

Prior-Constrained Scale-Space Mean Shift

Kazunori Okada¹ Maneesh Singh² Visvanathan Ramesh²

¹ Dept. of Computer Science, San Francisco State University

² Real-Time Vision & Modeling Dept., Siemens Corp. Research, Inc.
kazokada@sfsu.edu {msingh,rameshv}@scr.siemens.com

Abstract

This paper proposes a new variational bound optimization framework for incorporating spatial prior information to the mean shift-based data-driven mode analysis, offering flexible control of the mean shift convergence. Two forms of Gaussian spatial priors are considered. Attractive prior pulls the convergence toward a desired location. Repulsive prior pushes away from such a location. Using a generic variational optimization formulation via construction of quadratic lower and upper bounds, we show that the prior-constrained mean shift step can be interpreted as an information fusion of the data and prior terms in the sense of the best linear unbiased estimator. This approach is used to propose a mode parsing algorithm using the inhibition-of-return principle. The proposed algorithm is used for a semi-automatic 3D segmentation of lung nodules in CT data for evaluating its effectiveness. Our experiments demonstrate that the proposed solution can successfully segment challenging wall-attached cases.

1 Introduction

Mean shift is a popular data-driven technique for analyzing the mode structure of kernel-smoothed function surfaces. It is an adaptive gradient-ascent algorithm with automatic step-size selection and is convergent to a mode of the kernel-smoothed function. This framework provides an efficient solution to the general data clustering problem in the Parzen windows setting [7, 1, 4]. The mode-seeking property of the mean shift algorithm has also been successfully applied to a wide range of vision problems such as tracking [5, 2] and segmentation [3, 4, 13]. From an optimization-theory perspective, mean shift is a variational bound optimization algorithm [18, 4, 16, 6]. For a class of kernels with convex profiles, the convexity property yields quadratic lower-bounds of the smoothed function. These quadratic lower-bounds can then be readily optimized to define the next location, or *mean shift*, and its iteration is guaranteed to converge to a local maximum from any given point in the data space without adjusting the step sizes.

In this paper, we propose a variational optimization-based extended mean shift framework that incorporates spatial probabilistic priors in order to control where the mean shift converges. This paper focuses on Gaussian scale space [20, 9, 17] used as a positive kernel-smoothed function and on Gaussian model used as a spatial prior. However our contributions are generic beyond these choices of the scale-space and the prior. We derive such a convergent modified mean shift step by i) constraining the kernel-smoothed (density) function with the spatial priors, ii) deriving quadratic lower-bounds of the new constrained function, and iii) deriving the modified mean shift as the optimizer of the bounds. We demonstrate that the resulting formula takes a form of an information fusion of the data-driven and prior terms in the Best Linear Unbiased Estimator (BLUE)

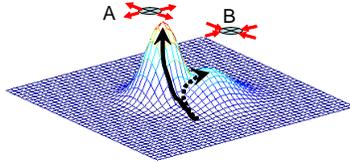


Figure 1: Conceptual illustration of the prior-constrained mean shift with a synthetic bimodal case. A: repulsive (negative) prior. B: attractive (positive) prior. Solid: data-driven convergence. Dashed: prior-constrained convergence.

sense [12, 21]. This framework offers a principled way to exploit independent information sources other than given data, such as i) user-interaction through a graphical user interface and ii) prior knowledge of domain-experts.

A key advantage of the proposed formulation is to detect hard-to-discover weak modes in multimodal data. By construction, mean shift assumes kernel-smoothing of objective functions. Such smoothing can mask smaller modes located close to larger ones. However, the small or weak modes can represent something important depending on specific application context. In Figure 1, we illustrate such a situation using a bimodal function with a pair of weak and strong modes/peaks located nearby. Although initialization is placed closer to the desirable mode, the data-driven convergence can fall to the farther but stronger mode, shown by the solid arrow. We propose two different types of priors: i) *repulsive prior*, depicted as an ellipse at point A, which pushes the convergence away from the data-driven convergence and ii) *attractive prior*, depicted at point B, which pulls the convergence toward the small target mode. The dashed arrow schematizes the convergence path to the desired weak mode by the prior-constrained mean shift.

In some application scenarios, furthermore, it is of great interest to automatically determine the spatial priors from the data itself for recursively parsing the multimodal data. To address this task, we present an *Inhibition Of Return (IOR)* algorithm which sequentially seeks data modes using the above prior-constrained mean shift together with convergence diagnostics. At each convergence, we fit a Gaussian to the local mode for constructing a repulsive prior using techniques proposed in [13]. This enables the IOR algorithm to sequentially visit all the data modes near the initialization point. This algorithm is conceptually similar to the well-known visual attention model in [8].

