BAYESIAN RATIONALE FOR THE VARIATIONAL FORMULATION

David Mumford
Harvard University
Department of Mathematics
Cambridge, MA 02138, USA

1. Introduction

One of the primary goals of low-level vision is to segment the domain D of an image I into the parts D_i on which distinct surface patches, belonging to distinct objects in the scene, are visible. Although this sometimes requires high level knowledge about the shape and surface appearance of various classes of objects, there are many low-level clues about the appearance of the individual surface patches and the boundaries between them. For example, the surface patches usually have characteristic albedo patterns, textures, on them, and these textures often change sharply as you cross a boundary between two patches. Therefore, one approach to the segmentation problem has been to try to merge all the low-level clues for splitting and merging different parts of the domain D and come up with probability measures $p(\{D_i\})$ of how likely a given segmentation $\{D_i\}$ is on the basis of all available low-level information, and what is the most likely segmentation. Alternately, one sets $E(\{D_i\}) = -\log(p(\{D_i\}))$, which one calls the 'energy' of the segmentation, and seeks the segmentation with the minimum energy. In general, these models have two parts: a *prior model* of possible scene segmentations, possibly including variables to describe other scene structures that are relevant (e.g. depth relationships), and a *data model* of what images are consistent with this prior model of the scene. If we write w for the variables used to describe the scene, e.g. the subsets D_i or the set of all their boundary points Γ, then the prior model is some probability space (Ω_w, p), where Ω_w is the set of all possible values of w. The model is specified by giving the probability distribution $p(w)$ on all these values. The data model is a larger probability space $(\Omega_{w,d}, p)$, where $\Omega_{w,d}$ is the set of

all possible values of w and of all possible observed images I. This model is completed by giving the conditional probabilities $p(I|w)$ of any image I given the scene variables w, resulting in the joint probability distribution:

$$p(I, w) = p(I|w) \cdot p(w)$$

The discussion above assumes implicitly that the spaces Ω_w and $\Omega_{w,d}$ are finite, although huge (e.g. the set of byte valued images I on a grid of size 256x256 has cardinality about 10^{150000}). In many situations, it is more convenient to consider images as real-valued functions of continuous variables, and to consider segmentations as sets of suitable measurable subsets of D: then more complicated probability spaces are needed, often using distributions as well as actual functions. We will not worry about this, as the expressions for the probability densities we use all look like $p = Z^{-1} \cdot p'$, where p' has a simple limiting expression as the exponential of an integral and Z is a normalizing constant introduced to make p into a probability measure. The only problem in the continuous limit is that $Z \to \infty$, hence $p \to 0$. But we can work with p' in the continuous limit, knowing that in finite approximations, the normalizing constant Z is finite. In terms of energy, this means that the E we work with should have an infinite constant added to it, which, in finite approximations, is finite.
These models are always used in conjunction with Bayes's theorem, which factors the joint probability distribution the other way:

$$\begin{aligned} & p(I, w) &=& \ p(w|I) \cdot p(I) \\ \text{hence} \quad & p(w|I) &=& \ \frac{p(I|w) \cdot p(w)}{p(I)} \\ & &\propto& \ p(I|w) \cdot p(w) \end{aligned}$$

The probability of w *given the data* I is called the *posterior probability* of w, and this what we want to calculate. In terms of energy, we can write:

$$\begin{aligned} E(w) &= -\log(p(w, I)) \\ &= -\log(p(I|w)) - \log(p(w)) \\ &= E_d(I, w) + E_p(w) \end{aligned}$$

where the goal is now to minimize $E(w)$. The log of the prior term, E_p, is sometimes called the 'regularizer' because it was initially conceived of as a way to make the variational problem of minimizing E_d well-posed. In what follows, it will play a much more central role of measuring how reasonable each scene model is, lower values being the more common scenes and higher values the less common ones.

