SIAM J. IMAGING SCIENCES
Vol. 1, No. 1, pp. 115–142

Segmentation under Occlusions Using Selective Shape Prior*

Sheshadri R. Thiruvenkadam†, Tony F. Chan‡, and Byung-Woo Hong§

Abstract. In this work, we address the problem of segmenting multiple objects, under possible occlusions, in a level set framework. A variational energy that incorporates a piecewise constant representation of the image in terms of the object regions and the object spatial order is proposed. To resolve occluded boundaries, prior knowledge of the shape of objects is also introduced within the segmentation energy. By minimizing the above energy, we solve the *segmentation with depth* problem, i.e., estimating the object boundaries, the object intensities, and the spatial order. The segmentation with depth problem was originally dealt with by the *Nitzberg–Mumford–Shiota* (NMS) variational formulation, which proposes segmentation energies for each spatial order. We discuss the relationships and show the computational advantages of our formulation over the NMS model, mainly due to our treatment of spatial order estimation within a *single* energy. A novelty here is that the spatial order information available in the image model is used to dynamically impose prior shape constraints *only* to occluded boundaries. Also presented are experiments on synthetic and real images that have promising results.

Key words. image segmentation, variational methods, level set methods

AMS subject classifications. 65K10, 68U10, 94A08

DOI. 10.1137/070695745

1. Introduction.

Image segmentation is an important step in understanding the composition of the original three-dimensional scene that gave rise to the image. However, it is often considered a difficult problem due to noise, which results in spurious edges and boundary gaps, and occlusions, which lead to an overlap of object boundaries. Low-level visual features such as intensity, color, and texture are generally not sufficient to overcome such difficulties that would make purely bottom-up segmentation approaches unsuccessful. This naturally leads to a need for integrating low-level features and high-level information in segmentation. Enforcing a prior knowledge of the shape of objects is a common way to facilitate segmentation especially under low contrasts, occlusions, and other undesirable noisy conditions.

In this paper, we address the problem of segmenting multiple objects with possible oc-

*Received by the editors June 29, 2007; accepted for publication (in revised form) January 7, 2008; published electronically March 20, 2008. A preliminary version of this work appeared in the conference proceedings of SSVM'07 [40]. This research was supported by ONR grant N00014-06-1-0345 and NSF grant DMS-0610079.
http://www.siam.org/journals/siims/1-1/69574.html

†Department of Mathematics, University of California, Los Angeles, Los Angeles, CA 90095 (sheshad@math.ucla.edu).

‡Department of Mathematics, University of California, Los Angeles, Los Angeles, CA 90095 and Directorate for Mathematics and Physical Sciences, The National Science Foundation, Arlington, VA 22230 (chan@math.ucla.edu).

§Computer Science Department, University of California, Los Angeles, Los Angeles, CA 90095 (hong@cs.ucla.edu). Current address: School of Computer Science and Engineering, Chung-Ang University, Seoul 156-756, Korea. The research of this author was supported by the Ministry of Knowledge Economy, Korea, under the Home Network Research Center–Information Technology Research Center support program supervised by the Institute of Information Technology Assessment.

clusions and recover the object spatial order whenever possible. Here, a segmentation energy incorporating shape prior knowledge is presented in a variational framework. The novelty here is that we identify the spatial order to be encoded within a piecewise constant representation of the given image, and hence are able to recover the spatial order as part of the segmentation process. We demonstrate how the above spatial order information can be used within the energy to impose the shape prior only to occluded boundaries.

Our model solves the *segmentation with depth* problem that aims to determine the boundaries of overlapping objects, along with their spatial ordering, based on intensity distributions in the object regions [1, 18, 21, 29, 30]. The segmentation with depth problem has been addressed before in a variational framework by Nitzberg, Mumford, and Shiota (NMS) in [34, 35], and numerical methods for minimizing the NMS model have been presented in [17, 46]. In the NMS approach, for each spatial order of the objects, a segmentation energy is minimized using curvature continuity to complete occluded boundaries. Then the smallest of the above energy-minima would give the spatial order of the objects along with their boundaries and intensities. In contrast, our formulation involves only *one* segmentation energy with the spatial order given by the estimation of a sequence of constants. Thus the main computational expense in the NMS approach which is of the order of $N.N!$ shape optimizations (for N objects in the scene) is reduced to just N shape optimizations in our case. We will provide a detailed comparison of our model with the NMS approach in the following sections.

In our work, the contours that segment the object boundaries are represented using the *level set method* [15, 16, 36] and are evolved by gradient descent of our variational energy. Deformable models, often called *Snakes*, initially proposed by Kaas, Witkin, and Terzopoulos [23], evolve an explicitly parametrized contour based on external and internal energies. Despite wide use and popularity due to its simplicity and speed, the Snakes approach has the drawbacks of being dependent on the initial contour placement and being restricted by the topology of the contour. There have been some efforts to overcome these limitations [3, 11, 20, 2, 31, 14].

In the level set framework, the contour is represented implicitly as the zero level of a higher dimensional embedding function, and the contour propagation is performed by evolving the embedding function. This enables one to easily handle topological changes of the boundary such as splitting and merging. Segmentation algorithms using the level set method have been developed based on the geometric heat equation taking into account the strength of the edgeness by Caselles et al. [4], Chopp [10], and Malladi, Sethian, and Vemuri [28]. The modified edge-based energy formulations of the above equation, termed *geodesic active contours*, have been proposed by Caselles, Kimmel, and Sapiro [5] and by Kichenassamy et al. [24, 25]. However, these approaches based on local image gradient features are still highly sensitive with respect to the initial contour placement. In order to overcome this undesired behavior, region-based energy formulations have been introduced in [32, 45, 7, 43, 37].

Most of the region-based approaches are based on the popular *Mumford–Shah functional* [32, 33], which finds a piecewise smooth representation for a given noisy image. A level sets–based active-contour implementation of the piecewise constant, two-phase version of the Mumford–Shah functional has been developed by Chan and Vese [7]. Their use of region-based information within an implicit framework gave the advantages of segmenting objects of unknown topology, robustness to noise, discontinuous edges and contour initialization, and detection of interior of objects. The Chan and Vese (CV) model has also been extended to

allow the segmentation of multiple disjoint objects/phases using a multiphase level set formulation [44]. Our segmentation model is similar to some of the above region-based approaches in seeking for a piecewise constant representation of the given image. But in addition to this representation, we also want to recover the actual object boundaries (possibly occluded) along with their spatial order (whenever well defined).

The above segmentation methods rely solely on image intensity, which is not sufficient to detect missing edges that could occur due to occlusions/low-contrast often observed in many practical applications. This consequently leads to the efforts that introduce prior shape information into segmentation schemes [22]. Chen et al. [8, 9] propose a level set segmentation algorithm based on geodesic active contours, incorporating an explicit registration term measuring shape similarity between the segmenting curve and a prior shape that is modeled by the distance function. The active shape model [12] using principal component analysis (PCA) from a set of training shapes has been used as prior shape information. The level set segmentation algorithm incorporating statistical shape priors has been proposed in [43, 41, 27, 42, 39, 19, 26], where PCA is performed based on a set of training shapes in the form of signed distance functions. In contrast to a conventional linear PCA, nonlinear shape statistics by means of kernel PCA are considered in [13]. In most works, a rigid transformation is considered in the comparison of the shape between the evolving level set function and the shape prior model, but a projective transformation is considered by slicing the signed distance function at various angles in [38]. A region-based level set segmentation algorithm with shape priors has been proposed by Chan and Zhu [6].

Although incorporating a shape prior within our segmentation model was inspired by the above works, our method has novelty in that the use of shape prior knowledge is automatically restricted *only* to occluded parts of the object boundaries. That is, the algorithm selectively activates the shape term within the energy functional *only* for occluded regions. Thus, the evolution of the segmenting level set function for the unoccluded regions is solely driven by image intensity, even though the governing energy functional also includes the shape term. This selective use of local prior shape avoids enforcing shape constraints on regions where the object boundary is clearly defined by image intensity.

The focus of this paper is to solve the segmentation from depth problem using prior knowledge of shape of objects. Using the spatial order information, we demonstrate one possible application where the shape constraints are activated only for occluded boundaries. The question of the kind of shape prior knowledge to be used depends on the application and is not dealt with here. In this work, for convenience, we deal with object boundaries that can be characterized by an *explicit prior shape* (given by a binary image). While assuming the availability of such prior knowledge may seem restrictive, its use has been demonstrated before by many shape-based segmentation methods, especially for medical imaging applications.

In a preliminary publication [40] of this work, an unconstrained energy is presented in a level set framework, and the spatial order is postprocessed from the recovered object intensities. In this paper, we formulate our model as a constrained energy and first show its equivalence with the NMS model. Second, we discuss the differences in spatial order estimation for our model and in the NMS model, and how this leads to a computationally less expensive implementation in our case. Finally, a comparison of results using the constrained energy and its simplified, unconstrained version is dealt with.

To summarize, we present a piecewise constant image representation of an occlusion scene, in terms of the object regions and their intersection regions. We then identify from the above representation that the sequence of constants corresponding to the intersection regions is unique to the object spatial order. The inverse problem is given by a constrained energy minimization that estimates the object boundaries, intensities, and the sequence of constants which determines the spatial order. We show that this constrained minimization is identical to the NMS minimization, but is computationally less expensive. Also, in our experiments, we approximate the above problem as an unconstrained energy minimization and have obtained identical results for images corrupted with white noise. A key contribution here is that the spatial order information available through the sequence of constants can be used within the energy to dynamically impose the shape prior only to occluded boundaries.

2. Nitzberg–Mumford–Shiota formulation. We will briefly review the related formulation of NMS [35] for segmentation with depth. In a later section, we will present detailed comparisons with our approach. We start with some basic definitions and assumptions. Suppose that $I : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^2$ is a two-dimensional image of a scene composed of N objects $\{O_p\}_{p=1}^N$. We define an *occlusion* relation ">" on object indices given by $i > j$ when O_i is in front of O_j (from the viewer's perspective). We denote a *spatial order* of the objects by the ordered set of object indices $Q^k = \{i_1 > i_2 > \cdots > i_N\}$, and denote by $\mathbf{Q} = \{Q^1, Q^2, \ldots, Q^{N!}\}$ the set of $N!$ possible spatial orders. Let $\{A_p\}_{p=1}^N$ be the regions formed on the image domain by the objects. Now suppose that we have the following assumptions:

- The image intensity formed by the objects O_k is close to a constant c_k, and the background intensity is close to a constant $\tilde{c}$.
- The objects are not twisted between themselves.

Due to the assumptions listed above, for a spatial order $\{i_1 > i_2 > \cdots > i_N\}$, $A_k - \cup_{j>k} A_j$ is the visible portion of the object O_k, with image intensity close to a constant c_k. For the topmost object, O_{i_1}, let $\cup_{j>i_1} A_j = \phi$. Thus for a given spatial order Q^k, the image I of the scene can be represented in terms of the visible object regions,

$$I = \sum_{q=1}^{N} c_q \chi_{A_q - \cup_{j>q} A_j} + \tilde{c} \prod_{q=1}^{N} (1 - \chi_{A_q}), \tag{2.1}$$

where χ_S is the characteristic function of a set S.