The effectiveness of the proposed algorithms are evaluated in a real world problem of semi-automatic 3D segmentation of lung tumor nodules. As pointed out by Okada et al. [14], the segmentation of the juxtapleural (wall-attached) tumor poses difficult challenges to the state-of-the-art [11, 13] because the target convex-shaped tumor, attached to large wall-like or tubular structures, appear as a weak mode surrounded by stronger non-target modes. They proposed a resampling-based approach for realizing the repulsive prior similar to our approach. We offer here a comprehensive theoretical framework to ground these ideas in more general probabilistic settings. Using the similar experimental setting in [14], we show that we can successfully segment such difficult wall-attached cases.

2 Scale-Space Mean Shift

Mean shift framework [7, 1, 4] was first introduced for iteratively seeking modes of kernel-based non-parametric density function, such as kernel density estimate of the form $\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n K_h(\mathbf{x} - \mathbf{x}_i)$, given a sample set $\mathcal{S}_n = \{\mathbf{x}_i | i = 1 \cdots n\}$ and the kernel bandwidth h . Let $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle_H = \mathbf{x}_1^T \mathbf{H}^{-1} \mathbf{x}_2$ denote the inner product of two vectors where \mathbf{H}

is a symmetric positive definite matrix. The associated norm is then denoted by $\|\cdot\|_H$ or by $\|\cdot\|_h$ when a diagonal bandwidth matrix $\mathbf{H} = h\mathbf{I}$ is considered. Then $K_H(\cdot)$ denotes a density kernel defined with its 1-dimensional profile $k(\cdot)$ via $K_H(\mathbf{x}) = |\mathbf{H}|^{-\frac{1}{2}}k(\|\mathbf{x}\|_H^2)$. Also $K_h(\cdot)$ denotes a radially-symmetric kernel with a diagonal bandwidth matrix. For the Epanechnikov kernel, the well-known mean shift algorithm is resulted [7]. For other kernels (e.g., the Gaussian kernel), the following mean shift step with a weighted arithmetic mean is defined $\mathbf{m}(\mathbf{x}, h) = \frac{\sum_{i=1}^n \mathbf{x}_i g(\|\mathbf{x} - \mathbf{x}_i\|_h^2)}{\sum_{i=1}^n g(\|\mathbf{x} - \mathbf{x}_i\|_h^2)} - \mathbf{x}$, where $g(\mathbf{x}) = k'(\mathbf{x})$.

Beyond the above original formulation, the mean shift framework can be derived for a general class of functions constructed using kernel-smoothing [19]. Since the Gaussian scale-space [17] can also be interpreted as a family of kernel-smoothed functions, it is possible to derive a mean shift algorithm in this domain [13, 2]. This provides a powerful robust tool for vision problems such as visual tracking [2] and medical image segmentation [13].

Let $I(\mathbf{x}) \in \mathbb{R}^+$ denote the positive-valued image data to be analyzed where $\mathbf{x} \in \mathbb{R}^d$ is the d -D coordinate indicating data (pixel/voxel) location. Gaussian scale-space $L(\mathbf{x}; h)$ of $I(\mathbf{x})$ is a solution to the diffusion equation $\partial_h L = 1/2\nabla^2 L$ with an initial condition $L(\mathbf{x}; 0) = I(\mathbf{x})$. It takes the form of convolution of the image data with a Gaussian kernel K_h having bandwidth $\mathbf{H} = h\mathbf{I}$,

$$L(\mathbf{x}; h) = I(\mathbf{x}) * K_h(\mathbf{x}) = \int I(\mathbf{x}') K_h(\mathbf{x} - \mathbf{x}') d\mathbf{x}' \quad (1)$$

where $K_h = |2\pi h\mathbf{I}|^{-\frac{1}{2}} \exp(-\frac{\|\mathbf{x}\|_h^2}{2})$. *Scale-space mean shift* is then derived by maximizing the quadratic lower-bound of $L(\mathbf{x}; h)$ constructed using the convexity of K_h 's profile. This results in the form proportional to the spatial gradient of the scale-space $\nabla L(\mathbf{x}; h)$,

$$\mathbf{m}(\mathbf{x}, h) = \frac{\int \mathbf{x}' K_h(\mathbf{x} - \mathbf{x}') I(\mathbf{x}') d\mathbf{x}'}{\int K_h(\mathbf{x} - \mathbf{x}') I(\mathbf{x}') d\mathbf{x}'} - \mathbf{x} = h \frac{\nabla L(\mathbf{x}; h)}{L(\mathbf{x}; h)} \quad (2)$$

An iterative procedure over the discretized data space \mathbf{x}_i is then defined analogous to the density mean shift: $\mathbf{y}_{k+1} = \mathbf{m}(\mathbf{y}_k, h) + \mathbf{y}_k = \frac{\sum_i \mathbf{x}_i K_h(\mathbf{y}_k - \mathbf{x}_i) I(\mathbf{x}_i)}{\sum_i K_h(\mathbf{y}_k - \mathbf{x}_i) I(\mathbf{x}_i)}$ with $\mathbf{y}_1 = \mathbf{x}_0$. The scale-space mean shift (2) can be understood as a mean shift with positive-valued weights of intensity because $\forall i, I(\mathbf{x}_i) > 0$ by construction. Its convergence to the local maximum of the scale-space L is also guaranteed following the result in [4].

