Uncertainty Driven Multi-Scale Optimization

Pushmeet Kohli[1] Victor Lempitsky[2] Carsten Rother[1]

[1]Microsoft Research Cambridge [2] University of Oxford

Abstract. This paper proposes a new multi-scale energy minimization algorithm which can be used to efficiently solve large scale labelling problems in computer vision. The basic modus operandi of any multi-scale method involves the construction of a smaller problem which can be solved efficiently. The solution of this problem is used to obtain a partial labelling of the original energy function, which in turn allows us to minimize it by solving its (much smaller) projection. We propose the use of new techniques for both the construction of the smaller problem, and the extraction of a partial solution. Experiments on image segmentation show that our techniques give solutions with low pixel labelling error and in the same or less amount of computation time, compared to traditional multi-scale techniques.

1 Introduction

Energy minimization and discrete optimization have become a cornerstone of computer vision. This has primarily been driven by their ability to efficiently compute the Maximum a Posteriori (MAP) solutions in models such as Markov and Conditional random fields (MRFs and CRFs), e.g. [1,3].

In recent years, advances in image acquisition technologies have significantly increased the size of images and 3D volumes. For instance, the latest commercially available cameras can capture images with almost 20 million pixels. In fact it is now possible to capture giga-pixel images of complete cities [7]. Similarly, latest medical imaging systems can acquire 3D volumes with billions of voxels. This type of data gives rise to large scale optimization problems which are very computationally expensive to solve and require large amounts of memory.

Multi-scale processing has long been a popular approach to reduce the memory and computational requirements of optimization algorithms (see [11,4,10] for a review). The basic structure of these methods is quite simple. In order to label a large image (or 3D volume) they first solve the problem at low resolution, obtaining a coarse labelling of the original high resolution problem. This labelling is refined by solving another optimization on a small subset of the pixels. A classic example of such a multi-scale method is the boundary band algorithm [9] for segmenting large images and 3D volumes. Given a solution on a coarse scale (fig. 1c), a partial solution (narrow band around segmentation) is extracted (fig. 1d) which is optimized again on high resolution (fig.1b). This algorithm suffers form the problem that it cannot efficiently recover from large errors present in the coarse labelling. For instance, if a thin foreground structure is missed in the coarse labelling, a large band of pixels will need to be analyzed at the fine scale. This would make the size of the resulting higher resolution problem large and reduce the computational benefits. An interesting method to resolve this problem was proposed by Sinop and Grady [13]. Motivated by the problem of segmenting thin structured objects they used the information from a Laplacian pyramid to

isolate pixels which might not have attained their correct labelling at the low resolution image. Recently, Lempitsky and Boykov [8] presented an interesting *touch-expand* algorithm that is able to minimize pseudo-boolean energy functions using a narrow band, while retaining the global optimality guarantees. On the downside, it has no bounds on the size of the band it may need to consider, and in the worst case the band can progressively grow to encompass the whole image. While for highly structured unary terms, concerned with the shape fitting task considered in [8], the bands are reasonably small, they are highly likely not to be so for the less structured unary terms in e.g. segmentation problems.

Our Contributions. The goal of this paper is to develop a multi-scale algorithm which can be used to minimize energy functions with a large number of variables. To do this, we need to answer the following questions: **(1)** How to construct the energy for the small scale problem? **(2)** After minimizing this energy, how do we then isolate variables which need to be solved at the finer resolution? As we will explain later in the paper, the answers to these two important questions are not independent. We will now provide a brief overview of our strategy. For ease of explanation, we will use the two-label interactive image segmentation problem as an example. However, our method is general and can be used for any labelling problem such as 3D reconstruction, stereo, object segmentation and optical flow.
Constructing the Low Resolution Energy. Ideally, we would want to construct the energy function in such a manner that its optimal solution, when projected to the full grid, matches the optimal solution of the original energy as closely as possible. Recent band-based methods for image segmentation such as [9] and [13] construct the small scale energy from a low-resolution version of the image to be segmented. In contrast, our approach constructs the small scale energy directly from the energy of the full resolution image. Experiments show that this strategy results in substantial improvements in running time and accuracy.
Uncertainty Driven Bands. The band-based multi-scale segmentation methods use the MAP solution of the small scale problem to isolate which pixels need to be solved at the fine scale. They ignore the confidence or uncertainty associated with the MAP solution. Intuitively, if a variable has low confidence in the MAP label assignment, the labels for its corresponding variables at the fine grid should be inferred by minimizing the original energy. Our method computes uncertainty estimates (fig. 1e) and uses them to choose which regions (fig. 1f) of the image should be included in the optimization at the finer level. Experimental results show that this technique enables us to identify thin structures of the object which had been misclassified in the solution of the coarse energy.

