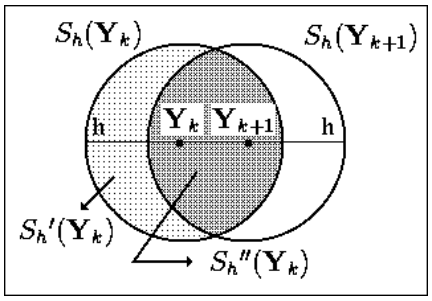


Figure 2: The d-dimensional windows used in the proof of convergence: $S_h(\mathbf{y}_k)$, $S_h{}'(\mathbf{y}_k)$, and $S_h{}''(\mathbf{y}_k)$. The point $\mathbf{y}_{k+1}$ is the mean of the data points falling in $S_h(\mathbf{y}_k)$.

$$\geq \frac{1}{nh^d} \sum_{\mathbf{x}_i \in S_h''(\mathbf{y}_k)} K_E\left(\frac{\mathbf{y}_{k+1} - \mathbf{x}_i}{h}\right)$$

$$= \frac{d+2}{2n(h^d c_d)} \sum_{\mathbf{x}_i \in S_h''(\mathbf{y}_k)} \left(1 - \frac{\|\mathbf{y}_{k+1} - \mathbf{x}_i\|^2}{h^2}\right). \quad (10)$$

Hence, knowing that $n_k^{'} = n_k - n_k^{''}$ we obtain

$$\hat{f}_E(k+1) - \hat{f}_E(k) \geq \frac{d+2}{2n(h^d c_d)h^2}$$

$$\left[\sum_{\mathbf{x}_i \in S_h(\mathbf{y}_k)} \|\mathbf{x}_i\|^2 - \sum_{\mathbf{x}_i \in S_h''(\mathbf{y}_k)} \|\mathbf{y}_{k+1} - \mathbf{x}_i\|^2 - n_k^{'} h^2\right], \quad (11)$$

where the last term appears due to the different summation boundaries.

Also, by definition $\|\mathbf{y}_{k+1} - \mathbf{x}_i\|^2 \geq h^2$ for all $\mathbf{x}_i \in S_h^{'}(\mathbf{y}_k)$, which implies that

$$\sum_{\mathbf{x}_i \in S_h^{'}(\mathbf{y}_k)} \|\mathbf{y}_{k+1} - \mathbf{x}_i\|^2 \geq n_k^{'} h^2. \quad (12)$$

Finally, employing (12) in (11) and using (8) we obtain

$$\hat{f}_E(k+1) - \hat{f}_E(k) \geq \frac{d+2}{2n(h^d c_d)h^2}$$

$$\left[\sum_{\mathbf{x}_i \in S_h(\mathbf{y}_k)} \|\mathbf{x}_i\|^2 - \sum_{\mathbf{x}_i \in S_h(\mathbf{y}_k)} \|\mathbf{y}_{k+1} - \mathbf{x}_i\|^2\right]$$

$$= \frac{d+2}{2n(h^d c_d)h^2} \left[2\mathbf{y}_{k+1}^T \sum_{\mathbf{x}_i \in S_h(\mathbf{y}_k)} \mathbf{x}_i - n_k \|\mathbf{y}_{k+1}\|^2\right]$$

$$= \frac{d+2}{2n(h^d c_d)h^2} n_k \|\mathbf{y}_{k+1}\|^2. \quad (13)$$

The last item of the relation (13) is strictly positive except when $\mathbf{y}_k = \mathbf{y}_{k+1} = 0$.

Being bounded and strictly monotonic increasing, the sequence $\hat{f}_E$ is convergent. Note that if $\mathbf{y}_k = \mathbf{y}_{k+1}$ then $\mathbf{y}_k$ is the limit of $\hat{f}_E$, i.e., $\mathbf{y}_k$ is the fixed point of the mean shift procedure.

4 Processing in Spatial-Range Domain

An image is typically represented as a 2-dimensional lattice of r-dimensional vectors (pixels), where r is 1 in the gray level case, 3 for color images, or $r > 3$ in the multispectral case. The space of the lattice is known as the *spatial* domain while the gray level, color, or spectral information is represented in the *range* domain. However, after a proper normalization with σ_s and σ_r, global parameters in the spatial and range domains, the location and range vectors can be concatenated to obtain a *spatial-range* domain of dimension $d = r + 2$.