3 Mean Shift Constrained by Gaussian Priors

We can interpret the original density mean shift in the following Maximum Likelihood (ML) sense. Given data sample \mathbf{x}_i from a random variable \mathbf{X} , a kernel density estimate is used to represent the likelihood distribution of \mathbf{X} . Mean shift algorithm can then be understood as an ML estimator if the distribution is unimodal. In case of a multimodal distribution of \mathbf{X} , each locally-maximum likelihood location can be parsed and *the* ML estimate can be found. Mode parsing will be discussed later in Section 4.

Suppose now that we have another source of information for \mathbf{X} . Under the condition of independence, the likelihood from this source can then be multiplied to constrain the likelihood from the kernel density estimator. The following derives the constrained mean shift formulae for Gaussian scale-space by treating $L(\mathbf{x}; h)$ as a pseudo-likelihood. In this setting, peaks in a scale-space image corresponds to the density modes. The results in the scale-space can be readily extended to the general true likelihoods in the ML sense.

3.1 Attractive Prior

Let such a prior for \mathbf{X} be given by a Gaussian distribution,

$$Q(\mathbf{x}) = |2\pi\mathbf{S}|^{-\frac{1}{2}} \exp\left(-\frac{\|\mathbf{x} - \mathbf{a}\|_{\mathbf{S}}^2}{2}\right) \quad (3)$$

with mean \mathbf{a} and bandwidth \mathbf{S} . We exploit this prior to *bias* or *attract* the solution toward a pre-selected point \mathbf{a} in the data space. Hence we call it an *attractive prior*.

By treating the Gaussian scale-space with $I(\mathbf{x}_i) > 0$ as a pseudo-likelihood, we multiplicatively modify the scale-space with the prior in (3), resulting in the constrained scale-space function $\tilde{L}_1(\mathbf{x}; h)$,

$$\tilde{L}_1(\mathbf{x}; h) = L(\mathbf{x}; h)Q(\mathbf{x}) \quad (4)$$

where $L(\mathbf{x}; h) = \sum_{i=1}^N I(\mathbf{x}_i)K_h(\mathbf{x} - \mathbf{x}_i)$ is the original scale-space defined over discretized data domain and N is the number of data points (pixels/voxels). When there are multiple independent priors, the likelihood function can be written as $\tilde{L}_1(\mathbf{x}; h) = L(\mathbf{x}; h) \prod_m Q_m(\mathbf{x})$.

Now let $\mathbf{z}_i(\mathbf{x}) = \frac{1}{2} [\|\mathbf{x} - \mathbf{x}_i\|_h^2 + \|\mathbf{x} - \mathbf{a}\|_{\mathbf{S}}^2]$. Then we can rewrite the constrained scale-space as $\tilde{L}_1(\mathbf{x}; h) = A \sum_i I(\mathbf{x}_i) \exp(-\mathbf{z}_i(\mathbf{x}))$ where $A = |2\pi(h\mathbf{I})|^{-1/2} \times |2\pi\mathbf{S}|^{-1/2}$. Using the convexity of the exponential function yields a lower-bound L_b

$$\tilde{L}_1(\mathbf{x}; h) - \tilde{L}_1(\mathbf{x}_0; h) \geq L_b(\mathbf{x}) = A \sum_i I(\mathbf{x}_i) \exp(-\mathbf{z}_i(\mathbf{x}_0)) [\mathbf{z}_i(\mathbf{x}_0) - \mathbf{z}_i(\mathbf{x})] \quad (5)$$

It is trivial to see that $L_b(\mathbf{x})$ has a unique maximum since it is quadratic in \mathbf{x} and $\nabla^2 L_b(\mathbf{x}) = -A \sum_i I(\mathbf{x}_i) \exp(-\mathbf{z}_i(\mathbf{x}_0)) [\mathbf{H}^{-1} + \mathbf{S}^{-1}] < 0$ ¹. Thus, solving $\nabla L_b(\mathbf{x}) = \mathbf{0}$ gives the desired constrained mean shift step denoted by $\mathbf{m}_a(\mathbf{x}, H, Q)$

$$\begin{aligned} \mathbf{m}_a(\mathbf{x}_0, h, Q) + \mathbf{x}_0 &= (\mathbf{H}^{-1} + \mathbf{S}^{-1})^{-1} \left[\mathbf{H}^{-1} \frac{\sum_i \mathbf{x}_i \exp(-\frac{\|\mathbf{x}_0 - \mathbf{x}_i\|_h^2}{2}) I(\mathbf{x}_i)}{\sum_i \exp(-\frac{\|\mathbf{x}_0 - \mathbf{x}_i\|_h^2}{2}) I(\mathbf{x}_i)} + \mathbf{S}^{-1} \mathbf{a} \right] \\ &= (\mathbf{H}^{-1} + \mathbf{S}^{-1})^{-1} [\mathbf{H}^{-1}(\mathbf{m}(\mathbf{x}_0, h) + \mathbf{x}_0) + \mathbf{S}^{-1} \mathbf{a}] \end{aligned} \quad (6)$$