2.1. NMS minimization. Given an image I_0, for each spatial ordering Q^k of objects, $(k = 1, 2, \ldots, N!)$, the NMS segmentation energy $\tilde{E}_{Q^k}[\mathbf{A}, C_{obj}]$ is written with respect to object regions A_p, constant intensities c_p, and background intensity $\tilde{c}$. Denote $\mathbf{A} = \{A_1, A_2, \ldots, A_N\}$, $C_{obj} = \{c_1, c_2, \ldots, c_N, \tilde{c}\}$,

$$\begin{aligned} \tilde{E}_{Q^k}[\mathbf{A}, C_{obj}] = \sum_{p=1}^{N} \int_{A_p - \cup_{j>p} A_j} (I_0 - c_p)^2 dx + \int_{\cap_{k=1}^{N} A_k^c} (I_0 - \tilde{c})^2 dx \\ + \sum_{p=1}^{N} \lambda \int_{\partial A_p} ds + \beta \int_{\Omega} \hat{S}(A_p) dx, \end{aligned} \tag{2.2}$$

where A_k^c is the complement of object region A_k, and λ and β are parameters that balance the formula.

Here, the fourth term is the prior knowledge that the NMS model incorporates to resolve occlusions. Specifically in the NMS model, $\hat{S}$ is used to complete occluded boundaries by continuity of curvature, and is given by

$$\hat{S}(A_p) = \phi(\kappa_p), \tag{2.3}$$

where the function $\phi(x)$ is to be a positive, convex, even function. κ_p is the curvature of ∂A_p. Our model also uses a prior shape term to resolve occlusions, although for a different choice of $\hat{S}$.

The NMS model is the minimization problem

$$\min_{k\in\{1,2,\ldots,N!\}} \min_{\mathbf{A},C_{obj}} \tilde{E}_{Q^k}. \tag{2.4}$$

A minimizer m determines the spatial ordering Q^m, the object regions $\mathbf{A}^m$, and the intensities C_{obj}^m.

3. Occlusion model. In the NMS model (2.4), the spatial order estimation is done via $N!$ segmentation energy minimizations and turns out to be computationally expensive in spite of a close initial guess for the object regions. The reason is that the spatial order is not characterized within the image model (2.1) and hence cannot be "adapted" based on the object regions.

To deal with the issue of spatial order estimation, we want to first look at an image representation that is *rich* enough to hold information on the object regions, object intensities, and the spatial order. In the following discussion, we observe a dependence between intensities in the intersection regions and the spatial order. Later, we utilize this dependence to incorporate spatial order within a *single* segmentation energy.

To motivate the form of the image I, in the case $N = 3$, suppose that, without loss of generality (w.l.o.g.), $1 > 3 > 2$ is the spatial order of the objects, and A_1, $A_3 - A_1$, and $A_2 - (A_1 \cup A_3)$ are the parts of the objects O_1, O_3, and O_2 visible in the image (Figure 1). Using the principle of inclusion and exclusion for (2.1) gives

$$\begin{aligned} I &= c_1\chi_{A_1} + c_2\chi_{A_3-A_1} + c_3\chi_{A_2-(A_1\cup A_3)} + \tilde{c}(1-\chi_{A_1})(1-\chi_{A_2})(1-\chi_{A_3}) \\ &= c_1\chi_{A_1} + c_2\chi_{A_2} + c_3\chi_{A_3} - c_2\chi_{A_1\cap A_2} - c_2\chi_{A_2\cap A_3} - c_3\chi_{A_1\cap A_3} + c_2\chi_{A_1\cap A_2\cap A_3} \\ &\qquad + \tilde{c}(1-\chi_{A_1})(1-\chi_{A_2})(1-\chi_{A_3}). \end{aligned} \tag{3.1}$$

In the above expression, for each intersection region, the corresponding constant factor equals the intensity of the farthest object in that intersection. Hence the spatial order is *encoded* within the sequence of intersection constants $\{c_2, c_2, c_3, c_2\}$.

For the general case of N objects with a spatial order $i_1 > i_2 > \cdots > i_N$, the image I of the scene in terms of the visible object regions is given by (2.1). Similarly to (3.1), rewriting the expression in terms of the visible regions A_k and their intersection regions gives the following general form of I, valid for *any* spatial ordering between the objects:

$$I = \sum_{p=1}^{N} c_p\chi_{A_p} + \sum_{p=2}^{N}\sum_{k=1}^{\binom{N}{p}}(-1)^{p-1}c_{p,k}\chi_{P_{p,k}} + \tilde{c}\prod_{k=1}^{N}(1-\chi_{A_k}). \tag{3.2}$$

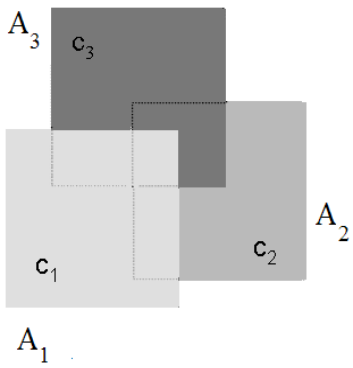

Figure 1. *Image formed for three objects, A_k with constant intensity c_k, for spatial order $1 > 3 > 2$. The dotted lines are the occluded boundaries of A_2 and A_3.*

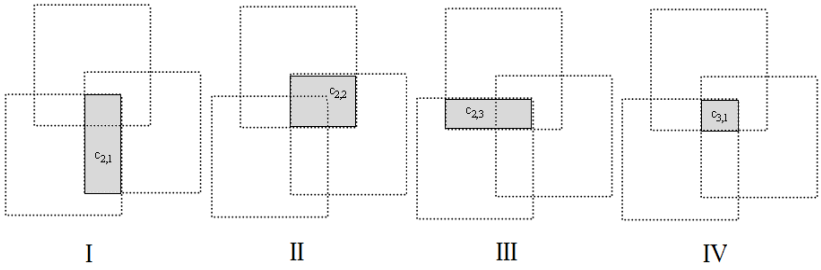

Figure 2. *For spatial order $1 > 3 > 2$ as in Figure 1 $\left(c_{2,1} = c_2,\ c_{2,2} = c_2,\ c_{2,3} = c_3,\ c_{3,1} = c_2\right)$. Intersection regions $P_{p,k}$ (gray) and corresponding constants $c_{p,k}$.* (I) $P_{2,1}$, (II) $P_{2,2}$, (III) $P_{2,3}$, (IV) $P_{3,1}$.

Here, $P_{p,k}$ $\left(p = 2, 3, \ldots, N, k = 1, 2, \ldots, \binom{N}{p}\right)$ is the kth unordered *intersection* of p *regions* from $A_1, A_2, \ldots, A_N$. $c_{p,k}$ are positive constants and take one of the values $c_1, c_2, \ldots, c_N$. In fact, $c_{p,k}$ takes the intensity of the farthest object appearing in the intersection, $P_{p,k}$. For $N = 3$, the regions $P_{p,k}$ and the constants $c_{p,k}$ are shown in Figure 2.

Thus, there are $N!$ possible sequences for $c_{p,k}$, with each sequence corresponding to a unique spatial ordering of the objects. For $i = 1, 2, \ldots, N!$, let S^i denote the sequence of constants $c_{p,k}$ corresponding to the spatial ordering of the objects, Q^i. We denote $\mathbf{S} = \{S^1, S^2, \ldots, S^{N!}\}$, the set of *spatial order sequences* corresponding to the set of spatial orders $\mathbf{Q} = \{Q^1, Q^2, \ldots, Q^{N!}\}$.

For the $N = 2$ case,

$$I = c_1\chi_{A_1} + c_2\chi_{A_2} - c_{2,1}\chi_{A_1\cap A_2} + \tilde{c}(1 - \chi_{A_1})(1 - \chi_{A_2}). \tag{3.3}$$

Here, $c_{2,1} = c_2$ if and only if $1 > 2$, and $c_{2,1} = c_1$ if and only if $2 > 1$.

Similarly for $N = 3$,

$$\begin{aligned} I = {} & c_1\chi_{A_1} + c_2\chi_{A_2} + c_3\chi_{A_3} - c_{2,1}\chi_{A_1\cap A_2} - c_{2,2}\chi_{A_2\cap A_3} - c_{2,3}\chi_{A_1\cap A_3} + c_{3,1}\chi_{A_1\cap A_2\cap A_3} \\ & + \tilde{c}(1 - \chi_{A_1})(1 - \chi_{A_2})(1 - \chi_{A_3}). \end{aligned} \tag{3.4}$$

Table 1

$N = 3$ case: Constants $c_{p,k}$ corresponding to intersection regions $P_{2,1} = A_1 \cap A_2$, $P_{2,2} = A_2 \cap A_3$, $P_{2,3} = A_1 \cap A_3$, $P_{3,1} = A_1 \cap A_2 \cap A_3$.

Spatial order Q^k	$S^k = \{c_{2,1}, c_{2,2}, c_{2,3}, c_{3,1}\}$
$1 > 2 > 3$	$\{c_2, c_3, c_3, c_3\}$
$1 > 3 > 2$	$\{c_2, c_2, c_3, c_2\}$
$2 > 1 > 3$	$\{c_1, c_3, c_3, c_3\}$
$2 > 3 > 1$	$\{c_1, c_3, c_1, c_1\}$
$3 > 2 > 1$	$\{c_1, c_2, c_1, c_1\}$
$3 > 1 > 2$	$\{c_2, c_2, c_1, c_2\}$

Table 1 shows the values $c_{p,k}$ for possible occlusion scenarios.

4. Energy formulation. As seen above, (3.2) gives a piecewise constant representation of the image in terms of the objects in the scene, along with their spatial order. Given an image I_0, we solve the inverse problem of recovering the object regions $\{A_p\}_{p=1}^N$, the object intensities $c_1, c_2, \ldots, c_N$, the background intensity $\tilde{c}$, and the intersection intensities $c_{p,k}$. Denote $\mathbf{A} = \{A_1, A_2, \ldots, A_N\}$, $C_{obj} = \{c_1, c_2, \ldots, c_N, \tilde{c}\}$, and sequence $C_{int} = \{c_{p,k} |\ p = 2, 3, \ldots, N,\ k = 1, 2, \ldots, \binom{N}{p}\}$. The sequence C_{int} is constrained to adhere to one of the $N!$ possible spatial order sequences S^i. We formulate the above problem as the following energy minimization. Here, a *prior shape* term is introduced within the energy to resolve occluded object boundaries:

$$E[\mathbf{A}, C_{obj}, C_{int}] = \int_\Omega \left(I_0 - \left(\sum_{p=1}^{N} c_p \chi_{A_p} + \sum_{p=2}^{N} \sum_{k=1}^{\binom{N}{p}} (-1)^{p-1} c_{p,k} \chi_{P_{p,k}} + \tilde{c} \prod_{k=1}^{N} (1 - \chi_{A_k}) \right) \right)^2 dx + \sum_{p=1}^{N} \left(\lambda \int_{\partial A_p} ds + \beta \int_\Omega \hat{S}(A_p) dx \right)$$

such that $C_{int} \in \mathbf{S} = \{S^1, S^2, \ldots, S^{N!}\}$.
(4.1)

We note that the above energy is quadratic with respect to the constants C_{obj} and C_{int}. Since the sequence of constants S^k corresponds to spatial order Q^k, a minimizing C_{int}^{min} for the energy would determine a spatial order for the objects. Whether a unique solution C_{int}^{min} exists or not would depend on the number of nonempty intersection regions between objects $\mathbf{A}$. However, as in many cases discussed in section 7, even partial information of elements of the sequence C_{int}^{min} is sufficient to give a unique projection onto the set $\mathbf{S}$ (thus solving the segmentation with depth problem). For example, we see from Table 1 that just estimating the first two elements of C_{int}, $c_{2,1} = c_2$ and $c_{2,2} = c_3$, would determine the spatial order as $1 > 2 > 3$.