The main novelty of this paper is to apply the mean shift procedure for the data points in the joint spatial-range domain. Each data point becomes associated to a point of convergence which represents the local mode of the density in the d-dimensional space. The process, having the parameters σ_s and σ_r, takes into account simultaneously both the spatial and range information. A similar idea was exploited differently in [16]. In Section 5.4 we will compare the two approaches.

The output of the mean shift filter for an image pixel is defined as the range information carried by the point of convergence. This process achieves a high quality, discontinuity preserving spatial filtering. For the segmentation task, the convergence points sufficiently close in the joint domain are fused to obtain the homogeneous regions in the image.

The proposed spatial-range filtering and segmentation are described in the sequel with results shown for both gray level and color images. The perceptually uniform $L^*u^*v^*$ space has been used to represent the color information, while for the gray level cases only the L^* component has been considered.

5 Filtering

Let $\{\mathbf{x}_j\}_{j=1\ldots n}$ and $\{\mathbf{z}_j\}_{j=1\ldots n}$ be the d-dimensional original and filtered image points in the spatial-range domain. The upperscripts s and r will denote the spatial and range parts of the vectors, respectively. The original data is assumed to be normalized with σ_s for the spatial part and σ_r for the range.

Mean Shift Filtering

For each $j = 1 \ldots n$

1. Initialize $k = 1$ and $\mathbf{y}_k = \mathbf{x}_j$.
2. Compute $\mathbf{y}_{k+1} = \frac{1}{n_k} \sum_{\mathbf{x}_i \in S_1(\mathbf{y}_k)} \mathbf{x}_i$, $k \leftarrow k+1$ till convergence.
3. Assign $\mathbf{z}_j = (\mathbf{x}_j^s, \mathbf{y}_{conv}^r)$.

The last assignment specifies that the filtered data at the spatial location of $\mathbf{x}_j$ will have the range components of the point of convergence $\mathbf{y}_{conv}$. The number of points in the window $S_1(\mathbf{y}_k)$ of radius 1 and centered on Y_k is n_k. The unit radius of the window is due to the normalization.

5.1 Arithmetic Complexity

In a practical implementation the lattice structure of the spatial domain is used for the efficient search of the points $\mathbf{x}_i \in S_1(\mathbf{y}_k)$. This search can obviously be limited to a rectangular window of size 2×2 in the normalized space, which corresponds to $(2\lfloor\sigma_s\rfloor + 1)^2$ image pixels, where $\lfloor\cdot\rfloor$ is the down-rounded integer.

By denoting with k_c the mean number of iterations needed for convergence, the arithmetic complexity of the mean shift filtering is about $k_c(2\lfloor\sigma_s\rfloor + 1)^2$ flops per image pixel.

5.2 Normalization Constants

The value of σ_s is related to the spatial resolution of the analysis while the value of σ_r defines the range (color) resolution.

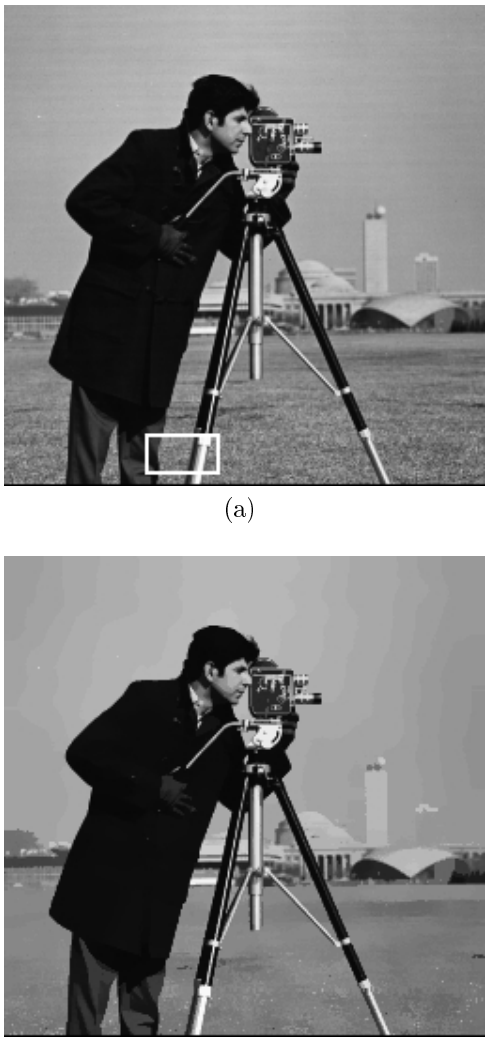
An asymptotically optimal (in the MISE sense) gradient estimate is obtained when the distribution in the joint space is normal. The radius of the searching window is a function of the number of data points n [12, p.152]. In our case, however, the data is far from being normal. Therefore, no theoretical constraints can be imposed on the values of σ_s and σ_r, which are task dependent and in practical settings their choice should incorporate a top-down, knowledge driven component.