The convergence of the associated mean shift procedure is guaranteed by construction [4]. Note that derived iterative step is equivalent to a BLUE fusion [12, 21] of the Gaussian prior $\mathcal{N}(\mathbf{a}, \mathbf{S})$ and the data-driven scale-space mean shift $\mathcal{N}(\mathbf{x}_{ms}, \mathbf{H})$ where $\mathbf{x}_{ms} = \mathbf{m}(\mathbf{x}_0, h) + \mathbf{x}_0$. The confidence level we associate with the prior can be controlled by varying \mathbf{S} . As we decrease $\|\mathbf{S}\|$, the confidence in the prior increases. With the highest confidence, i.e. $\mathbf{S} = \mathbf{0}$, the solution reduces to \mathbf{a} : the most likely spatial point indicated by the prior. On the other hand, as $\|\mathbf{S}\|$ increases, the confidence in the prior decreases. With $\|\mathbf{S}\| \rightarrow \infty$, the mean shift step converges to the original scale-space mean shift step \mathbf{x}_{ms} .

3.2 Repulsive Prior

In some situations, the available prior information can take a form of a negation. For example, it might be *unlikely* for a certain parameter to take some specific values. Such a prior can be used to *repel* or *push away* the convergence from a point \mathbf{a} . Thus we call it a *repulsive prior*. We represent such a prior by an inverted Gaussian

$$\tilde{Q}(\mathbf{x}) = |2\pi\mathbf{S}|^{-\frac{1}{2}} - Q(\mathbf{x}) \quad (7)$$

¹ < 0 is used as a shorthand for negative definiteness of the Hessian.

where $Q(\cdot)$ is the positive Gaussian prior defined in (3). We treat $\tilde{Q}(\mathbf{x})$ as a likelihood function by appropriately truncating and imposing a normalization over a finite domain. Then we can use it to multiplicatively constrain the scale-space as in Section 3.1, resulting in the following constrained scale-space

$$\tilde{L}_3(\mathbf{x}; h) = L(\mathbf{x}; h)\tilde{Q}(\mathbf{x}) = |2\pi\mathbf{S}|^{-\frac{1}{2}}L(\mathbf{x}; h) - L(\mathbf{x}; h)Q(\mathbf{x}) \quad (8)$$

Variational bound optimization of $\tilde{L}_3(\mathbf{x}; h)$ is not straightforward. Our strategy is to derive a quadratic lower-bound $L_c(\mathbf{x})$ of $\tilde{L}_3(\mathbf{x}; h)$ and derive the modified mean shift by solving $\nabla L_c(\mathbf{x}) = \mathbf{0}$. This construction assures the convergence of the resulting procedure. Now the convexity property of the exponential kernel, used for the positive case, similarly provides a lower-bound of the first positive term of \tilde{L}_3 . However, it cannot lower-bound the second negative term because of the negation. In fact a quadratic *upper-bound* of the exponential is required for deriving the desired variational lower-bound L_c .

We derive an analytic expression of such a quadratic upper-bound using Taylor series expansion of the kernel function. First we consider the exact 1st-order Taylor series expansion about \mathbf{x}_0 of a differentiable function $f(\mathbf{x})$ as a sum of the 1st-order Taylor polynomial T_1 and Lagrange remainder R_1 so that $f(\mathbf{x}) = \sum_{|k|=0}^1 \frac{f^{(k)}(\mathbf{x}_0)}{k!}(\mathbf{x} - \mathbf{x}_0)^k + R_1$ where k is a multi-index. The remainder is defined by using the mean value theorem with $\zeta \in [\mathbf{x}_0, \mathbf{x}]$ so that $R_1 = \frac{f^{(2)}(\zeta)}{2!}(\mathbf{x} - \mathbf{x}_0)^2 \exists \zeta \in [\mathbf{x}_0, \mathbf{x}]$. The desired quadratic upper-bound of $f(\mathbf{x})$ is then given by maximizing the Lagrange remainder over the free variable ζ such that $f(\mathbf{x}) = T_1(\mathbf{x}) + R_1(\mathbf{x}, \zeta) \leq T_1(\mathbf{x}) + \max_{\zeta} R_1(\mathbf{x}, \zeta)$. For a Gaussian with mean $\mathbf{0}$ and width Σ , this procedure results in the following quadratic upper-bound

$$\begin{aligned} \exp\left(-\frac{\|\mathbf{x}\|_{\Sigma}^2}{2}\right) &\leq \exp\left(-\frac{\|\mathbf{x}_0\|_{\Sigma}^2}{2}\right) + \exp\left(-\frac{\|\mathbf{x}_0\|_{\Sigma}^2}{2}\right)[- \Sigma^{-1}\mathbf{x}_0]^T(\mathbf{x} - \mathbf{x}_0) \\ &\quad + \exp\left(-\frac{3}{2}\right)\|\mathbf{x} - \mathbf{x}_0\|_{\Sigma}^2 \end{aligned} \quad (9)$$