2 Multi-scale Energy Minimization

Most of the our ideas can be applied to any pairwise MRF and CRF. Since we use interactive image segmentation as the example application, we introduce the energy model only for this application. The energy $E : \mathcal{L}^n \rightarrow \mathbb{R}$ can be written as a sum of unary and pairwise functions:

$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} \phi_i(x_i) + \sum_{(i,j) \in \mathcal{E}} \phi_{ij}(x_i, x_j). \quad (1)$$

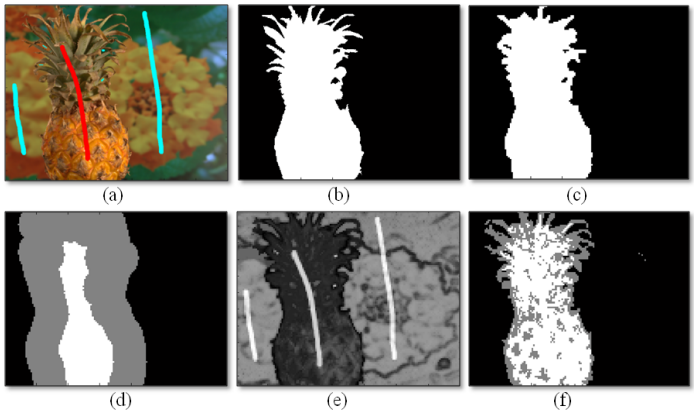

Fig. 1. Uncertainty driven multi-scale image segmentation. (a) Image with user marked brush strokes for different segments. (b) Segmentation obtained by minimizing a conventional segmentation energy (1) defined over the image grid. (c) Segmentation obtained by minimizing an energy defined over a coarse level gird which is constructed using our method (see section 3). Observe that many pixels take labels different from the MAP labels shown in (b). To correct such errors we need to mark such pixels as unlabelled, and find their labels at the fine scale. (d) Partial segmentation obtained by marking pixels in the band around the segmentation boundary of (c) as unlabelled (marked gray). The size of the band is chosen to include all incorrectly labelled pixels. (e) Min-marginal based confidence values for pixels taking the MAP label (bright pixels are more confident) — see section 4 for more details. (f) Partial labelling obtained by marking pixels below a confidence score as unlabelled. As for (d) the confidence threshold is chosen to include all incorrect pixels. For this example, with the uncertainty based scheme 3 times less number of pixels need to be marked as unlabelled compared to the number marked with the boundary band. As less number of variable need to be solved at the fine resolution, we get a larger speed-up.

The set $\mathcal{V}$ corresponds to the set of all image pixels, $\mathcal{E}$ is set of all edges between pixels in a 4 or 8 neighborhood. The random variable X_i denotes the labelling of pixel i of the image. The label set $\mathcal{L}$ consists of two labels: foreground (fg) and background (bg). Every possible assignment of the random variables $\mathbf{x}$ defines a segmentation. The unary potential is given as $\phi(x_i = \mathcal{S}) = -\log \Pr(I_i|x_i = \mathcal{S})$ where $\mathcal{S} \in \{\text{fg,bg}\}$, and initialized using a standard GMM model [12]. The pairwise terms ϕ_{ij} of the CRF take the form of a contrast sensitive Ising model, i.e. $\phi(x_i, x_j) = g(i,j)[x_i \neq x_j]$, where $[arg]$ is 1 if "arg" is true and 0 otherwise. The function $g(i,j)$ is an edge feature based on the difference in colors of neighboring pixels [1]. It is typically defined as:

$$g(i,j) = \theta_p + \theta_v \exp(-\theta_\beta ||I_i - I_j||^2), \tag{2}$$

where I_i and I_j are the colour vectors of pixel i and j respectively. The energy (1) is submodular, hence the global optimum can be computed efficiently with min-cut/maxflow, also known as graph cut, [1].