A challenging issue not considered in this paper is the adaptive definition of the normalization constants. To take into account the nonstationarity of the input adaptive kernel estimation techniques were proposed in the statistical literature [15], however for less complex data. Beside exploiting a priori information (often available for low level vision) robust image understanding methods can also be helpful.

5.3 Experiments

Mean shift filtering with $(\sigma_s, \sigma_r) = (8, 4)$ has been applied to the often used 256×256 gray level *cameraman* image (Figure 3a), the processed image being shown in Figure 3b. The regions containing the grass field have been almost completely smoothed while details such as the tripod and the buildings in the background were preserved.

The entire processing time was a few seconds on a standard laptop with a 233 MHz *Pentium II* processor. We used a Java implementation of the algorithm. The mean number of iterations necessary for convergence was very low, around 3, due to the relatively small number of data points falling in the searching window.



(a)

(b)

Figure 3: *Cameraman* image. (a) Original. (b) Mean shift filtered $(\sigma_s, \sigma_r) = (8, 4)$.

To illustrate the effectiveness of the filtering process, the region marked in Figure 3a is represented in three dimensions in Figure 4a. In Figure 4b the mean shift paths associated with each pixel from the central plateau and the line are shown. Note that the convergence points (black dots) are situated in the opposite direction relative to the edge, while the shifts on the line remain on it. As a result, the filtered data (Figure 4c) shows clean quasi-homogeneous regions.

A second filtering example is given in Figure 5b. The original, 512×512 color image *baboon* has been processed with a mean shift filter having $(\sigma_s, \sigma_r) = (16, 16)$. While the texture of the fur has been cleaned,

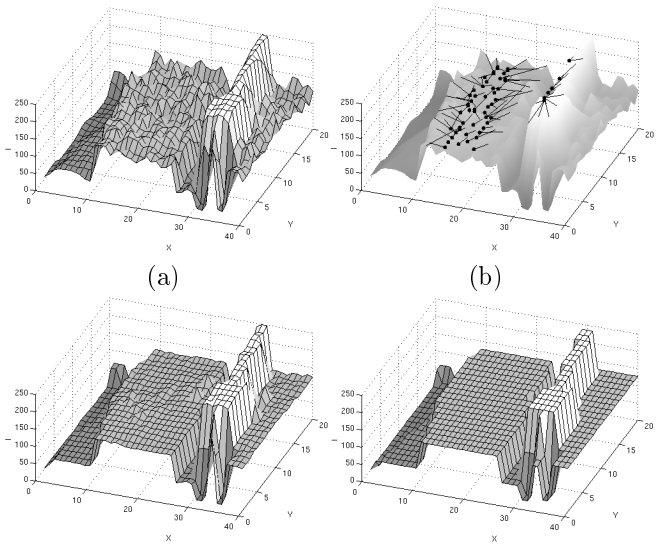


Figure 4: A 40×20 window from the image *cameraman*. (a) Original data (rotated and flipped over for better visualization). (b) Mean shift paths for the points in the central and top (white) plateaus. (c) Filtering result $(\sigma_s, \sigma_r) = (8, 4)$. (d) Segmentation result (see Section 6 for details).

the details of the eyes and the whiskers remained crisp.