We apply this upper-bound (9) to the function of form $\exp(-\frac{1}{2}\mathbf{z})$ where $\mathbf{z} = \|\mathbf{x} - \mathbf{x}_i\|_h^2 + \|\mathbf{x} - \mathbf{a}\|_{\Sigma}^2$, resulting in a quadratic lower-bound to the negative term in (8). The complete lower-bound L_c is given by combining this result with the one for the positive term. After some algebra, we can show that solving $\nabla L_c(\mathbf{x}) = \mathbf{0}$ results in the following constrained mean shift step denoted by $\mathbf{m}_r(\mathbf{x}_0, h, Q)$

$$\begin{aligned} \mathbf{m}_r(\mathbf{x}_0, h, Q) + \mathbf{x}_0 &= [\mathbf{H}^{-1} + \alpha_0(\mathbf{H}^{-1} + \mathbf{S}^{-1})]^{-1} \\ &\quad [\mathbf{H}^{-1}[(1 - \tilde{Q}_0)\mathbf{x}_{ms} + \tilde{Q}_0\mathbf{a}] + \alpha_0(\mathbf{H}^{-1} + \mathbf{S}^{-1})[(1 + \frac{\tilde{Q}_0}{\alpha_0})\mathbf{x}_0 - \frac{\tilde{Q}_0}{\alpha_0}\mathbf{a}]] \end{aligned} \quad (10)$$

where $\alpha_0 = \frac{2\exp(-\frac{3}{2})\sum_i I(\mathbf{x}_i)}{\sum_i I(\mathbf{x}_i)\exp(-\frac{1}{2}\|\mathbf{x}_0 - \mathbf{x}_i\|_h^2)}$ and $\tilde{Q}_0 = |2\pi\mathbf{S}|^{\frac{1}{2}}Q(\mathbf{x}_0)$. The derived optimization step can be again interpreted as a BLUE fusion of two Gaussian sources with more complex mean and covariance terms than the attractive case: $\mathcal{N}((1 - \tilde{Q}_0)\mathbf{x}_{ms} + \tilde{Q}_0\mathbf{a}, \mathbf{H})$ and $\mathcal{N}((1 + \frac{\tilde{Q}_0}{\alpha_0})\mathbf{x}_0 - \frac{\tilde{Q}_0}{\alpha_0}\mathbf{a}, (\mathbf{H}^{-1} + \mathbf{S}^{-1})^{-1}/\alpha_0)$. Note that when the current point \mathbf{x}_0 is asymptotically far from the prior located at \mathbf{a} , the mean-shift step reduces to be the BLUE fusion of \mathbf{x}_{ms} and \mathbf{x}_0 since $\tilde{Q}_0 \rightarrow 0$.

3.3 Synthetic Example

Figure 2 shows 1D simulation of the prior-constrained mean shift algorithms with the attractive and repulsive priors. The 1D bimodal image data is $I(x) = \max[\exp(-(x +$

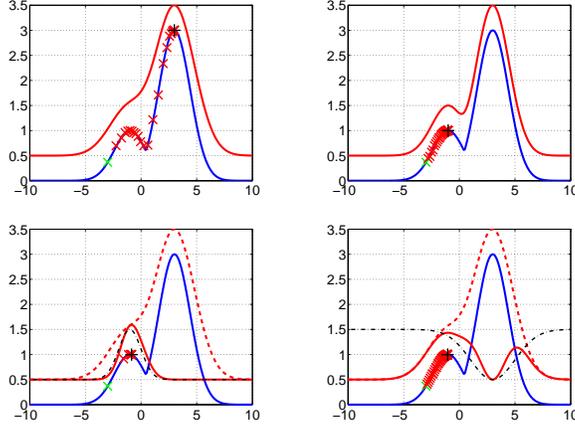


Figure 2: Prior-constrained mean shift with 1D bimodal data. The input data (solid blue) consists of two Gaussian components centered at -1 and 3 . The solid red curve with a 0.5 vertical bias show the function to be optimized for each case. "x": iterations initialized at $x_p = -3$. "+" : convergence. Top-left: scale-space mean shift (SSMS) in (2) with $h=1.2$. Top-right: SSMS with $h=0.2$. Bottom-left: variational attractor-constrained mean shift with $h=1.2$ and $(a, S) = (-1, 1)$, $L(x, h)$: dashed, $Q(x)$: dot-dashed, $\tilde{L}_1(x; h)$: solid red. Bottom-right: variational repeller-constrained mean shift with $h=1.2$ and $(a, S) = (3, 4)$, $L(x, h)$: dashed, $\tilde{Q}(x)$: dot-dashed, $\tilde{L}_3(x; h)$: solid red.