We now provide an overview of multi-scale methods for energy minimization. These algorithms have the following basic steps:

Construction the smaller problem. A new energy function $E^l : \mathcal{L}^{n\omega_s} \rightarrow \mathbb{R}$ is constructed over a smaller grid $(\mathcal{V}^l, \mathcal{E}^l)$ where $\mathcal{V}^l$ denotes the set of lattice points, and $\mathcal{E}^l$ denotes the corresponding edge set. This grid has $|\mathcal{V}^l| = n\omega_s$ variables (original energy E had n variables), where ω_s is the scaling parameter $(0 \leq \omega_s \leq 1)$. Let $\mathbf{X}^l = \{X_i^l, i \in \mathcal{V}^l\}$ denote the vector of variables defined on $\mathcal{V}^l$. We will denote their labelling by $\mathbf{x}^l = \{x_1^l, x_2^l, ..., x_{n\omega_s}^l\}$.

Computation of a partial labelling. The coarse energy E^l is minimized to extract a partial labelling $\mathbf{x}^*$ for the original random variables $\mathbf{X}$. Formally, each variable X_i is assigned one label from the extended label set $\mathcal{L} \cup \{\epsilon\}$. The assignment $x_i^* = \epsilon$ indicates that variable X_i has not been assigned any label.

Solving the Partial Labelling induced Projection. The final solution of the original problem is obtained by minimizing a projection of the original energy function E. This energy projection $E' : \mathcal{L}^{n_\epsilon} \rightarrow \mathbb{R}$ is constructed from $E(\cdot)$ by fixing the values of the labelled variables as: $E'(\mathbf{x}) = E(\mathbf{x}_p)$ where n_ϵ is the number of unlabelled variables i.e. those assigned label ϵ.

2.1 Partial Labelling Quality

We will now discuss the question of *what is a good partial labelling?* If all variables in the partial solution $\mathbf{x}^*$ are labelled, then the projection E' of the energy will take no variables as argument (a constant function) and would be trivially minimized. On the other hand, if all variables are unlabelled, the projection of the energy will be the same as the original energy ($E' = E$) and we would not obtain any speed-up. While constructing the partial labelling $\mathbf{x}^*$, we also want to make sure that all labelled variable are assigned the MAP label i.e. $\mathbf{x}_i^* \neq \epsilon \Rightarrow \mathbf{x}_i^* = \mathbf{x}_i^{\text{opt}}$. This will ensure that $\min E'(\mathbf{x}) = \min E(\mathbf{x})$.

We will measure the quality of a partial labelling using two measures: (1) Percentage of unlabelled variables (P_u) (lower the better), and (2) Percentage of correct label assignments (P_c). Formally, these are defined as:

$$P_u = \frac{100}{|\mathcal{V}|} \sum_{i \in \mathcal{V}} [x_i^p \neq \epsilon], \text{ and } P_c = \frac{100 \sum_{i \in \mathcal{V}} [x_i^p = x_i^{\text{opt}}]}{\sum_{i \in \mathcal{V}} [x_i^p \neq \epsilon]}, \tag{3}$$

where $[arg]$ is as defined above.

Computation Complexity. Let us denote the complexity of the algorithm we are using to minimize the original energy E by $\mathrm{O}(\mathcal{T}(n))$, where $\mathcal{T}(n)$ is any function of n. For instance, the complexity of max-flow based algorithms for minimizing submodular functions of the form (1) is $\mathrm{O}(n^3)$, so $\mathcal{T}(n) = n^3$. The computation time for the multi-scale algorithm can be divided into two parts. **(1)** The time taken for computing the partial solution by minimizing the coarse energy E^l. More specifically, $\mathcal{T}(n\omega)$ for minimizing the energy over $n\omega$ variables and a linear term (n) for extracting the partial solution, thus resulting in the complexity $\mathrm{O}(\mathcal{T}(n\omega) + n)$. **(2)** Time taken for minimizing the projection of the energy which is $\mathrm{O}(\mathcal{T}(P_u))$. The final complexity is: $\mathrm{O}(\mathcal{T}(n\omega) + n + \mathcal{T}(P_u))$.