The second term is a length regularization term, commonly used in shape optimization problems, and the third term constrains the shape of the boundaries of A_p. There are various

choices for the shape term such as Euler's elastica energy (to enforce curvature continuity of A_p) and statistical/explicit shape priors depending on the kind of application. In this work, we mainly consider the application of shape explicitly, with the term $\int_\Omega \hat{S}(A_p)$ measuring the area dissimilarity of the object A_p from a given binary shape S. λ and β are parameters that balance the formula. Later, in section 8, we also show how one can selectively impose shape constraints only to occluded object boundaries.

In our work, for computational simplicity, we have assumed the constants $c_{p,k}$ to be independent and minimize (4.1) without constraints:

$$\min_{\mathbf{A}, C_{obj}, C_{int}} E. \tag{4.2}$$

Then, the minimizing constants C^m_{int} can be postprocessed to deduce a unique spatial order for images with unambiguous occlusions. Our experiments indicate that, for images with uncorrelated noise, the above simplified approach (4.2) takes solutions very close to that of the original constrained minimization problem (4.1).

5. Related works. The form of (4.1) indicates its relationship to some of the region-based segmentation approaches, e.g., [44]. Without the constraint on C_{int} and for $\beta = 0$, (4.1) is just the *cartoon* Mumford–Shah functional that seeks a piecewise constant approximation to a given I_0. But this is not sufficient to recover object boundaries that are occluded. For a nonzero β, one can also recover occluded boundaries due to the shape term, and the estimated C_{int} in many cases defines a unique spatial order. The constraint on C_{int} gives the model additional robustness to the segmentation of nonwhite noise images. However, as observed in section 11, the minimizers obtained for white noise images do not depend on the constraint.

5.1. Comparisons with the NMS model. Here, we compare our energy (4.1) with the NMS model (2.4). The NMS segmentation energy for a spatial order Q^k, (2.2), is given by

$$\begin{aligned}\tilde{E}_{Q^k}[\mathbf{A}, C_{obj}] = &\sum_{p=1}^{N} \int_{A_p - \cup_{j>p} A_j} (I_0 - c_p)^2 dx + \int_{\cap_{k=1}^{N} A_k^c} (I_0 - \tilde{c})^2 dx \\ &+ \sum_{p=1}^{N} \lambda \int_{\partial A_p} ds + \beta \int_\Omega \hat{S}(A_p) dx.\end{aligned} \tag{5.1}$$

The first term of the above energy is rewritten as

$$\begin{aligned}\sum_{p=1}^{N} \int_{A_p - \cup_{j>p} A_j} (I_0 - c_p)^2 dx &= \int_\Omega \left(I_0 - \sum_{p=1}^{N} c_p \chi_{A_p - \cup_{j>p} A_j} \right)^2 dx \\ &= \int_\Omega \left(I_0 - \left(\sum_{p=1}^{N} c_p \chi_{A_p} + \sum_{p=2}^{N} \sum_{k=1}^{\binom{N}{p}} (-1)^{p-1} c_{p,k} \chi_{P_{p,k}} \right) \right)^2 dx,\end{aligned} \tag{5.2}$$

where the above sequence, $C_{int} = \{c_{p,k} |\ p = 2, 3, \ldots, N,\ k = 1, 2, \ldots, \binom{N}{p}\} = S^k$, corresponds to the spatial order Q^k. Thus we have

$$\tilde{E}_{Q^k}[\mathbf{A}, C_{obj}] = E[\mathbf{A}, C_{obj}, C_{int} = S^k]. \tag{5.3}$$

It follows that

$$\begin{aligned}\min_{k\in\{1,2,\ldots,N!\}} \min_{\mathbf{A},C_{obj}} \tilde{E}_{Q^k} &= \min_{k\in\{1,2,\ldots,N!\}} \min_{\mathbf{A},C_{obj},C_{int}=S^k} E \\ &= \min_{\mathbf{A},C_{obj},C_{int}\in\{S^1,S^2,\ldots,S^{N!}\}} E. \end{aligned} \tag{5.4}$$

Hence the NMS model (2.4) and the minimization of (4.1) are equivalent. The key contribution here is that we have written the $N!$ minimizations of energies $\tilde{E}_{Q^k}[\mathbf{A}, C_{obj}]$ as a *single* energy minimization of $E[\mathbf{A}, C_{obj}, C_{int} \in \{S^1, S^2, \ldots, S^{N!}\}]$ by introducing additional variables C_{int} to estimate the spatial ordering of the objects. The main computational expense in both models (4.1) and (2.2) is the minimization with respect to regions $\mathbf{A}$. The NMS model is more expensive because of the $N.N!$ shape minimizations it has to deal with, in contrast to just N shape minimizations in our case. Further, in the NMS model, in spite of a close initial guess for object boundaries, one has to go through each of the $N!$ minimizations in order to find the correct spatial order. Whereas our energy, which is quadratic with respect to the variable C_{int}, allows one to utilize projection–gradient-type algorithms for fast convergence to the correct spatial order even for rough initial guesses for object boundaries. The problem is further simplified for white noise images when the unconstrained energy (4.2) can be used and the estimated C_{int} can be postprocessed to infer the spatial order.

We will further illustrate the above differences through examples in section 11. In the subsequent sections, we discuss the implementation details and experimental results for the unconstrained energy (4.2).

6. Level set formulation. The regions A_k are represented as the interior of level set functions ϕ_k, i.e., $H(\phi_k) = \chi_{A_k}$, $k = 1, 2, \ldots, N$, where $H(t)$ is the Heaviside function. Thus, we try to recover ϕ_k from I_0. Let $\Phi = (\phi_1, \phi_2, \ldots, \phi_N)$ and, as before, $C_{obj} = \{c_1, c_2, \ldots, c_N, \tilde{c}\}$ and sequence $C_{int} = \{c_{p,k} |\ p = 2, 3, \ldots, N,\ k = 1, 2, \ldots, \binom{N}{p}\}$. We reformulate the problem (4.2) as the minimization with respect to Φ and C_{obj}, C_{int}, of the following energy:

$$\begin{aligned} E[\Phi, C_{obj}, C_{int}] = \int_\Omega \left(I_0 - \sum_{p=1}^{N} c_p\, H(\phi_p) - \sum_{p=2}^{N} \sum_{k=1}^{\binom{N}{p}} (-1)^{p-1} c_{p,k} S_{p,k} - \tilde{c}\tilde{S}_b \right)^2 dx \\ + \lambda \int_\Omega \sum_{k=1}^{N} |\nabla H(\phi_k)| + \beta \int_\Omega \sum_{k=1}^{N} \hat{S}(\phi_k) dx, \end{aligned} \tag{6.1}$$

where $S_{p,k}$ is the kth unordered product of p functions from $H(\phi_1), H(\phi_2), \ldots, H(\phi_N)$ and $\tilde{S}_b = \prod_{k=1}^{N} (1 - H(\phi_k))$. The second term regularizes Φ, and the last term constrains the shape of the 0 level set of ϕ_k. λ and β balance the formula.

For simplicity, we will illustrate the $N = 2$ case. The above energy reduces to

$$E[\phi_1, \phi_2, c_1, c_2, c_{2,1}, \tilde{c}]$$
$$= \int_\Omega \left(I_0 - (c_1 H(\phi_1) + c_2 H(\phi_2) - c_{2,1} H(\phi_1) H(\phi_2) + \tilde{c}(1 - H(\phi_1))(1 - H(\phi_2)))\right)^2 dx$$
$$+\lambda \left(\int_\Omega |\nabla H(\phi_1)| + \int_\Omega |\nabla H(\phi_2)| \right) + \beta \int_\Omega \hat{S}(\phi_1) + \hat{S}(\phi_2) dx. \tag{6.2}$$

The fourth term is used to impose shape-based constraints, such as curvature, and explicit shape on the objects $H(\phi_i)$. This term is used to avoid local minima of the first term in (6.2) that occur particularly under occlusions. In our implementation, we deal only with imposing constraints on length (to resolve linear edges that are occluded) and explicit shape (given by a binary image).

Note. In (6.1), the N level set functions are used to partition Ω into 2^N piecewise constant, disjoint regions, i.e., N visible regions of the objects; $2^N - N - 1$ occlusion regions; and the background. Here N level set functions are used for N objects to allow for the maximum possible number of $2^N - N - 1$ occlusions between the objects. In fact, to represent an occlusion scenario of N objects with M occlusions ($M \leq 2^N - N - 1$), and the background, we need an optimal number of $\log_2(N+M+1)$ level set functions. For instance, the multiphase version of the CV model assumes the objects/phases to be disjoint and hence needs only $\log_2 N$ (the background is also treated as an object) level set functions to partition Ω.

7. Spatial order estimation.

Once (6.1) is minimized, the solutions C_{int}^{min} and C_{obj}^{min} are compared to infer the spatial order. To start, we consider nonempty intersection regions to infer the occlusion relation ">" for possible object pairs. Then, for certain classes of images without occlusion ambiguities and with sufficient number of nonempty object intersections, the occlusion relation determines the spatial order.

7.1. Disoccluding intersection regions.

First, the mean intensities in nonempty intersection regions are compared with the object intensities to infer the object in front. For instance, in Figure 3, we can infer that $1 > 3$ since the mean intensity in P_1 is closer to c_1 than to c_3. A similar comparison of the mean intensity in P_4 with c_2 and c_3 gives $1 > 2$ and $1 > 3$. For the mean intensity in an intersection to give consistent information about object occlusions, we need the following assumptions:

- The object intensity in the regions A_p is close to a constant c_p. Further, objects that occlude each other have *different* intensities. This allows us to identify, based on mean intensity, the topmost object in an intersection region.
- Objects do not twist between themselves; i.e., if O_i occludes O_j in some region, O_i is *not occluded* by O_j in any region. This is to define a valid order relation for the objects based on their intersection regions.

Then the region covering all the objects, $A = \bigcup_{p=1}^N A_p$, can be written as a disjoint union of $\sum_{n=1}^N \binom{N}{n} = 2^N - 1$ regions, where the image intensity is close to a constant in each of those regions. Of these regions, N regions are the visible, nonintersecting parts of the objects O_k, with intensity c_k. The rest of the $2^N - N - 1$ regions is where occlusions can possibly occur. The intensities in these regions are close to one of the object intensities, c_k, i.e., the intensity

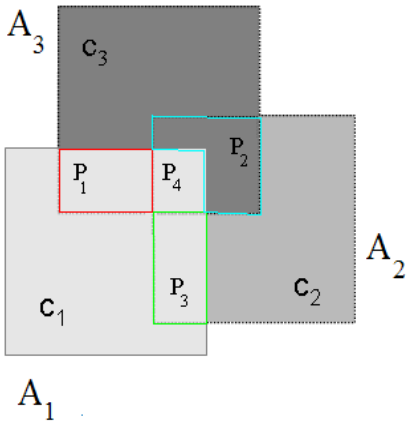

Figure 3. *Spatial order estimation. Nonempty, disjoint intersections* $P_1 = (A_1 \cap A_3) - A_2$, $P_2 = (A_2 \cap A_3) - A_1$, $P_3 = (A_1 \cap A_2) - A_3$, $P_4 = A_1 \cap A_3 \cap A_2$ *shown.*

of the object in front. So we look at the mean intensity μ_P in a nonempty intersection region P and infer that the object, say O_f, whose intensity c_f is the closest to μ_P, is in front of the other objects that occur in the intersection.