$1)^2/4, 3\exp(-(x-3)^2/4]$, illustrating a situation similar to Figure.1. As shown in the figure, the scale-space mean shift initialized at $x_p = -3$ converges to the stronger right peak (top-left), missing the weaker left peak that can be found with a smaller analysis bandwidth (top-right). Using the same bandwidth, the prior-constrained mean shift algorithms are controlled to converge to the weaker peak by setting the attractive (bottom-left) and repulsive (bottom-right) priors appropriately (shown in dot-dashed curves).

4 Local Data Parsing by Inhibition of Return

Automatic data and mode parsing algorithms are useful for representing and processing structures in multimodal data. However they suffer from the same inability to detect weak data modes as was discussed in the introduction. To solve the parsing problem with the presence of weak modes, we propose a novel *inhibition of return (IOR)* framework using the prior-constrained mean shift. The main idea is to sequentially perform the negative prior-constrained mean shift procedure by setting repulsive priors at each consecutive convergence point of the constrained mean shift. We use the robust multiscale Gaussian fitting solution [13] for setting the priors. At each convergence, we *implicitly* fit a 3D Gaussian intensity model $I(\mathbf{x}) \simeq [\alpha \mathcal{N}(\mathbf{x}; \mathbf{u}, \Sigma) + \beta]_{\mathbf{x} \in \Omega}$ to the data, by using scale-space mean shift (2), where $\alpha, \beta > 0$ and Ω denotes the basin of attraction. Then the center and the approximate 3D shape of a blob-like peak at the convergence are characterized by the mean \mathbf{u}^* and the covariance Σ^* of the fitted Gaussian. We set a negative prior (7) by inverting the fitted Gaussian with (\mathbf{u}^*, Σ^*) .

This procedure traverses from mode to mode starting from an arbitrary initial point, parsing all the blob-like data structures located nearby. This parsing process is efficient because each detected mode is inhibited from being revisited by setting a repulsive prior. This IOR concept was first proposed for a computational model of visual attention [8]. However, their neural network-based implementation differs greatly from our approach,

although we share their basic concept. The following describes the algorithm.

Initialization data $I(\mathbf{x}_i)$, initial point \mathbf{x}_p , and $Q_{m=0}=\text{NIL}$:

Loop over m until no structure is found near \mathbf{x}_p :

1. Perform Gaussian fitting with the prior-constrained mean shift $\mathbf{m}_r(\mathbf{x}, h)$ constrained by repellers $\{Q_{m'} | m' = 1, \dots, m\} \rightarrow (\mathbf{u}^*, \Sigma^*)$.
2. Set the m -th parsed structure $\text{Str}_m \leftarrow (\mathbf{u}^*, \Sigma^*)$.
3. Set the $m+1$ -th prior $Q_{m+1} \leftarrow (\mathbf{a}, \mathbf{S}) = (\mathbf{u}^*, \Sigma^*)$.

This IOR-based algorithm results in a set of blob structures $\{\text{Str}_m\}$ located near \mathbf{x}_p .

5 Experiments

To evaluate the effectiveness of the proposed framework, we apply a 3D implementation of the IOR-based local data analysis to the semi-automatic segmentation of lung nodules.

5.1 Lung Nodule Segmentation

In the past decades, a large amount of work has been done in computer-aided detection and segmentation of lung nodules toward improving mortality rate of lung cancers [15]. Recently, more sophisticated solutions [11, 10, 13] for this purpose have been proposed by taking advantage of increased image quality and resolution with the enhanced CT scanners. Despite the improvements, however, the problem remains open due to the data’s natural complexity. For example, many lung nodules are attached or located nearby other non-target structures such as lung wall and vessels [10, 14]. Because the intensity values of these structures can be similar to those of the nodules, separating them from such structures is not trivial. The robust solution by Okada et al. [13] approached this issue by incorporating geometric constraints, however it still fails to segment many nodules located near or attached to neighboring structures. These failures are caused because the target nodules are located near rib bones which have much higher intensity values than nodules, thus falsely attracting the scale-space mean shift convergence. This is precisely the situation, as illustrated in Figure 1, in which the prior-constrained mean shift framework is designed to be effective. Therefore we will extend the Okada’s approach by replacing the scale-space mean shift with its prior-constrained counter part, improving the performance for the attached cases. Figure 4 compares the results of the original and extended solutions.

An IOR-based solution is proposed for this purpose as described in Section 4. After the first iteration, we perform a goodness of fit test based on chi-square measure [13] after the step 2. If the initial fit is verified as a good fit, no further process is performed. When the initial fit fails according to the measure, however, a repulsive prior is set as in Section 3.2 and the next iteration of the IOR algorithm is carried out. In this study, we use two iterations.