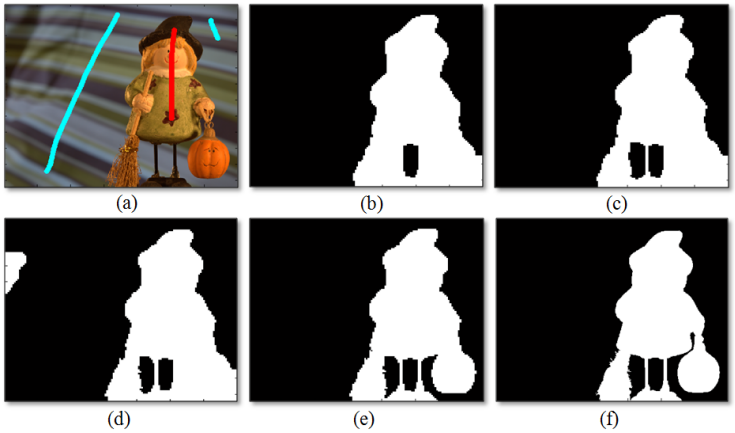

Fig. 2. Constructing the coarse energy. The figure shows the results of using different methods for constructing the coarse energy function. (a) The original image with user marked brush strokes for the different segments (here $\omega_s = 0.04; \theta_s = 5$). (b) The solution $\mathbf{I}$ obtained by constructing the energy using a low resolution version of the original image. (c) Solution $\mathbf{I^c}$ of the energy with the scale-corrected parameter values for the pairwise potentials. (d) Solution $\mathbf{E}$ obtained by using the energy constructed from the original energy. (e) Solution $\mathbf{E^c}$ obtained by using the pairwise potential definition in eqn. (4). (f) The solution obtained by minimizing the energy function (1) defined over the full-resolution image grid.

3 Constructing the Low Resolution Problem

We now explain how a smaller energy minimization problem over the coarse grid $\mathcal{V}^l$ is constructed from the original large scale problem, defined over $\mathcal{V}$.

There is a many-one mapping between points in $\mathcal{V}$ and $\mathcal{V}^l$. We denote the set of indices of nodes in $\mathcal{V}$ which map to the node $i \in \mathcal{V}^l$ by $\mathcal{V}(i)$ which we will call the child set of i. We also define the function $k : \mathcal{V} \rightarrow \mathcal{V}^l$ which given a node i in the original grid, returns the index of its parent node in the reduced grid $\mathcal{V}^l$. For image labelling problems, the traditional approach is to map a square $\theta_s \times \theta_s$ grid of nodes in $\mathcal{V}$ to a single node in the small scale grid $\mathcal{V}^l$, where $\theta_s^2 = \frac{1}{\omega_s}$. We also use this, however, we can extend to other mappings using e.g. super-pixels.

The energy E^l defined over $\mathcal{V}^l$ has the same form as the original energy E (1), with new unary ϕ_i^l and pairwise ϕ_{ij}^l potentials, as defined next.

Scale Dependent Parameter Selection. Traditional band-based multi-scale methods for image segmentation (e.g. [9,13]) define the energy potentials using a low-resolution version $\mathcal{I}^l$ of the original image $\mathcal{I}$. These methods typically overlook the fact the energy should be adjusted and simple use the original energy, i.e. $\phi_i^l = \phi_i$ and pairwise $\phi_{ij}^l = \phi_{ij}$. Figure 2b shows a result, and we refer to the solution with the symbol $\mathbf{I}^1$.

Using the work of Boykov and Kolmogorov [2] it is clear that the strength of the pairwise potentials has to be adjusted when changing resolution. This is due to the fact that the length of the segmentation boundary, in pixel terms, is

[1] It indicates that the coarse scale energy was constructed from the low-res image.

reduced when we move from the original image $\mathcal{I}$ to the low-resolution image $\mathcal{I}^l$. This reduction is inversely proportional to $\theta_s = \frac{1}{\sqrt{\omega_s}}$. Thus, we need to reduce the strength of the pairwise potentials by the same amount, hence the terms in eqn. (2) are chosen as $\{\theta_p^l, \theta_v^l, \theta_\beta^l\} = \{\sqrt{\omega_s}\theta_p, \sqrt{\omega_s}\theta_v, \theta_\beta\}$[2]. Figure 2c shows an example, which we denote by the symbol $\mathbf{I}^c$.