If P is an intersection of p objects, w.l.o.g.,

$$P = \{\cap_{s=1}^{p} A_s\} - \{\cup_{s=p+1}^{N} A_s\}, \quad p > 1, \tag{7.1}$$

we can infer the object in front, O_f, from

$$f = \arg\min_{1 \le s \le p} (\mu_P - c_s)^2. \tag{7.2}$$

Thus inspecting the intersection P gives $f > s$ $(s = 1, 2, \ldots, f-1, f+1, \ldots, p)$. Also, μ_P, the mean image intensity in P, can be linearly expressed in terms of the constants $c_q,\ c_{q,k},\ q \le p$. To see this, from (3.2),

$$\begin{aligned} I &= \sum_{P \subset A_q} c_q \chi_{A_q} + \sum_{P \subset P_{q,k}} (-1)^{q-1} c_{q,k} \chi_{P_{q,k}} + I\chi_{P^c} \\ &= \left(\sum_{P \subset A_q} c_q + \sum_{P \subset P_{q,k}} (-1)^{q-1} c_{q,k} \right) \chi_P + I\chi_{P^c}. \end{aligned} \tag{7.3}$$

It follows that $\mu_P = \frac{\int_P I dx}{\int_P dx}$ is a linear expression in terms of $c_q,\ c_{q,k},\ q \le p$.

To illustrate the procedure, for the $N = 2$ case, from (3.2) we have

$$\begin{aligned} I &= c_1 \chi_{A_1} + c_2 \chi_{A_2} - c_{2,1} \chi_{A_1 \cap A_2} + \tilde{c} \chi_{A_1^c \cap A_2^c} \\ &= c_1 \chi_{A_1 - A_2} + c_2 \chi_{A_2 - A_1} + (c_1 + c_2 - c_{2,1}) \chi_{A_1 \cap A_2} + \tilde{c} \chi_{A_1^c \cap A_2^c}, \end{aligned} \tag{7.4}$$

where A_1^c and A_2^c are the complements of A_1 and A_2. Thus, if $P = A_1 \cap A_2 \neq \{\phi\}$, then

$$\mu_P = \frac{\int_P I dx}{\int_P dx} = c_1 + c_2 - c_{2,1}. \tag{7.5}$$

To infer the occlusion relation for O_1 and O_2, we notice that if

$$(c_2 - c_{2,1})^2 = (\mu_P - c_1)^2 < (\mu_P - c_2)^2 = (c_1 - c_{2,1})^2, \tag{7.6}$$

then O_1 occludes O_2 $(1 > 2)$ and vice versa.

7.2. Spatial order. Previously, we deduced the occlusion relation from intersection regions. Now for the above relation to actually give us a unique spatial ordering of the objects, we need the following assumptions:

- For the objects O_i, O_j, O_k, if $i > j$ and $j > k$, then $i > k$ (i.e., *transitivity*). This falls in line with the assumption that the objects do not twist; hence $k \not> i$. This assumption also allows us to extend the relation to objects that do not directly intersect.
- Each pair of objects can be compared (either directly or through the transitivity assumption). In other words, there is a sufficient number of relevant, nonempty intersection regions to uniquely determine the spatial order.

To summarize, after minimizing (6.1), we look at the mean intensities (which depend linearly on the minimizers C_{int}^{min} and C_{obj}^{min}) in the object intersection regions to define the occlusion relation. When the given image satisfies the assumptions listed above, a unique spatial order for the objects is recovered. In Figure 3, by inspection of P_1 and P_2, $1 > 3$ and $3 > 2$ follow. By transitivity, the spatial order $1 > 3 > 2$ follows.

Note 1. In this work, due to the constant intensity assumption in the regions A_p, we use mean intensity to test for occlusions. Hence we used the distance measure $D(I_0|_P, I_0|_{A_s}) = (\mu_P - c_s)^2$ in (7.2) to compare the image intensity in the intersection region P with the intensity in an object region A_s. D can be easily extended to other distance measures to compare general image distributions.

Note 2. After the 2^N constants C_{obj}, C_{int} are computed, one obvious way to infer the spatial order is to directly compare the sequence C_{int} with the $N!$ sequences of the set $\mathbf{S}$, which has a high complexity $O(2^N N!)$. In the projection method described here, a factorial cost is avoided by using mean intensities in nonempty intersection regions to infer the order. There are two steps in our method. First, the cost of converting C_{int} into mean intensities is $O(2^{2N})$ (since the transformation matrix is triangular (7.3)). Second, in (7.2) we compare the mean intensity in nonempty intersections with the object intensities to infer the object in front, which requires p computations for a p-object intersection. Finding the object in front for a sufficient number of intersection regions determines the order. The above procedure's cost is bounded by the complexity of finding the object in front for each intersection region, which is $O(N2^N)$. Thus our projection method is $O(2^{2N} + N2^N) = O(2^{2N})$.

8. Selective shape term. Since we are dealing with occlusions in multiobject segmentation, prior shape information of the objects is used to fill in missing boundaries in occluded regions. However, just adding a shape term as in (6.2) means that the shape term might influence boundary shapes even in unoccluded regions, where the boundary is unambiguously

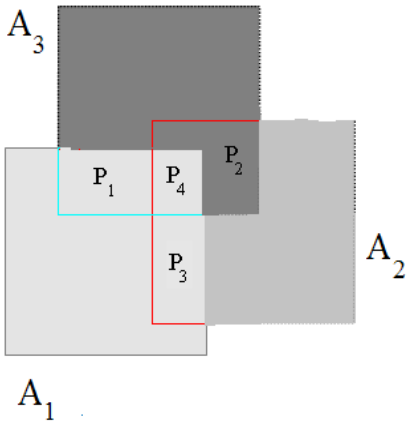

Figure 4. *Selective use of shape. We want the shape to influence* only *the red and cyan curves (occluded boundaries of A_2 and A_3).*

defined by image intensity. Hence, we introduce our shape term in a selective manner. That is, the shape term is allowed to take effect only for occluded boundaries. From the previous section, we see that the constants C_{int} encode spatial order information, which we use here to impose shape constraints selectively.

To motivate the term, in Figure 4, we want the shape term to *influence only* the red and cyan curves (occluded boundaries of A_2 and A_3). Let $P_1 = A_1 \cap A_3 - A_2$, $P_2 = A_2 \cap A_3 - A_1$, $P_3 = A_1 \cap A_2 - A_3$, and $P_4 = A_1 \cap A_3 \cap A_2$. To construct a selective shape term that influences only occluded boundaries, we consider each region P_k and, based on the constants C_{int}, apply shape to the part of P_k's boundary that is occluded. We are also careful to avoid repetitions in our terms for occluded boundaries across P_k's. Thus we consider only the A_1, A_3 boundaries of P_1, the A_2, A_3 boundaries of P_2, and so on.

In the given example in Figure 4, for $P_1 = A_1 \cap A_3 - A_2$, we want to influence only the part of the boundary of P_1 which belongs to the occluded object, i.e., the A_3 *boundary*. Similarly, we want to influence only the A_2 boundary of P_2, the A_2 boundary of P_3, and the A_2,A_3 boundaries of P_4. For P_1, a comparison of μ_{P_1} with c_1 and c_3 would tell us that A_3 is occluded. Then to influence the shape of only the A_3 boundary of P_1, we can restrict the shape term to only the region $A_1 - A_2$. Thus for P_1, we can use a local shape term,

$$\int_{A_3 - A_2} (\mu_{P_1} - c_1)^2 \hat{S}(A_1)\, dx + \int_{A_1 - A_2} (\mu_{P_1} - c_3)^2 \hat{S}(A_3)\, dx. \tag{8.1}$$

Since μ_{P_1} is very close to c_1, the first term is close to zero, and the shape term does not take effect for A_1. For the second term, $(\mu_{P_1} - c_3)^2$ is considerably larger, so the shape term for A_3 is active in the region $A_1 - A_2$ and hence influences only the A_3 boundary of P_1. Similarly to the above expression, we can define local shape terms for P_2, P_3, and P_4. Thus we would have a local influence of shape only to the occluded boundaries of A_2 and A_3.

We follow the above idea for the general case. For each intersection region of p objects,

$p > 1$, w.l.o.g.,

$$P = \{\cap_{k=1}^{p} A_k\} - \{\cup_{k=p+1}^{N} A_k\}, \tag{8.2}$$

consider the shape term

$$\sum_{s=1}^{p} \int_{P_s} (\mu_P - c_s)^2 \hat{S}(A_s)\, dx. \tag{8.3}$$

Here, $P_s = \bigcap_{\substack{t=1 \\ t\neq s}}^{p} A_t - \{\cup_{s=p+1}^{N} A_s\}$. The above shape term localizes use of shape only to boundaries of P that belong to occluded objects. First, the terms $\hat{S}(A_s)$, which constrain the shape of A_s, are weighted by $(\mu_P - c_s)^2$, which is larger for occluded objects and is minimal for the object that is in front. Second, the shape term (8.3) is defined only on P_s, the region that occludes the A_s boundary of P.

Now, for $N = 2$, we present the modified energy with the local shape term. For the intersection $P = A_1 \cap A_2$ with mean intensity μ, the local shape term defined by (8.3) in level set formulation is

$$\int_{\Omega} H(\phi_2)(\mu - c_1)^2 \hat{S}(\phi_1) + H(\phi_1)(\mu - c_2)^2 \hat{S}(\phi_2) dx. \tag{8.4}$$

Denote $I = c_1 H(\phi_1) + c_2 H(\phi_2) - c_{2,1} H(\phi_1) H(\phi_2) + \tilde{c}(1 - H(\phi_1))(1 - H(\phi_2))$. Since $\mu = c_1 + c_2 - c_{2,1}$ from (7.5), the energy (6.2) becomes, after adding the local shape term,

$$\begin{aligned} &E[\phi_1, \phi_2, c_1, c_2, c_{2,1}, \tilde{c}] = \\ &\int_{\Omega} (I_0 - I)^2 dx + \lambda \left(\int_{\Omega} |\nabla H(\phi_1)| + \int_{\Omega} |\nabla H(\phi_2)| \right) + \beta \int_{\Omega} \hat{S}(\phi_1) + \hat{S}(\phi_2) dx \\ &\qquad + \tilde{\beta} \int_{\Omega} H(\phi_2)(c_{2,1} - c_2)^2 \hat{S}(\phi_1) + H(\phi_1)(c_{2,1} - c_1)^2 \hat{S}(\phi_2) dx. \end{aligned} \tag{8.5}$$

Here, the fourth term is the shape term used to globally influence the shape of the segmented objects to avoid local minima. β and $\tilde{\beta}$ balance the shape terms with $\tilde{\beta} \gg \beta$.

9. Numerical implementation. In this paper, given the binary image S of a prior shape, we use the symmetric area measure to compare shapes. Hence, in (6.1), the shape term is $\hat{S}(\phi_k) = (H(\phi_k) - S \circ T_k)^2$, where $S \circ T_k$ is the deformed image S under rigid transformations T_k, which have to be determined during minimization. To minimize (6.1), we use a finite difference scheme to solve the resulting Euler–Lagrange equations.