5.2 Toy Data

For testing the feasibility, we first evaluated a 2D version of the system on synthetic data as shown in Figure 3. Emulating the wall-attached nodule case, the data is constructed with a step function (lung wall), a large Gaussian with higher intensity (bone), and a small truncated Gaussian with lower intensity (wall-attached nodule). The results demonstrate that our solution correctly finds and segments the difficult target structure in spite of the presence of a stronger mode, white noise, and variation in the initialization.

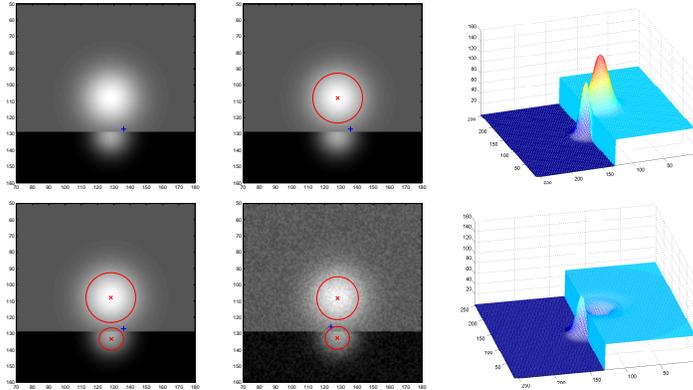


Figure 3: Two-step IOR-based data analysis evaluated with 2D synthetic data. From top-left to bottom-right: A) input data, B) initial Gaussian fit from the initialization at “+”, C) initial weights, D) second step with the prior-constrained MS capturing the tumor correctly, E) second step with the same data with white noise and different initialization, F) weights constrained by the repulsive prior.

5.3 Chest HRCT Data

We use a clinical data set of the thin-section (1.25 mm slice thickness) chest high-resolution computed tomography (HRCT) images including 39 patients with a total of 1310 nodules. The data is recorded with Multislice CT scanners and anonymized. Each volumetric image consists of 12-bit positive values over an array of 512x512 lattices. The implementation of the Gaussian fitting algorithm follows the settings described in [13]. This resulted in 106 verified failure cases. Most of these failures were the wall-attached cases and they were correctly segmented by our data parsing solution with the prior-constrained mean shifts. Figure 4 shows 8 illustrative examples for such cases. The left three columns show the original failed segmentation results in three orthogonal cross-sections respectively while the right three columns show the segmentation results with the prior constraints. Note that the original and our results are shown in different cross sections of the same volumes thus they appear differently. See the caption for details.

6 Conclusions

This paper proposes a novel framework for incorporating attractive and repulsive spatial priors in the Gaussian form to the mean shift-based data-driven mode analysis. Using the variational bound optimization scheme, we derived two convergent optimizers for the constrained Gaussian scale-space functions, resulting in a BLUE fusion of two independent Gaussian sources. Further, we proposed a mode parsing algorithm based on the IOR approach using the prior-constrained mean shift. This algorithm is then successfully applied to the semi-automatic 3D lung tumor segmentation problem to segment difficult wall-attached cases.

We consider the presented work as a step toward deriving a general framework for combining data-driven statistical analysis with prior information in order to provide effective and efficient vision solutions. We believe that our contribution is general and the proposed formulation can be extended in the following ways: i) exploiting the prior-constrained mean shift in the kernel density estimate domain, ii) applying to problems other than segmentation such as tracking, iii) adapting the IOR-based mode parsing algorithm to more flexible visual search [8] by updating the initialization at each convergence.