Construction from the Original Energy. A simple method to compute the unary potential for a variable X_i^l is to sum the unary potentials of all the variables in its child set $\mathcal{V}(i)$. Similarly, the pairwise potential for an edge $(u, v) \in \mathcal{E}^l$ is computed by summing the pairwise potentials defined between their children. The segmentation result is shown in fig. 2(d), and denoted solution $\mathbf{E}$

At first glance this strategy seems reasonable, however, it ignores the definition of the pairwise potentials defined on variables X_i and X_j $(i, j \in \mathcal{V})$ which have the same parent i.e. $k(i) = k(j)$. In fact, it can be verified that this approximation is correct only if we assume that Ising model pairwise potentials with infinite cost are defined over every pair of variables X_i and X_j $(i, j \in \mathcal{V})$ which share the same parent.

This situation can be more easily understood by considering the maxflow problem corresponding to the original energy minimization problem. As an example, consider a multi-scale decomposition where variables in a 2×2 square on the original grid share the same parent. The pairwise potential definition would translate into a capacity in the max-flow graph that would allow flow coming into any child node to pass through to any other child node and flow out from it. Obviously, this is a very bad assumption since in reality the child nodes in the graph corresponding to the original energy may be disconnected from each other[3]. This added phantom capacity would make the pairwise potentials very strong and result in over-smooth segmentations (as seen in fig. 2d). We resolve the problem of excess flow capacity by computing a lower bound on the flow that can be passed between child nodes constituting any two sides of the child-set square of a coarse variable $X_i^l (i \in \mathcal{V}^l)$. This capacity is used as the upper bound on the capacity of the edges which connect a particular parent node to other parent nodes. We estimate the lower bound by finding the minimum capacity edge in the child set.

Coming back to the energy formulation, instead of eqn. (2) we use

$$g_e^l(i,j) = \mathcal{R}(i,j) \min_{\substack{k \in \{i,j\}, (u,v) \in \mathcal{E} \\ u \in \mathcal{V}(k), v \in \mathcal{V}}} \theta_p + \theta_v \exp(-\theta_\beta ||I_u - I_v||^2), \tag{4}$$

where $\mathcal{R}(i, j)$ is the number of edges between child-sets of the two coarse level variables X_i and X_j $(i, j \in \mathcal{V}^l)$, i.e. $\mathcal{R}(i,j) = \sum_{(u,v) \in \mathcal{E}: u \in \mathcal{V}(i), v \in \mathcal{V}(j)} 1$. A result is shown in fig. 2e, marked with the symbol $\mathbf{E}^c$.

4 Computing Partial Labellings

Conventional multi-scale methods use the lowest cost solution $\mathbf{x}^{l*}$ of the coarse energy E^l for defining the partial labelling. For instance, the method proposed

[2] Note that the parameter values depend on the topology of the graph, and this equation would be different for the 3D voxel segmentation problem.

[3] This is true if the Ising model penalty for taking different labels is zero.