We present the numerical implementation for the $N = 2$ case, shown in (8.5). Denote $T = [T_1, T_2]$, where $T_k = [\mu_k, \theta_k, \mathbf{t}_k]$, $k = 1, 2$, are rigid transformations with scale μ_k, rotation θ_k, and translation $\mathbf{t}_k$. Also, $\Phi = [\phi_1, \phi_2]$, and $C = [c_1, c_2, c_{2,1}, \tilde{c}]$. Rewriting (8.5) using the shape term shown above,

$$E[\Phi, C, T]$$

$$= \int_\Omega (I - I_0)^2 dx + \lambda \left(\int_\Omega |\nabla H(\phi_1)| + \int_\Omega |\nabla H(\phi_2)| \right) + \int_\Omega \{\beta + \tilde{\beta} H(\phi_2)(c_{2,1} - c_2)^2\}$$

$$(H(\phi_1) - S \circ T_1)^2 dx + \int_\Omega \{\beta + \tilde{\beta} H(\phi_1)(c_{2,1} - c_1)^2\}(H(\phi_2) - S \circ T_2)^2 dx. \tag{9.1}$$

Given (Φ, T), the minimizing constants C of the above energy are easily computed as the solution of a triangular linear system. For an index $k \in \{1, 2\}$, let $\bar{k}$ denote its complement. For $k = 1, 2$, the Euler–Lagrange equations for (9.1) are

$$\delta(\phi_k)\{(I - I_0)(c_{\bar{k}} - c_{2,1} H(\phi_{\bar{k}}) - \tilde{c}(1 - H(\phi_{\bar{k}}))) - \lambda \nabla \cdot \frac{\nabla \phi_k}{|\nabla \phi_k|}$$

$$+ \{\beta + \tilde{\beta} H(\phi_{\bar{k}})(c_{2,1} - c_{\bar{k}})^2\}(H(\phi_k) - S \circ T_k) + \tilde{\beta}(c_{2,1} - c_k)^2 (H(\phi_{\bar{k}}) - S \circ T_{\bar{k}})^2\} = 0,$$

$$\frac{\partial \phi_k}{\partial \vec{n}} = 0 \quad \text{on} \quad \partial\Omega, \tag{9.2}$$

$$\mu_k \int_\Omega \{\beta + \tilde{\beta} H(\phi_{\bar{k}})(c_{2,1} - c_{\bar{k}})^2\}(S \circ T_k - H(\phi_k)) \nabla S_{T_k x} \cdot R'_{\theta_k} x \, dx = 0, \tag{9.3}$$

$$\int_\Omega \{\beta + \tilde{\beta} H(\phi_{\bar{k}})(c_{2,1} - c_{\bar{k}})^2\}(S \circ T_k - H(\phi_k)) \nabla S_{T_k x} \cdot R_{\theta_k} x \, dx = 0, \tag{9.4}$$

$$\int_\Omega \{\beta + \tilde{\beta} H(\phi_{\bar{k}})(c_{2,1} - c_{\bar{k}})^2\}(S \circ T_k - H(\phi_k)) \nabla S_{T_k x} \, dx = 0, \tag{9.5}$$

where R_{θ_k} is the rotation matrix for θ_k, and R'_{θ_k} is its derivative at θ_k. Given initial values Φ_0, C_0, and T_0, we use an iterative scheme to minimize (9.1) by sequentially updating Φ, C, and T using gradient descent.

10. Experimental results. In this section, we present experimental results using (6.1) on synthetic images and real examples, including an electron microscope (EM) image of erythrocytes and other images, with multiple occluded objects. We demonstrate use of prior shape information in the form of length that penalizes longer boundary (Figure 5) and explicit shape (Figures 6–10) to handle occlusions. Once the mean intensities are obtained for occlusion regions using the minimizing constants C_{int} and C_{obj}, we show in examples (Figure 7) how the occlusion relation can be deduced for the objects. Finally in (Figures 11 and 12), we see how these relationships have been used to impose shape constraints selectively.

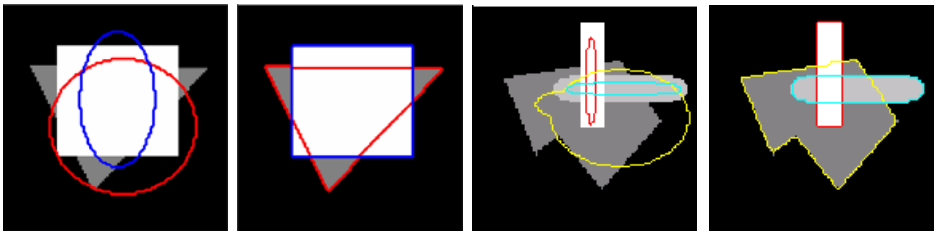

Figure 5. *Segmentation of objects with occluded linear boundaries, using* only *the length term.*

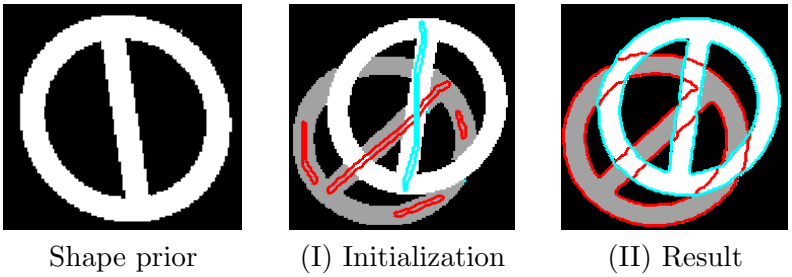

Figure 6. *Occluded boundaries for* 2 *objects are detected due to the use of explicit shape.*

10.1. Segmentation with length and explicit shape. In the images in Figure 5, we assume that only *linear segments* of objects are occluded. Setting $\beta = 0$ in (6.2), we see that the length term alone can be used to resolve segmentation in occluded regions. Starting with initial level sets $\{\phi_k\}_{k=1}^{N}$ (N$\leq$ max. number of objects expected) shown in (I), (III), we minimize (6.2) to obtain segmentation results in (II), (IV). The initial level sets have to be overlaid at least partially on the corresponding objects to avoid local minima. The constants C_{int} and C_{obj} can be arbitrarily initialized.

In the next set of examples in Figures 6 and 7, we assume that explicit shape information on the objects is available. In Figure 6, given an image with occlusions (I), in which the objects can be described by a prior shape (binary image shown on left), we minimize (8.5) using (9.2)–(9.4) to get the result in (II). The only initial values required here are the initial level sets (shown in (I)). Using this, a few iterations of the (9.3)–(9.4) are used to get an initial guess for the rigid transformation T_k. Figure 7 shows an example for the 3 objects case. Notice that use of prior shape information has resulted in a good segmentation in spite of a lack of any intensity information (e.g., Image B).

In Figure 8 (I), we see a given EM image of erythrocytes (red blood cells), which are generally known to have a circular surface structure. Naturally, we use a prior shape of a circle in this case to resolve occlusions. Starting with an initial guess for the active contours in (II), we arrive at result (IV), which is not possible without prior shape information (III).

In Figures 9 and 10, we demonstrate our segmentation algorithm on real images with multiple occluded objects as shown in (I) incorporating explicit shape priors (II) to overcome occlusions. The initial condition for the active contour to evolve is presented in (III), and the segmentation result with correct spatial order is obtained in (IV) where occluded parts of the objects are properly recovered by using explicit prior shapes.

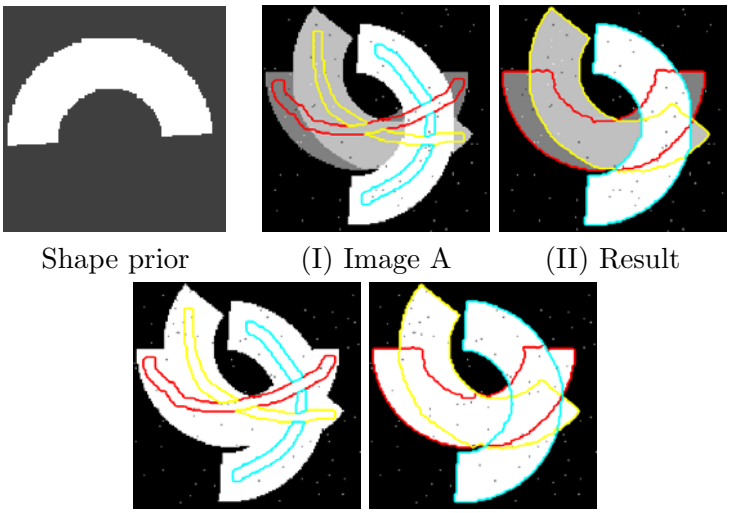

Figure 7. *Segmentation of* 3 *occluded objects with explicit shape. For Images* A *and* B, (I) *and* (III) *show the input image with the initial guess, and* (II) *and* (IV) *show the result. The correct result is obtained for Image B despite lack of contrast.*

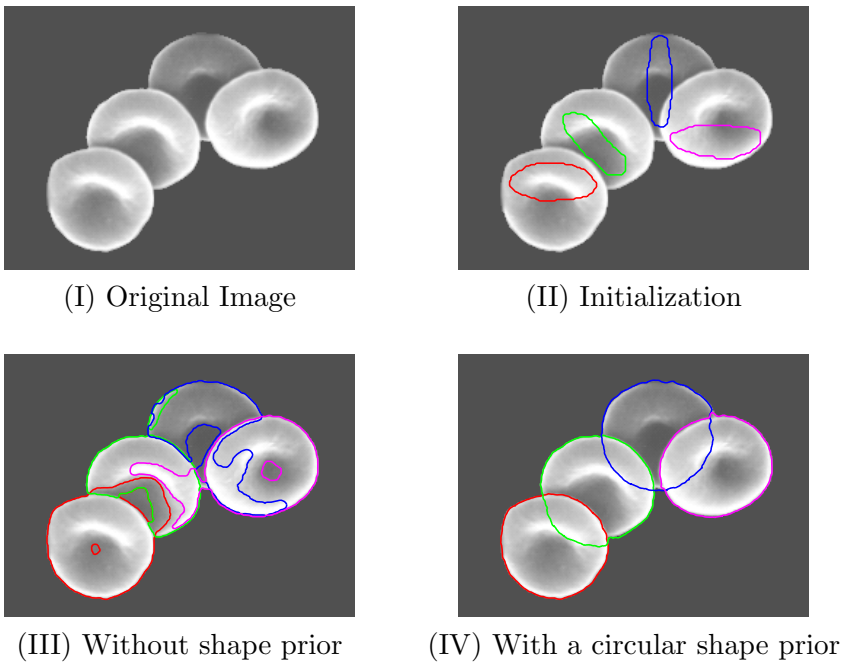

Figure 8. *Segmentation of erythrocytes in EM image.* (I) *Input image with circular shaped objects.* (II) *Initial contours that have to be placed close to the objects.* (III) *Incorrect result without use of shape due to occlusions.* (IV) *Correct result with use of a circular shape prior.*

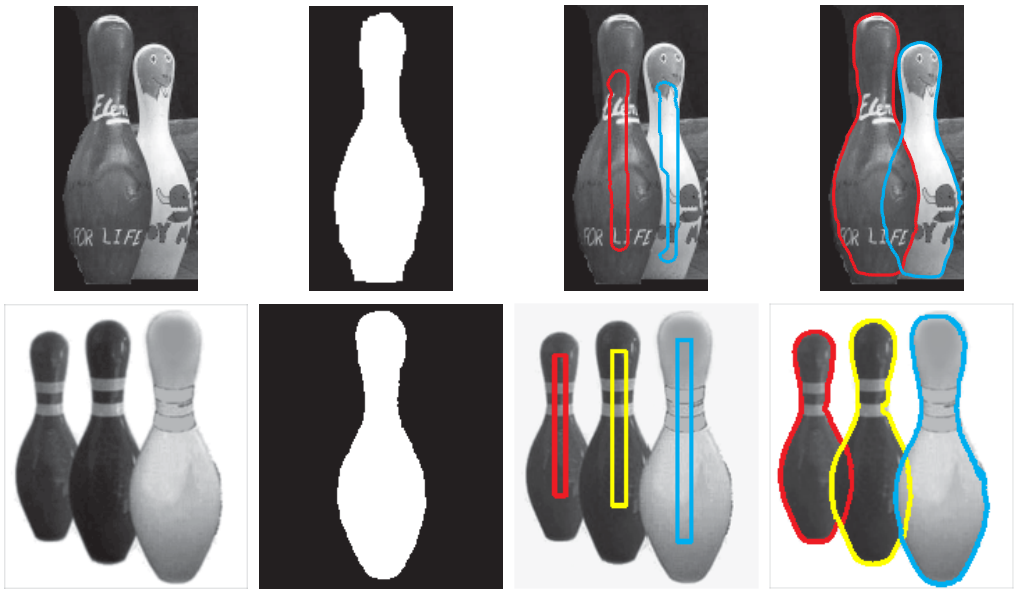

Figure 9. *Segmentation results for real images of multiple objects under occlusion.* (I) *Input image to segment.* (II) *Explicit shape prior to be used for segmentation.* (III) *Initial contours that have to be placed close to the objects* (IV). *Correct result with use of a shape prior.*

10.2. Spatial order estimation. An immediate application of computing the constants C_{int}, C_{obj} by minimizing (6.1) is to deduce the occlusion relation between the objects. This relation defines a unique object spatial order for images that satisfy the assumptions discussed in section 7. In the examples shown so far, a unique spatial order exists for Figures 5, 6, 7 (Image A), and 9, but does not exist for Figures 7 (Image B), 8, and 10. A unique spatial order cannot be determined because of lack of contrast at the intersection regions in Figures 7 (Image A) and 10 (aircraft), and due to an insufficient number of overlapping objects in Figures 8 and 10 (soccer ball).