5.4 Comparison to Bilateral Filtering

We note here two important differences between the mean shift and bilateral filtering proposed by Tomasi and Manduchi [16]. Both methods are based on the same principle, the simultaneous processing of both the spatial and range domains. However, while the bilateral filtering uses a static window in the two domains, the mean shift window is *dynamic*, moving in the direction of the maximum increase in the density gradient. Therefore, the mean shift filtering has a more powerful adaptation to the local structure of the data.

In addition, the filtering iterations proposed in [16] do not have a stopping criterion. After a sufficient number of iterations, the processed image collapses to a flat surface. The same observation is valid for other adaptive smoothing techniques [10, 11]. The process defined by mean shift is run till convergence and maintains the structure of the data.

6 Segmentation

The mean shift segmentation in the spatial-range domain has the same simple design as the filtering process. Again, we assume the input data to be normalized with (σ_s, σ_r). Let $\{\mathbf{x}_j\}_{j=1\ldots n}$ be the original image points, $\{\mathbf{z}_j\}_{j=1\ldots n}$ the points of convergence, and

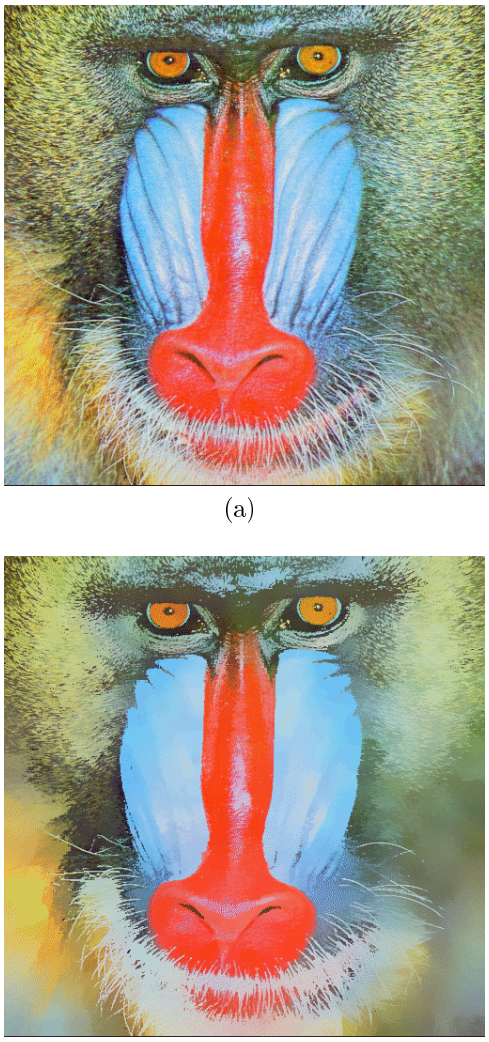


Figure 5: *Baboon* image. (a) Original. (b) Mean shift filtered $(\sigma_s, \sigma_r) = (16, 16)$.

$\{L_j\}_{j=1\ldots n}$ a set of labels (scalars).

Mean Shift Segmentation

1. For each $j = 1 \ldots n$ run the mean shift procedure for $\mathbf{x}_j$ and store the convergence point in $\mathbf{z}_j$.
2. Identify clusters $\{\mathbf{C}_p\}_{p=1\ldots m}$ of convergence points by linking together all $\mathbf{z}_j$ which are closer than 0.5 from each other in the joint domain.
3. For each $j = 1 \ldots n$ assign $L_j = \{p \mid \mathbf{z}_j \in \mathbf{C}_p\}$.
4. Optional: Eliminate spatial regions smaller than M pixels.

The first step of the segmentation is a filtering process. However all the information about the d-dimensional convergence point is stored now in $\mathbf{z}_j$, not

only its range part. Note also that the number of clusters m is controlled by the parameters (σ_s, σ_r).

The arithmetic complexity of the segmentation is similar to that of the mean shift filtering, its first step being the most computationally expensive.

6.1 Experiments

We employed the algorithm described above with $(\sigma_s, \sigma_r, M) = (8, 7, 20)$ to segment the 256×256 gray level image *MIT* (Figure 6a). The segmentation is presented in Figure 6b with the associated contours in Figure 6c. A number of 225 homogeneous regions were identified. The high quality contours allow the delineation of the walls, sky, steps, inscription on the building, etc.