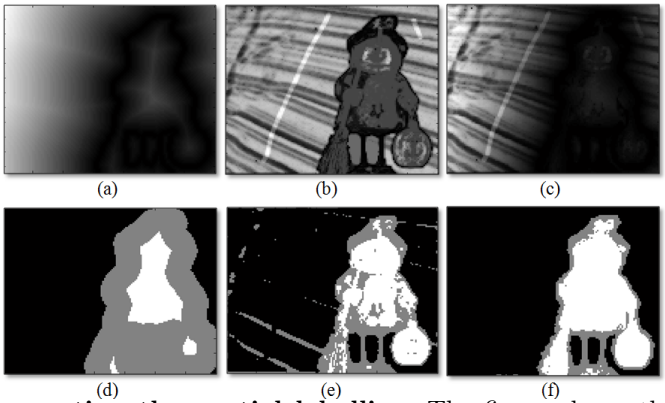

Fig. 3. Computing the partial labelling. The figure shows the results of using different techniques for computing the partial labelling. For this experiment, we used the image shown in figure (2a). We constructed an energy on a coarse level grid with scaling parameter $\omega_s = 0.04$ ($\theta_s = 5$) using the method explained in sec. 3, with the result in fig. (2e). Images (a), (b) and (c) depict the response of a boundary distance function, confidence score $\mu(\cdot)$, and hybrid score $\mathcal{H}(\cdot)$ respectively. These functions were used to obtain the partial solutions shown in images (d), (e) and (f) is the same order. The thresholds for marking pixels as unlabelled was chosen to ensure all marked pixels took the MAP label of the original energy. It can be seen that the hybrid approach requires less number of unlabelled pixels compared to the confidence function which in turn requires less pixels compared to the boundary band approach.

in [9] first defined a full labelling of the original variables $\mathbf{x}^e$ as: $x_i^e = x_{k(i)}^{l*}$, where recall $k(i)$ returns the parent of any variable $X_i, i \in \mathcal{V}$. From the solution $\mathbf{x}^e$, a new set $\mathcal{P}_B(\delta_B)$ is derived, which comprise of pixels that are a maximum distance δ_B away from the boundary. The original problem is then solved only for pixels in $\mathcal{P}_B(\delta_B)$. This band has to be large enough, so that no thin structures are lost, example in fig. 1d. This band-based approach for extracting partial labellings does not take into account the confidence or uncertainty associated with the label assignment for any variable $x_i^l,\ i \in \mathcal{V}^l$, which we will do next.

Partial Labelling from Min-marginals. Given an energy function, the min-marginal encodes the confidence associated with a variable being assigned the MAP label. More concretely, the min-marginal $\psi_{i;a}$ returns the energy value obtained by fixing the value of variable X_i to label a ($x_i = a$) and minimizing over all remaining variables. Formally, $\psi_{i;a} = \min_{\mathbf{x}, x_i = a} E(\mathbf{x})$. Min-marginals naturally encode the uncertainty of a labelling and have been successfully used for solving a number of vision and learning problems, e.g. [5]. The exact min-marginals associated with graph cut solutions can be efficiently computed using dynamic graph cuts in roughly 3-4 times the time taken for minimizing the energy itself [6].

We use the absolute difference between the min-marginals corresponding to the fg and bg labels as our confidence score function $\mu : \mathcal{V} \rightarrow \mathbb{R}$. Formally, $\mu(i) = ||\psi_{i;fg} - \psi_{i;bg}||$. If the difference between min-marginals of any variable

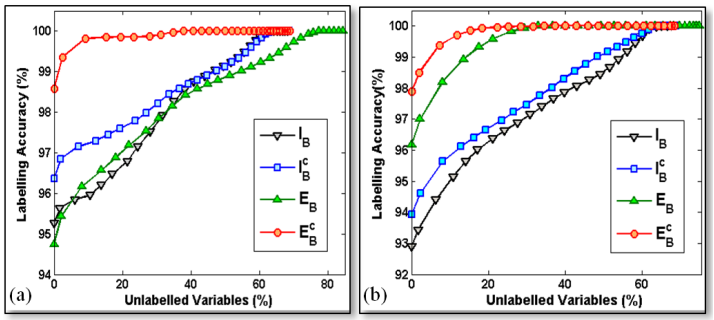

Fig. 4. Results of different multi-scale energy constructions. The graphs show how the accuracy P_c of a partial solution changes as we increase the percentage of unlabelled variables (P_u). Graphs (a) and (b) show the results of using the band based approach on the solutions generated from different energy construction methods. **Key: $\mathbf{I_B}$** : result of energy constructed from the low resolution image, $\mathbf{I_B^c}$: same energy with scale dependent parameters, $\mathbf{E_B}$: smaller problem constructed from the original energy, $\mathbf{E_B^c}$: smaller problem constructed using the lower bound on pairwise potentials.