We demonstrate the procedure described in section 7 for Figure 6 (Image A). The segmented object regions are given by the characteristic functions $\mathrm{H}(\phi_\mathrm{k}) := \mathrm{H_k}$. In Figure 6(II), $\mathrm{H_1}$, $\mathrm{H_2}$, and $\mathrm{H_3}$ are the interiors of the red, yellow, and cyan curves, respectively. Let $\bar{\mathrm{H}}_\mathrm{k} := 1 - \mathrm{H_k}$. Now the object regions $\mathrm{H_k}$ define 4 occlusion regions,

$$\mathrm{P_1} = \mathrm{H_1H_2}\bar{\mathrm{H}}_3, \quad \mathrm{P_2} = \mathrm{H_1}\bar{\mathrm{H}}_2\mathrm{H_3}, \quad \mathrm{P_3} = \bar{\mathrm{H}}_1\mathrm{H_2H_3}, \quad \mathrm{P_4} = \mathrm{H_1H_2H_3}.$$

The computed constants C_{obj}, C_{int} are

$$c_1 = 128.9,\ c_2 = 191.9,\ c_3 = 254.3,\ c_{2,1} = 129,\ c_{2,2} = 129.1,\ c_{2,3} = 192,\ c_{3,1} = 129.1.$$

Table 2 shows a comparison for these intensities in different occlusion regions. The third column shows the mean intensities μ_k corresponding to the regions P_k, which can be computed from C_{obj}, C_{int} (as shown in the second column).

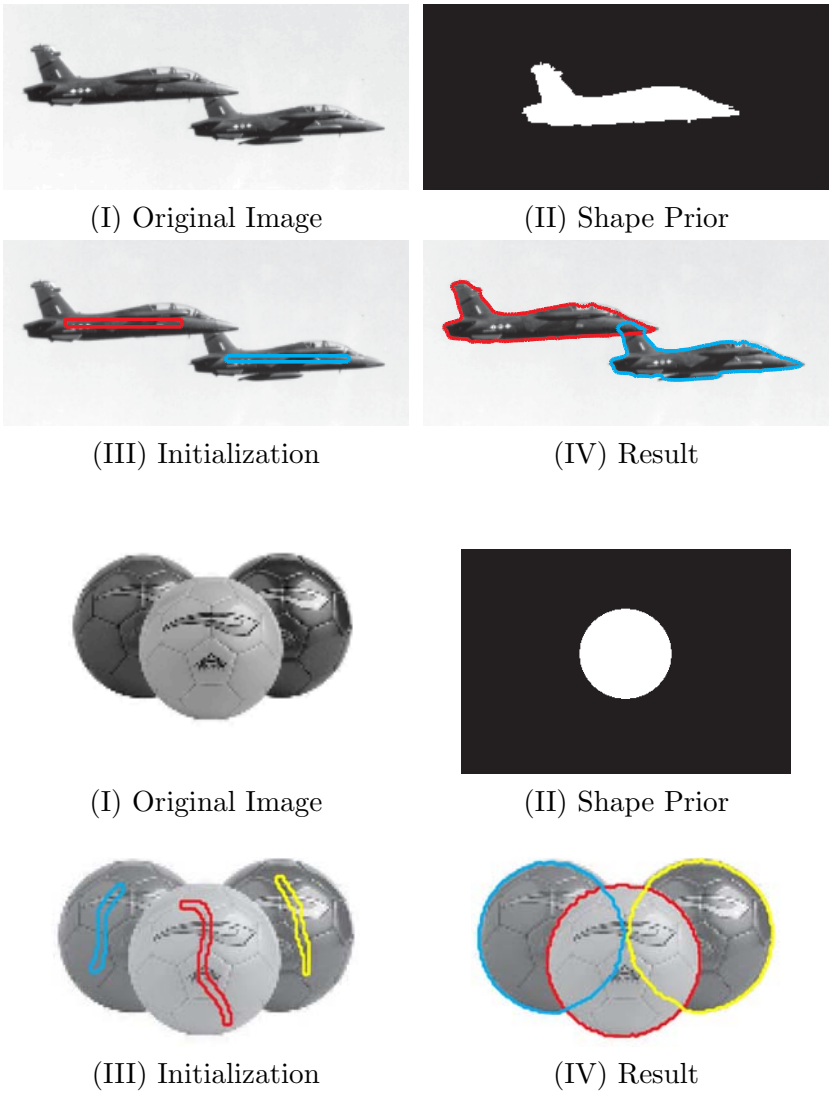

Figure 10. *Segmentation results for real images of multiple objects under occlusion.* (I) *Input image to segment.* (II) *Explicit shape prior to be used for segmentation.* (III) *Initial contours that have to be placed close to the objects.* (IV) *Correct result with use of a shape prior.*

When we compare the mean intensities μ_k with C_{obj}, we can infer the occlusion relation as

$$|\mu_1 - c_2| < |\mu_1 - c_1| \text{ gives } 2 > 1, \text{ and}$$
$$|\mu_3 - c_3| < |\mu_3 - c_2| \text{ gives } 3 > 2,$$

which due to the transitivity assumption gives the unique spatial ordering of the objects as $3 > 2 > 1$, i.e., O_3, O_2, O_1 (with increasing depth).

Table 2
Comparison of mean intensities in different regions.

Occlusion region	Mean-intensity (μ_k)	Values
P_1	$c_1 + c_2 - c_{2,1}$	191.8
P_2	$c_1 - c_{2,2} + c_3$	254.1
P_3	$-c_{2,3} + c_2 + c_3$	254.1
P_4	$c_{3,1} + c_{2,1} + c_{2,2} + c_{2,3} - c_1 - c_2 - c_3$	254

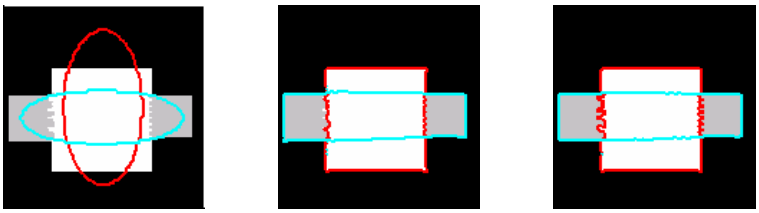

Figure 11. *Selective use of length only for occluded boundaries.* (I) *Initialization.* (II) *Without selective use of length, the sharp features in the white object are lost.* (III) *With a selective length term, the features of the white object are recovered.*

10.3. Selective use of shape. Finally, we will see examples where we have used the occlusion relation to impose shape constraints selectively. In Figure 11, we see an image of two objects, say W (white) and G (gray), with W occluding G. In addition, the object in front, W, has sharp features in the intersection region, which we want the segmentation to preserve. Assuming occluded boundaries to be linear, only the length term is used to resolve occlusions. In (II), we minimized (6.2) with $\beta = 0$. The resulting segmentation has correctly filled in the missing linear boundaries for G, but has not segmented W correctly, since the use of the length term evenly for both objects has resulted in the loss of sharp features in W. When we use a selective length term as in (8.5) with $\hat{S}(\phi_k) = |\nabla H(\phi_k)|$ and setting the parameters $\tilde{\beta} = 0, \lambda \ll \beta$, the required boundaries are computed as needed (III). Notice that the length term is effected only for the object that is occluded G, hence preserving the features of W.

A similar example is presented in Figure 12, with use of explicit shape. Here (I) shows two *square* shaped objects W and G, with W occluding G, and each with one of its corners chipped off. We want our segmentation to be able to complete the missing boundary for G in the occlusion region, and also to be able to preserve edges that are not occluded. (II) shows the result when a shape term is not used; the length term completes the occluded part of G linearly. (III) shows the result with a uniform shape term as in (8.5) using a *square* as the shape prior. Notice that the corners of both W and G are not segmented properly, due to the influence of the shape term even to nonoccluded regions. Finally, we get the correct segmentation in (IV), using a selective shape term as in (9.1). First, use of the shape term has filled the missing boundary of G that has been occluded. Second, the shape term is applied only to the occluded object G. Hence the corner of W is recovered. Third, the shape term applied to G is affected only within the object W, thus localizing the effect of shape on G.

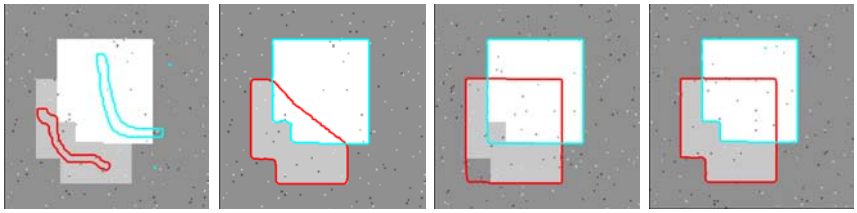

Figure 12. *Selective use of shape.* (I) *Initialization.* (II) *Without shape. Only the length term is active.* (III) *Without selective shape, the chipped corners are not segmented due the influence from the shape term.* (IV) *With selective shape, the corners are detected, and the occluded boundary of the gray object is also recovered.*

Hence the corner of G is also recovered.

11. Comparisons. In section 5, we showed that the NMS minimization (2.4)

$$\min_{k\in\{1,2,\ldots,N!\}} \min_{\mathbf{A},C_{obj}} \tilde{E}_{Q^k} \tag{11.1}$$

is equivalent to our constrained minimization (4.1)

$$\min_{\mathbf{A},C_{obj},C_{int}\in\{S^1,S^2,\ldots,S^{N!}\}} E. \tag{11.2}$$

In the previous sections, we dealt with the implementation details and experiments for the unconstrained minimization (4.2)

$$\min_{\mathbf{A},C_{obj},C_{int}} E. \tag{11.3}$$

Now, we will provide a detailed comparison of computational complexities of (2.4) with (4.1). Then for a typical example with varying noise levels, we compare the solutions obtained by the energy (4.1) with its unconstrained version (4.2) that we use for our experiments.