Compare the segmentation in Figure 6 with the segmentations of the same image through clustering [4, Figure 4] or using Gibbs random field [9, Figure 7].

Returning to the *cameraman* image, Figure 7 shows the reconstructed image after the regions corresponding to the sky and grass were replaced with white. Observe the preservation of the details. The mean shift segmentation has been applied with $(\sigma_s, \sigma_r, M) = (8, 4, 10)$. Figure 4d shows the segmentation (with the same parameters) of the selected rectangular window in Figure 3a.

The segmentation with $(\sigma_s, \sigma_r, M) = (16, 7, 40)$ of the 512×512 color image *lake* is shown in Figure 8b. Compare this result with that of the multiscale approach in [14, Figure 11]. Finally, one can compare the contours of the color image *hand* presented in Figure 9 with those from [17, Figure 15] obtained through a complex global optimization.

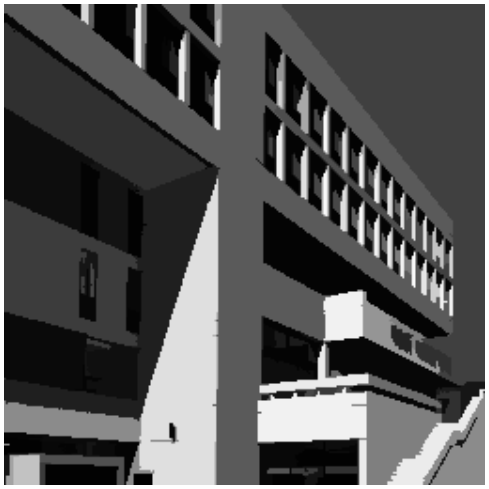
6.2 Discussion

It is interesting to contrast the mean shift segmentation with those based on the attraction force field [14] and edge flow propagation [8]. While all the three methods employ a vector field to detect regions in the spatial domain, only the mean shift based segmentation has strong statistical foundations. Our method associates the current pixel with a mode of the density located in its neighborhood (measured in both spatial and range domains).

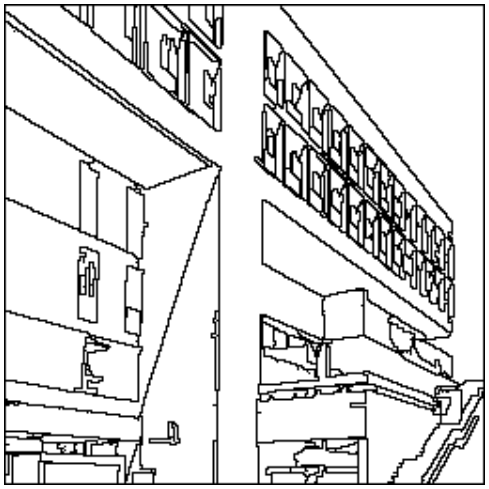
The attraction force field defined in [14] is computed at each pixel as a vector sum of pairwise affinities between the current pixel and all other pixels. No theoretical evidence of the existence of such a force field is given.

The edge flow in [8] is obtained at each location for a given set of directions as the magnitude of the gradient of a smoothed image. The quantization of the edge flow direction, however, may introduce artifacts.

(a)



(b)



(c)

Figure 6: *MIT* image. (a) Original. (b) Segmented $(\sigma_s, \sigma_r, M) = (8, 7, 20)$. (c) Contours.

Figure 7: Segmentation with $(\sigma_s, \sigma_r, M) = (8, 4, 10)$ and reconstruction of the *cameraman* image after the elimination of regions representing sky and grass.

Recall that, by contrast, the direction of the mean shift is dictated solely by the data.

7 Conclusions

This paper suggests that effective image analysis can be implemented based on the mean shift procedure. The nonparametric estimation of the density gradient in the spatial-range domain is a useful tool for bottom-up computer vision tasks such as edge preserving filtering and segmentation. The methods we proposed can be easily extended to the processing of other low level image features like the texture or optical flow.

Acknowledgment

This research was supported by the NSF under the grant IRI-9530546.

References

[1] G.R. Bradski, "Real Time Face and Object Tracking as a Component of a Perceptual User Interface", Proc. *IEEE Workshop on Applications of Comp. Vis.*, Princeton, 214-219, 1998.