X_i corresponding to taking the MAP label and any other label is large, then the variable is assigned a high confidence score. The set of variables assigned the label ϵ in the partial solution is now computed by finding the set of nodes $\mathcal{P}_M(\delta)$ whose confidence scores are less than some constant δ_μ i.e. $x_i^e = \epsilon, \forall i \in \mathcal{V}, \mu(i) \leq \delta_\mu$. Formally, the set is defined as $\mathcal{P}_M(\delta_\mu) = \{\ i : i \in \mathcal{V}, \mu(i) \leq \delta_\mu\}$. Similar to the boundary band-width parameter δ_B, the value of the confidence threshold δ_μ can be used to the change the number of unlabelled variables (see fig. 3e).

Although, the min-marginals based confidence function is able to obtain good partial labellings, we observed that it sometimes selects variables which are spatially distant from the main foreground segment. This motivated us to test a new hybrid measure which combines the boundary and uncertainty based techniques described above. We construct the new function $\mathcal{H} : \mathcal{V} \rightarrow \mathbb{R}$ which is defined as:$\mathcal{H}(i) = \mu(i)\mathcal{D}(i)$, where $\mathcal{D}$ is the boundary distance function. As before, the set of variables assigned the label ϵ is now computed by finding the set of nodes $\mathcal{P}_H(\delta) = \{\ i : i \in \mathcal{V}, \mathcal{H}(i) \leq \delta_H\}$. Formally, the partial solution is defined as: $x_i^e = \epsilon, \forall i \in \mathcal{V}, \mathcal{H}(i) \leq \delta_H$. (see fig. 3f)

5 Experiments

Relating Speed with Accuracy. The speed and accuracy of a multi-scale method are inversely proportional to each other. The correctness of the partial labellings can be easily changed by changing the threshold parameters δ_B, δ_μ, and δ_H.[4] The key matter we want to investigate is, how the percentage of variables (P_u) unlabelled in the partial solutions produced by the different multi-scale minimization techniques affect correctness P_c of the solution. We divide our experiments into two parts to investigate how the performance is affected by the use of different: **(1)** Methods for constructing the smaller energy minimization

[4] For instance, setting $\delta_B = \sqrt{I_{width}^2 + I_{height}^2}$ will make sure that all variables in the partial solution are unlabelled. Here I_{width} and I_{height} are the width and height of the image I to be segmented respectively.

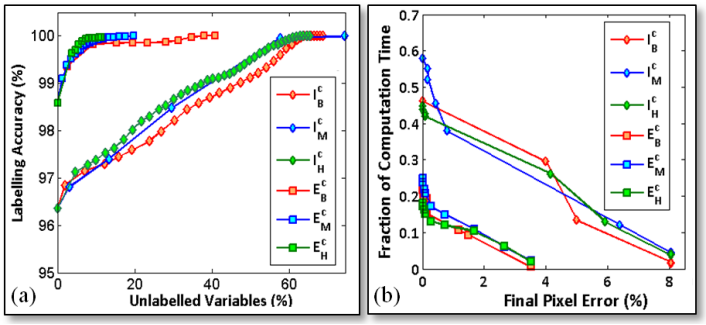

Fig. 5. Computation time and accuracy of different multi-scale methods. Graph (a) shows how the accuracies P_c of partial solutions extracted using different methods change as we increase the percentage of unlabelled variables (P_u). It depicts the results of using different partial solution extraction methods. They key is the same as the one used in graph 4. Subscripts B, M and H denote that the partial solutions were extracted using boundary distance, min-marginal based uncertainty, and the hybrid uncertainty boundary bands respectively. Graph (b) shows the fraction of computation required to achieve a particular pixel labelling accuracy in the final segmentation solution.

problem (section 3), **(2)** Methods for extracting the partial labelling from the smaller energy (section 4).

Comparing Energy Construction Methods. We compared the quality of partial labellings generated from different coarse energy constructions using the boundary band method. The results for the images shown in fig. 1(a) and 2(a) are shown in graphs in fig. 4(a) and (b). It can be seen from the results that using scale dependent parameters is better than the traditional approach. Further, the method for constructing coarse energy directly from the original energy function outperforms other methods. It is able to achieve a correctness of $P_c = 99.5\%$ with less than 10% of unlabelled variables.