Here, for the implementation of (2.4) and (4.1), we used an iterative gradient descent scheme as shown for (4.2). To impose the constraint on C_{int} for (4.1), during each iteration n, we project the sequence C_{int}^n onto $\mathbf{S}$ when such a projection is uniquely defined. To avoid local minima issues, if such a projection is not uniquely defined, we simply retain the value of C_{int}^n for the next iteration.

11.1. Complexity comparisons of (4.1) with NMS energies. The key difference between the NMS energy and our formulation is the estimation of the spatial order of the objects. In our model, first we note that there is a one-one correspondence between the spatial orders $\mathbf{Q} = \{Q^1, Q^2, \ldots, Q^{N!}\}$ and the sequences $\mathbf{S} = \{S^1, S^2, \ldots, S^{N!}\}$. Thus in our constrained formulation (4.1), the spatial order is given by the estimated value of C_{int}. We show an example to illustrate the spatial order estimation in our case. In Figure 13, (I) is an image with 3 objects with a "crescent" shape. The initial contours are shown overlaid on the objects. (II) shows the segmented contours using the gradient descent of (4.1) which converged in about 1000 iterations.

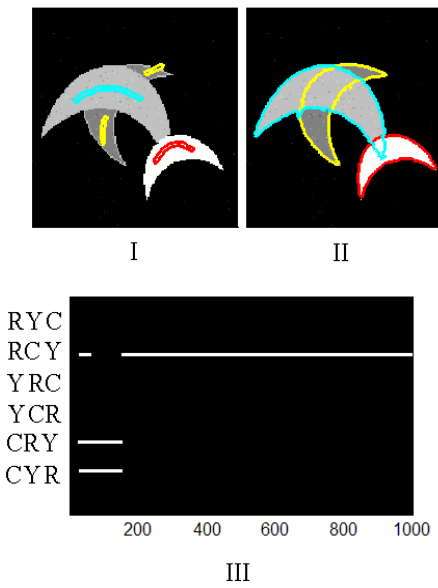

Figure 13. *Spatial order estimation through variable C_{int}.* (I) *Image with initial guess.* (II) *Segmented contours.* (III) *Plot of possible spatial orders (y-axis) against iteration (x-axis) (e.g., between iterations* $100-200$*, CRY and CYR are the possible spatial orders).*

In graph (III) (*iteration versus spatial order*), we see the spatial order corresponding to the estimated C_{int}^n. The spatial order (y-axis) between the objects (labelled by colors (red (R), yellow (Y), cyan (C))) is indicated by the letter order (e.g., RYC denotes the order R $>$ Y $>$ C). For the first few iterations when the evolving object regions have not yet intersected, C_{int}^n is unknown. Then, until about 200 iterations, the projection of C_{int}^n to the spatial order set $\{S^1, S^2, \ldots, S^{N!}\}$ is not unique. Finally after 200 iterations, when a sufficient number of intersection regions between the evolving objects are nonempty, the projection of C_{int}^n to $\mathbf{S}$ is unique and hence defines the spatial order.

In the NMS approach (2.4), there are $N!$ energies $\tilde{E}_{Q^k}(C_{obj}, \mathbf{A}), k = 1, 2, \ldots, N!$, that one minimizes. The spatial order is then given by the smallest of the above energy-minima. Even in the case of a good initial guess for the object boundaries, one would have to go through the $N!$ minimizations. Thus, even though the search space is the same for both NMS and our energies, unlike the NMS approach, our model determines the spatial order by gradient descent of C_{int}. Since the spatial order is integrated within a *single energy*, our model involves N level set minimizations in contrast to $N.N!$ level set minimizations in the NMS model. Below we compute the computational complexity of one gradient descent iteration of our energy (4.1) with that of the NMS model (2.4). Here M^2 is the grid size.

- Energy (4.1): In our case, there are 3 steps every iteration: solving for the 2^N constants C_{obj} and C_{int}, evolving N level set functions, and projecting the minimizer C_{int} onto the spatial order set $\mathbf{S}$.

In (6.1), computing products involving $H(\phi_1), H(\phi_2), \ldots, H(\phi_N)$, i.e., $S_{p,k}$, costs $O(N2^N M^2)$. The linear system for C_{obj}, C_{int} involves integrals of $S_{p,k}$, costing $O(2^N M^2)$. Furthermore, the above linear system is triangular, costing $O(2^{2N})$ to solve. Also, each level set evolution is a sum of 2^N terms involving $S_{p,k}$, thus costing $O(2^N M^2)$. Finally as discussed in section 7, the cost of projecting the minimizer C_{int} onto the spatial order set $\mathbf{S}$ is $O(2^{2N})$. Hence the total cost for one gradient descent iteration of (4.1) is $O(N2^N M^2 + 2^{2N})$.

- Energy (4.2): For the unconstrained energy (4.2), there is no projection step for C_{int} during gradient descent. The minimizer C_{int}^{min} is projected onto $\mathbf{S}$ after minimization to infer the spatial order. Thus the complexity is the same as that for (4.1), $O(N2^N M^2 + 2^{2N})$.
- NMS: Similarly to the above computations, it is easy to show that for each NMS energy $\tilde{E}_{Q^k}(C_{obj}, \mathbf{A}), k = 1, 2, \ldots, N!$, solving for the constants C_{obj} costs $O(N^2 M^2)$. Furthermore, one level set evolution would involve N terms and hence costs $O(NM^2)$. Thus the total cost for $N!$ energies is $O(N^2 N! M^2)$.

To summarize, in the NMS case, spatial order is determined by minimization of $N!$ energies. In our case spatial order is determined by gradient descent of C_{int}. Second, in the NMS case there are $N.N!$ level set evolutions that one has to go through, each costing $O(NM^2)$ because of N terms in the energy. In our case, there are only N level set evolutions, but each with a higher cost $O(2^N M^2)$ because of additional terms in the energy that model intensity in intersection regions. There is also the cost $O(2^{2N})$ for computing the constants C_{int} and projecting it onto the spatial order set $\mathbf{S}$. One way to reduce the additional cost per level set evolution in our energy (4.1) is to use the current spatial order estimate (whenever available) to reduce the number of terms in the level set evolution for subsequent iterations.

Note. Our experiments further indicate that the energy convergence rate of gradient descent is the *same* for (4.1), and for *one* of the $N!$ NMS energies (i.e., the energy corresponding to the object spatial order).

11.2. Comparison of (4.1) and its unconstrained version (4.2). For section 10, we used the energy (4.2), $E[\mathbf{A}, C_{obj}, C_{int}]$, which is the unconstrained version of (4.1), i.e., $E[\mathbf{A}, C_{obj}, C_{int} \in \mathbf{S}]$. For images corrupted with white noise, our experiments indicate that the minimizers computed using the above energies are very close. This is not true in general for nonwhite (*spectral*) noise images, since (4.2) due to the additional constants C_{int} intrinsically allows the intensities in the intersection regions to be different from the object intensities. Thus the resulting segmentation of occluded objects could be influenced by edges introduced by noise. For instance, in Figure 14 (I) shows an image with spectral noise (in the light gray region) that corrupts the intersection of the 2 square objects (white and dark gray regions). (II) shows the incorrect segmentation obtained by the energy (4.2). To make the intersection region close to constant intensity, the image term competes with the shape term and has influenced the shape of the cyan curve (occluded object). Thus one would have to tune the shape parameter β in (4.2) to get to a minimizer close to that of (4.1) (shown in (III)). Whereas the result computed using (4.1) is robust to such types of noise, because once the spatial order is determined (i.e., projection of C_{int} onto $\mathbf{S}$ is unique), occluded boundaries are driven *only* by the shape term. Hence tuning of β is not required in this case.

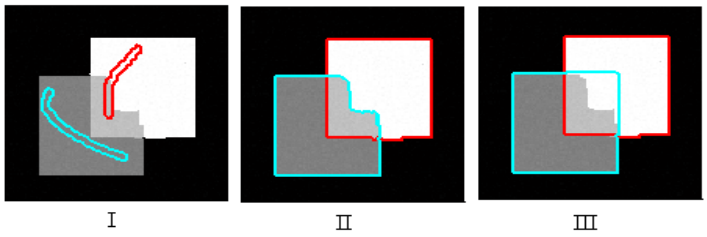

Figure 14. *Solutions computed using* (4.1) *and* (6.1) *for spectral noise images.* (I) *Image of* 2 *objects (white and dark gray) with noise induced edges (due to the light gray region).* (II) *Incorrect result using* (4.1) *(segmentation of cyan curve influenced by image term).* (III) *Correct result using* (4.1) *(segmentation of cyan curve influenced only by shape term).*

Now we illustrate 2 examples in Figure 15 for $N = 2, 3$ to compare solutions of the above models for varying white-noise levels. The results for both models were obtained for the same values of the shape parameter β chosen nominally without any careful tuning. In both examples, we see *identical* segmentation results using (4.2) (row A) and (4.1) (row B) for images corrupted with increasing levels of white noise. For $N = 2, 3$, the plot shown is the relative-error ξ (%) (y-axis) with respect to noise level (x-axis). Here $\xi = \text{mean}(\frac{|A_i - B_i|}{|B_i|}, i = 1, 2, \ldots, N)$, where A_i are the segmented regions using (4.2) and B_i are the corresponding segmented regions using (4.1). Likewise, the errors obtained for the object intensities were very low, and the estimated spatial orders were identical for both models. The low errors observed above indicate that for white-noise images, the minimizers are not dependent on the constraint on C_{int}.

12. Discussion. We have presented a variational energy based framework (4.1) to segment multiple occluded objects using prior shape information of the objects. We note that the spatial order between the objects is encoded within the constants C_{int} which we estimate as part of the minimization process. Our model as we demonstrated is computationally less expensive than the NMS approach (2.4) which formulates segmentation energies for each spatial order. Second, we showed how the estimated constants C_{int} and C_{obj} can be used to impose shape priors selectively, just to occluded boundaries.

Also in our experiments, we have approximated the energy (4.1) by its unconstrained version (4.2), and the results for the above energies are observed to be identical for white noise images. As future work, we want to prove the above observation and provide error estimates for spectral noise images. Second, in our experiments the initial contours have to be placed near the objects to avoid local minima. This might not be practicable for applications with a large number of objects. Hence another direction is to automate the contour initialization procedure to extend the EM (I) example (Figure 8) to applications such as automatic counting of blood cells. Finally, we have used one level set per object irrespective of the occlusion scene. A better way would be to use an optimal number of level set functions depending on the occlusion scene. For instance, when N objects occlude each other in all possible regions, we definitely need N level sets, whereas when no occlusions occur, one can work with $\log_2 N$ level sets (e.g., multiphase CV).