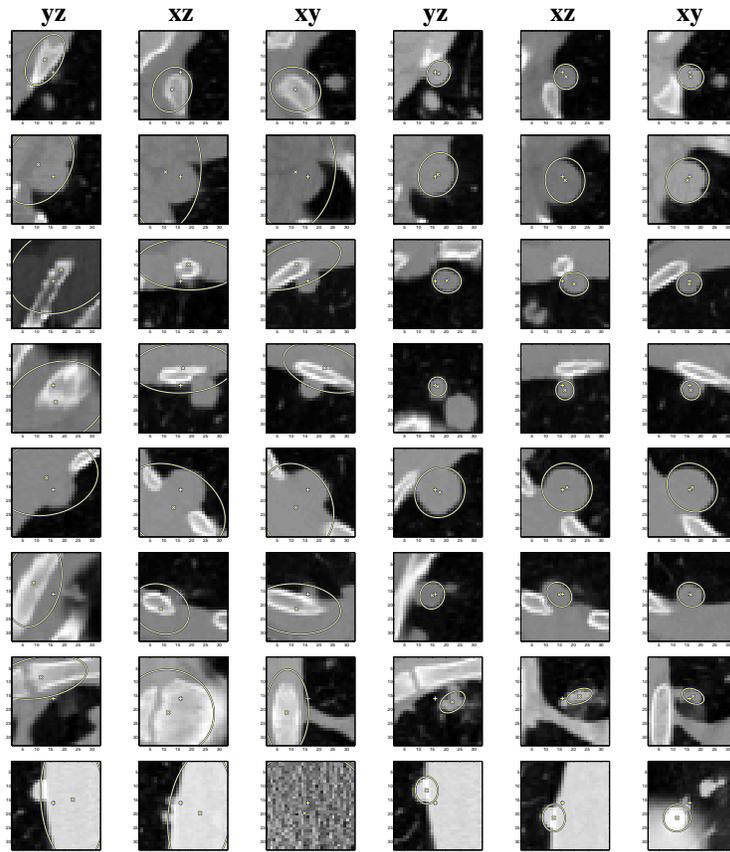


Figure 4: Eight examples of 3D tumor segmentation results with and without the repulsive prior. Each row shows the results for each example. Left 3 columns: failure segmentation due to false data-driven convergence at rib bones shown in 3 orthogonal cross sections. Right 3 columns: corresponding segmentation results using the failed fit as a repelling prior. ”+”: initialization marker \mathbf{x}_p . ”x”: estimated center \mathbf{u} . Ellipses: image-plane intersection of 35% confidence ellipsoid of the estimated Gaussian. *Note that the original and our results are shown in different cross sections thus they appear differently.*

Acknowledgments

The authors wish to thank Alok Gupta and Arun Krishnan for their support and encouragement.

References

- [1] Y. Cheng. Mean shift, mode seeking, and clustering. *IEEE Trans. Pattern Anal. Machine Intell.*, 17(8):790–799, 1995.
- [2] R. T. Collins. Mean-shift blob tracking through scale space. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages II:234–240, 2003.
- [3] D. Comaniciu and P. Meer. Mean shift analysis and applications. In *Int. Conf. Computer Vision*, pages 1197–1203, 1999.

- [4] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Machine Intell.*, 24(5):603–619, 2002.
- [5] D. Comaniciu, V. Ramesh, and P. Meer. Real-time tracking of non-rigid objects using mean shift. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 142–149, 2000.
- [6] M. Fashing and C. Tomasi. Mean shift is a bound optimization. *IEEE Trans. Pattern Anal. Machine Intell.*, 27(3):471–474, 2005.
- [7] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, San Diego, 1990.
- [8] L. Itti and C. Koch. Computational modeling of visual attention. *Nature Reviews Neuroscience*, 2(3):194–203, Mar 2001.
- [9] J. J. Koenderink. The structure of images. *Biol. Cybern.*, 50:363–370, 1984.
- [10] W. J. Kostis, A. P. Reeves, D. F. Yankelevitz, and C. I. Henschke. Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. *IEEE Trans. Medical Imaging*, 22(10):1259–1274, 2003.
- [11] Y. Lee, T. Hara, H. Fujita, S. Itoh, and T. Ishigaki. Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique. *IEEE Trans. Medical Imaging*, 20(7):595–604, 2001.
- [12] X. R. Li, Y. M. Zhu, J. Wang, and C. Z. Han. Optimal linear estimation fusion-part i: unified fusion rules. *IEEE Trans. Information Theory*, 49(9):2192–2208, 2003.
- [13] K. Okada, D. Comaniciu, and A. Krishnan. Robust anisotropic Gaussian fitting for volumetric characterization of pulmonary nodules in multislice CT. *IEEE Trans. Medical Imaging*, 24(3):409–423, 2005.
- [14] K. Okada, V. Ramesh, A. Krishnan, M. Singh, and U. Akdemir. Robust pulmonary nodule segmentation in CT: Improving performance for juxtapleural cases. In *Int. Conf. Medical Image Computing and Computer Assisted Intervention*, 2005.
- [15] A. P. Reeves and W. J. Kostis. Computer-aided diagnosis of small pulmonary nodules. *Seminars in Ultrasound, CT, and MRI*, 21(2):116–128, 2000.
- [16] M. Singh, H. Arora, and N. Ahuja. A robust probabilistic estimation framework for parametric image models. In *Proc. 8th ECCV*, volume 1, pages 508 – 522, 2004.
- [17] J. Sporring, M. Nielsen, P. Johansen, and L. Florack. *Gaussian Scale-Space Theory*. Kluwer Academic Publishers, 1997.
- [18] M. J. Wainwright and M. I. Jordan. Graphical models, exponential families, and variational inference. Technical Report 649, Department of Statistics, University of California, Berkeley, 2003.
- [19] M. P. Wand and M. C. Jones. *Kernel Smoothing*. Chapman & Hall, 1995.
- [20] A. Witkin. Scale-space filtering. In *Int. Joint. Conf. Artificial Intell.*, pages 1019–1021, Karlsruhe, 1983.
- [21] X. S. Zhou, D. Comaniciu, and A. Gupta. An information fusion framework for robust shape tracking. *IEEE Trans. Pattern Anal. Machine Intell.*, 27(1):115–129, 2005.