Comparing Methods for Partial Solution Extraction. The relative performance of different techniques for extracting the partial solution is now analyzed. Consider the problem of segmenting the image in fig. 2(a). Figure 3 shows the different partial labellings extracted from the coarse energy. The size of the sets $\mathcal{P}_B$, $\mathcal{P}_M$, and $\mathcal{P}_H$ was chosen to ensure that the partial labelling were fully correct ($P_c = 100\%$), i.e. this gives the optimal solution of the original problem. The percentage of unlabelled variables required for the boundary band, uncertainty, and hybrid approaches were 35.29%, 17.36%, and 9.03% respectively. The results on the image shown in fig. 1(a) are shown in fig. 5(a). Due to space, we only show the better performing methods for constructing the energy, i.e. I^c and E^c. It can be seen that the hybrid partial labelling technique results in a much smaller problem to be solved while still obtaining the exact global minimum.

Relating Computational Speedup and Accuracy. We now discuss the speed-up obtained by our multi-scale methods. As explained in section 2, the total computation time $\mathcal{T}$ of a multi-scale method has two primary components: time for partial solution computation t_s, and that for solving the resulting projection (t_p). The size of the projection (and thus t_p) is dependent on the level of accuracy required by the user, while t_s is independent.

For the boundary band method, t_s is equal to the time needed to minimize the coarse energy. For the min-marginal based confidence and hybrid extraction methods, t_s is the time needed to find the min-marginals, which is a much more expensive operation. For instance, for the image shown in figure 1(a), it takes only 1 msec to minimize the coarse energy, while it takes 10 msec to compute all the min-marginals. However, for any given solution accuracy, the min-marginal based methods produce a smaller partial solution compared to the boundary band method. For high levels of accuracy, the size of the projection is large and thus t_p is the dominant time. Thus, min-marginals based methods are able to out-perform band based methods. However, for low levels of accuracy, the size of the projection is very small, which makes t_s to dominate. In such cases, the boundary band based approach outperforms the min-marginals based approach. The performance of all the methods can be seen in the graph shown in fig. 5(b).

6 Discussion and Conclusions

In this paper, we presented a uncertainty driven approach for multi-scale energy minimization. We showed that this strategy allows us to compute solutions close to the globally optimal in a fraction of the time required by a conventional energy minimization algorithm. The method proposed in this paper is general and can be applied to any labelling problem. In future work we would like to investigate how general energies defined over variables with large label sets can be minimized in a multi-scale fashion.

References

1. Y. Boykov and M. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In *ICCV*, pages I: 105–112, 2001. 1, 3
2. Y. Boykov and V. Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *ICCV*, pages 26–33, 2003. 5
3. Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *PAMI*, 2001. 1
4. B. Gidas. A renormalization group approach to image processing problems. *PAMI*, 1989. 1
5. B. Glocker, N. Paragios, N. Komodakis, G. Tziritas, and N. Navab. Optical flow estimation with uncertainties through dynamic mrfs. In *CVPR*, 2008. 7
6. P. Kohli and P. Torr. Efficiently solving dynamic markov random fields using graph cuts. In *ICCV*, volume II, pages 922–929, 2005. 7
7. J. Kopf, M. Uyttendaele, O. Deussen, and M. F. Cohen. Capturing and viewing gigapixel images. *ACM Trans. Graph.*, 2007. 1
8. V. S. Lempitsky and Y. Boykov. Global optimization for shape fitting. In *CVPR*, 2007. 2
9. H. Lombaert, Y. Sun, L. Grady, and C. Xu. A multilevel banded graph cuts method for fast image segmentation. In *ICCV*, 2005. 1, 2, 5, 7
10. P. Pérez and F. Heitz. Restriction of a markov random field on a graph and multiresolution statistical image modeling. *IEEE Trans. on Inf. Theory*, 1996. 1
11. J. Puzicha and J. Buhmann. Multiscale annealing for grouping and unsupervised texture segmentation. *CVIU*, 1999. 1
12. C. Rother, V. Kolmogorov, and A. Blake. Grabcut: interactive foreground extraction using iterated graph cuts. In *SIGGRAPH*, pages 309–314, 2004. 3
13. A. K. Sinop and L. Grady. Accurate banded graph cut segmentation of thin structures using laplacian pyramids. In *MICCAI*, 2006. 1, 2, 5