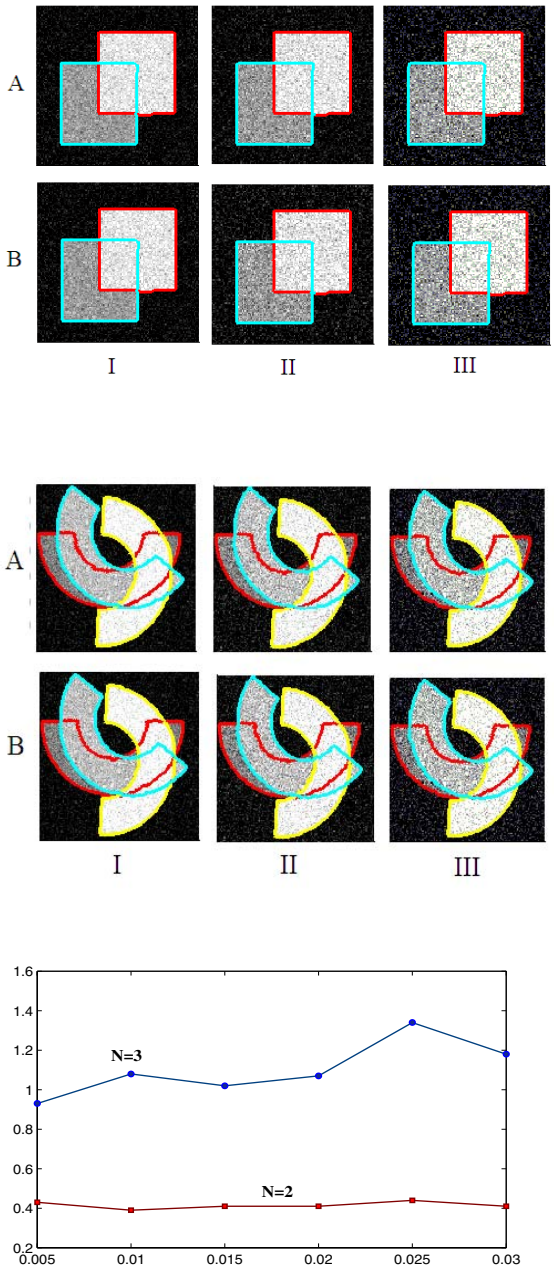

Figure 15. *Comparison of solutions computed using* (4.2) *(row A) and* (4.1) *(row B) for noise levels* (I) 0.01, (II) 0.02, (III) 0.03. *The plots for* $N = 2, 3$ *show relative-error* ξ *(%) (y-axis) with respect to noise level (x-axis).*

Our work is based on some assumptions which limit the application of our model to many real life images. First, the object intensity is assumed to be close to a constant. For many real images that contain objects with nonconstant intensity and textures, one could modify the energy (4.1) by replacing the constants C_{obj}, C_{int} with components to handle smooth intensity variations along with textures. In this case, the spatial order estimation becomes challenging since it would now be nontrivial to compare the intensity in the intersection regions with that of the objects. Another possibility in a related direction would be to account for transparencies of occluding objects within the model. Second, we assume that the objects do not twist between themselves. To allow for twisting between objects would require a new representation of the intersection region intensities and their relationship with the object intensities, since the intersection intensities are now no longer close to a constant. Additionally, one would also have to estimate spatial orders locally around each intersection. Finally, we assume that our prior shape is a binary image differing from the objects by a rigid transformation. Thus the shape term would penalize even small local deformations from the given shape. One possibility would be to extend the search space of T_k in (9.1) to transformations that have a global affine component along with a local nonrigid component that allows for local shape variations. Alternatively we can introduce richer shape prior knowledge (when available) using statistical techniques such as PCA or its variations instead of using a single binary shape, similar to the work of Leventon, Grimson, and Faugeras [27].

REFERENCES

[1] L. Bergen and F. Meyer, *A novel approach to depth ordering in monocular image sequences*, in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, 2000, pp. 536–541.

[2] G. Bertrand and G. Malandain, *A new topological segmentation of discrete surfaces*, in Proceedings of the Second European Conference on Computer Vision, 1992, pp. 710–714.

[3] A. Blake and A. Zisserman, *Visual Reconstruction*, MIT Press, Cambridge, MA, 1987.

[4] V. Caselles, F. Catte, T. Coll, and F. Dibos, *A geometric model for active contours*, Numer. Math., 66 (1993), pp. 1–31.

[5] V. Caselles, R. Kimmel, and G. Sapiro, *Geodesic active contours*, Int. J. Comput. Vision, 22 (1997), pp. 61–79.

[6] T. Chan and W. Zhu, *Level set based shape prior segmentation*, in Proceedings of the IEEE Computer Science Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, pp. 1164–1170.

[7] T. F. Chan and L. A Vese, *Active contours without edges*, IEEE Trans. Image Process., 10 (2001), pp. 266–277.

[8] Y. Chen, S. Thiruvenkadam, H. D. Tagare, F. Huang, D. Wilson, and E. A. Geiser, *On the incorporation of shape priors into geometric active contours*, in IEEE Workshop on Variational and Level Set Methods, 2001, pp. 145–152.

[9] Y. Chen, H. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. S. Gopinath, R. Briggs, and E. A. Geiser, *Using prior shapes in geometric active contours in a variational framework*, Int. J. Comput. Vision, 50 (2002), pp. 315–328.

[10] D. L. Chopp, *Computing minimal surfaces via level set curvature flow*, J. Comput. Phys., 106 (1993), pp. 77–91.

[11] L. D. Cohen, *On active contour models and balloons*, CVGIP: Image Understanding, 53 (1991), pp. 211–218.

[12] T. Cootes, C. Taylor, D. Cooper, and J. Graham, *Active shape models—Their training and application*, CVIU, 61 (1995), pp. 38–59.

[13] D. Cremers, T. Kohlberger, and C. Schnörr, *Shape statistics in kernel space for variational image segmentation*, Pattern Recognition, 36 (2003), pp. 1929–1943.
[14] H. Delingette and J. Montagnat, *New algorithms for controlling active contours shape and topology*, in Proceedings of the 6th European Conference on Computer Vision, 2000, pp. 381–395.
[15] A. Dervieux and F. Thomasset, *A finite element method for the simulation of a Rayleigh-Taylor instability*, in Approximation Methods for Navier–Stokes Problems (Paderborn, 1979), Lecture Notes in Math. 771, Springer-Verlag, Berlin, 1980, pp. 145–158.
[16] A. Dervieux and F. Thomasset, *Multifluid incompressible flows by a finite element method*, in Seventh International Conference on Numerical Methods in Fluid Dynamics, W. Reynolds and R. W. MacCormack, eds., Lecture Notes in Phys. 141, Springer-Verlag, Berlin, 1980, pp. 158–163.
[17] S. Esedoglu and R. March, *Segmentation with depth but without detecting junctions*, J. Math. Imaging Vision, 18 (2003), pp. 7–15.
[18] L. H. Finkel and P. Sajda, *Object discrimination based on depth-from-occlusion*, Neural Comput., 4 (1992), pp. 901–921.
[19] A. Foulonneau, P. Charbonnier, and F. Heitz, *Affine-invariant geometric shape priors for region-based active contours*, IEEE Trans. Pattern Anal. Mach. Intell., 28 (2006), pp. 1352–1357.
[20] P. Fua and Y. G. Leclerc, *Model driven edge detection*, Machine Vision Appl., 3 (1990), pp. 45–56.
[21] R.-X. Gao, T.-F. Wu, S.-C. Zhu, and N. Sang, *Bayesian inference for layer representation with mixed Markov random field*, in Energy Minimization Methods in Computer Vision and Pattern Recognition, Lecture Notes in Comput. Sci. 2679, Springer-Verlag, Berlin, 2007, pp. 213–224.
[22] U. Grenander, Y. Chow, and D. Keenan, *HANDS: A Pattern Theoretic Study of Biological Shapes*, Springer-Verlag, 1990.
[23] M. Kass, A. Witkin, and D. Terzopoulos, *Snakes: Active contour models*, Int. J. Comput. Vision, 1 (1988), pp. 321–331.
[24] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, *Gradient flows and geometric active contour models*, in Proceedings of the Fifth International Conference on Computer Vision, IEEE Computer Society, Washington, DC, 1995, pp. 810–815.
[25] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, *Conformal curvature flows: From phase transitions to active vision*, Arch. Rational Mech. Anal., 134 (1996), pp. 275–301.
[26] F. Lecellier, S. Jehan-Besson, J. Fadili, G. Aubert, M. Revenu, and E. Saloux, *Region-based active contour with noise and shape priors*, in Proceedings of the IEEE International Conference on Image Processing, 2006, pp. 1649–1652.
[27] M. Leventon, W. L. Grimson, and O. Faugeras, *Statistical shape influence in geodesic active contours*, in Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, Vol. 1, 2000, pp. 316–323.
[28] R. Malladi, J. A. Sethian, and B. C. Vemuri, *A topology-independent shape modeling scheme*, in Proceedings of the SPIE Conference on Geometric Methods in Computer Vision II, Vol. 2031, 1993, pp. 246–258.
[29] K. V. Mardia, W. Qian, D. Shah, and K. M. A. de Souza, *Deformable template recognition of multiple occluded objects*, IEEE Trans. Pattern Anal. Mach. Intell., 19 (1997), pp. 1035–1042.
[30] S. Masnou, *Disocclusion: A variational approach using level lines*, IEEE Trans. Image Process., 11 (2002), pp. 68–76.
[31] T. McInerney and D. Terzopoulos, *Topologically adaptable snakes*, in Proceedings of the Fifth IEEE International Conference on Computer Vision, Los Alamitos, CA, 1995, pp. 840–845.
[32] D. Mumford and J. Shah, *Boundary detection by minimizing functionals*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1985, pp. 22–26.
[33] D. Mumford and J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math., 42 (1989), pp. 577–684.
[34] M. Nitzberg and D. Mumford, *The* 2.1-*D sketch*, in Proceedings of the Third IEEE International Conference on Computer Vision, 1990, pp. 138–144.
[35] M. Nitzberg, D. Mumford, and T. Shiota, *Filtering, Segmentation and Depth*, Lecture Notes in Comput. Sci. 662, Springer-Verlag, Berlin, 1993.
[36] S. Osher and J. Sethian, *Fronts propagating with curvature dependent speed: Algorithms based on the Hamilton–Jacobi formulation*, J. Comp. Phys., 79 (1988), pp. 12–49.

[37] N. Paragios and R. Deriche, *Coupled geodesic active regions for image segmentation: A level set approach*, in Proceedings of the 6th European Conference on Computer Vision, Part II, 2000, pp. 224–240.

[38] T. Riklin-Raviv, N. Kiryati, and N. Sochen, *Unlevel-sets: Geometry and prior-based segmentation*, in Computer Vision—ECCV'04, Lecture Notes in Comput. Sci. 3024, Springer-Verlag, Berlin, 2004, pp. 50–61.

[39] M. Rousson and N. Paragios, *Shape priors for level set representations*, in Computer Vision—ECCV'02, Lecture Notes in Comput. Sci. 2351, Springer-Verlag, Berlin, 2002, pp. 78–92.

[40] S. R. Thiruvenkadam, T. F. Chan, and B. W. Hong, *Segmentation under occlusions using selective shape prior*, in Scale Space and Variational Methods in Computer Vision, Springer-Verlag, Berlin, 2007, pp. 191–202.

[41] A. Tsai, A. Yezzi, Jr., W. Wells, III, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, *Model-based curve evolution technique for image segmentation*, in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, 2001, pp. 463–468.

[42] A. Tsai, A. Yezzi, Jr., W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, *A shape-based approach to the segmentation of medical imagery using level sets*, IEEE Trans. Med. Imaging, 22 (2003), pp. 137–154.

[43] A. Tsai, A. Yezzi, and A. S. Willsky, *Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification*, IEEE Trans. Image Process., 10 (2001), pp. 1169–1186.

[44] L. Vese and T. Chan, *Multiphase level set framework for image segmentation using the Mumford and Shah model*, Int. J. Comput. Vision, 50 (2002), pp. 271–293.

[45] A. Yezzi, A. Tsai, and A. Willsky, *A statistical approach to snakes for bimodal and trimodal imagery*, in Proceedings of the IEEE International Conference on Computer Vision, Vol. 2, 1999, pp. 898–903.

[46] W. Zhu, T. Chan, and S. Esedoglu, *Segmentation with depth: A level set approach*, SIAM J. Sci. Comput., 28 (2006), pp. 1957–1973.