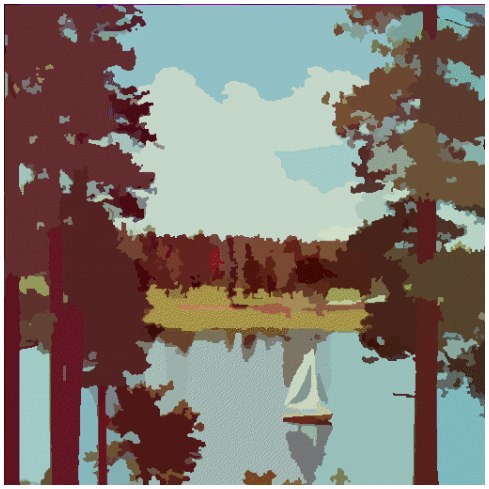
[2] Y. Cheng, "Mean Shift, Mode Seeking, and Clustering", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, 790-799, 1995.

[3] K. Cho, P. Meer, "Image Segmentation from Consensus Information", *Comp. Vision and Image Understanding*, vol. 68, 72-89, 1997.

[4] D. Comaniciu, P. Meer, "Robust Analysis of Feature Spaces: Color Image Segmentation", Proc. *IEEE Conf. on Comp. Vis. and Pattern Recognition*, Puerto Rico, 750-755, 1997.

[5] D. Comaniciu, P. Meer, "Distribution Free Decomposition of Multivariate Data", To appear. *Pattern Anal. and Application*, 1999.

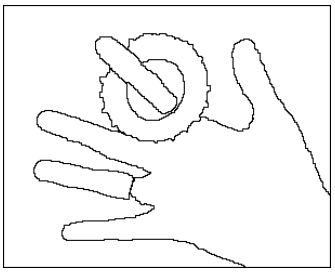
(a)



(b)

Figure 8: *Lake* image. (a) Original. (b) Segmented $(\sigma_s, \sigma_r, M) = (16, 7, 40)$.

(a)



(b)

Figure 9: *Hand* image. (a) Original. (b) Contours $(\sigma_s, \sigma_r, M) = (16, 19, 40)$.

[6] P.F. Felzenszwalb, D.P. Huttenlocher, "Image Segmentation Using Local Variation", Proc. *IEEE Conf. on Comp. Vis. and Pattern Recognition*, Santa Barbara, 98-103, 1998.

[7] K. Fukunaga, L.D. Hostetler, "The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition", *IEEE Trans. Info. Theory*, vol. IT-21, 32-40, 1975.

[8] W.Y. Ma, B.S. Manjunath, "Edge Flow: A Framework of Boundary Detection and Image Segmentation", Proc. *IEEE Conf. on Comp. Vis. and Pattern Recognition*, Puerto Rico, 744-749, 1997.

[9] T.N. Pappas, "An Adaptive Clustering Algorithm for Image Segmentation", *IEEE Trans. Signal Process.*, vol. 40, 901-914, 1992.

[10] P. Perona, J. Malik, "Scale-Space and Edge Detection Using Anisotropic Diffusion", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, 629-639, 1990.

[11] P. Saint-Marc, J.S. Chen, G.G. Medioni, "Adaptive Smoothing: A General Tool for Early Vision", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, No. 6, 514-529, 1991.

[12] D.W. Scott, *Multivariate Density Estimation*, New York: Wiley, 1992.

[13] B.W. Silverman, *Density Estimation for Statistics and Data Analysis*, New York: Chapman and Hall, 1986.

[14] M. Tabb, N. Ahuja, "Multiscale Image Segmentation by Integrated Edge and Region Detection", *IEEE Trans. Image Process.*, vol. 6, 642-655, 1997.

[15] G.R. Terrell, D.W. Scott, "Variable Density Estimation", *The Annals of Statistics*, vol. 20, 1236-1265, 1992.

[16] C. Tomasi, R. Manduchi, "Bilateral Filtering for Gray and Color Images", Proc. *Int'l Conf. on Comp. Vis.*, Bombay, India, 839-846, 1998.

[17] S.C. Zhu, A. Yuille, "Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, 884–900, 1996.