What we want to do in the rest of this chapter is to present four such energy models for image analysis of increasing sophistication, which attempt to capture more and more of the subtleties of actual world scenes. It will be clear that none of these models captures all the important scene variables and that this is just the beginning of the exploration of probabilistic models for low-level vision. For instance, none of these models include explicit illumination variables, which are often essential to disambiguating scene structure. Also, we have not included any models of this type which build on *multiple* images, i.e. stereo pairs or motion sequences. Multiple images without a doubt make it infinitely easier to properly segment images (how many animals understand the content of photographs?). Notable models of this type are due to Belhumeur [29] and to Weber and Malik [375]. It is my belief that a robust solution to the general low-level vision problem can be found using this approach. The main obstacle is to find more effective and faster ways of estimating the w minimizing $E(w)$ than those presently available.
Although this energy approach seems on the surface to be totally different from the non-linear PDE's investigated in the rest of this book, they are in fact closely linked to each other. Some of these links can be seen in the contributions by Mitter and Richardson and by Nordstrom to this book. Others can be found in Geiger and Yuille's work [130].

2. Four Probabilistic Models

2.1. The Ising model

This is a model which comes directly from statistical physics, where it is used to describe a two-dimensional crystal of iron atoms subject to an external magnetic field. It was a very influential model in statistical physics, because it was the first mathematical model which was rigorously proven to model phase transitions (for discrete but infinite domains D) [283]. In vision, it models images which are made up of a set of white blobs against a dark background (or vice versa) and where one seeks to describe the white blobs by an auxiliary binary image, or, equivalently, by a subset $S \subset D$.
The prior model is very simple: we ask that S consists of a small number of compact blobs, i.e. that the length of the perimeter of S is as small as possible. More precisely, define ∂S as the boundary of S: in the discrete case, this is the set of pairs of horizontally or vertically adjacent pixels (α, β), one of which is in S and the other of which is not. In the continuous case, this is set of points (x, y) which are in the closure of S and $D - S$. Define $|\partial S|$ as the cardinality of ∂S in the discrete case and as the length (= 1-dimensional Hausdorff measure) of S in the continuous case. Then the prior model is determined by setting $E_p(S) = \nu|\partial S|$: this means that the

shorter the perimeter of S, the more likely S is to be the model of the scene. Let χ_S be the characteristic function of S, and let I be the observed image. The blobs are supposed to be characterized by being more or less bright compared to the background. However, we assume that the intensity of neither the blobs nor the background is uniform, but that it fluctuates randomly and is corrupted by many kinds of noise. If the image were simpler, we could recover S by simply thresholding I. However, if there is a substantial amount of noise present, no matter what threshold is chosen, we may find not S but S with lots of extra specks and hairs, minus small holes and cracks. This is exactly what the model seeks to correct. Assume for simplicity that on the blobs S the image I tends to have values bigger than 0.5 and that on the background, the image has values less than 0.5. (The data model can be modified for other thresholds: the original Ising model used 0). Then the data model is given by $E_d(I,S) = \mu \iint_D (I-\chi_S)^2 d\vec{x}$. This is equivalent to assuming that $I = \chi_S + n$, where n is white noise.
We may summarize the model by:

$$\begin{aligned} w &= S, S \subset D \\ E_p &= \nu|\partial S| \\ E_d &= \mu \int\int_D (I-\chi_S)^2 d\vec{x}, \qquad \chi_S = \text{ char.function of } S \end{aligned}$$

In figure 5.1, we have illustrated this model by an image I consisting of a scene with a cow, tree and foreground in deep shade against a background of the sky and more distant parts of the scene. The figure shows both the original scene and the binary image given by the Ising model optimal S (for suitable ν, μ).

2.2. The Cartoon Model

The cartoon model is the model which has been most used in vision. It was invented in the discrete case, independently by S. and D. Geman in their influential paper [132] and by A. Blake and A. Zisserman [34]. J.Shah and I then investigated the corresponding continuous model in [264]. In fact, the variational problem of minimizing energy functionals of this continuous kind had also been independently invented by De Giorgi and his school at about the same time in modeling materials with 2 phases and a free interface [83].

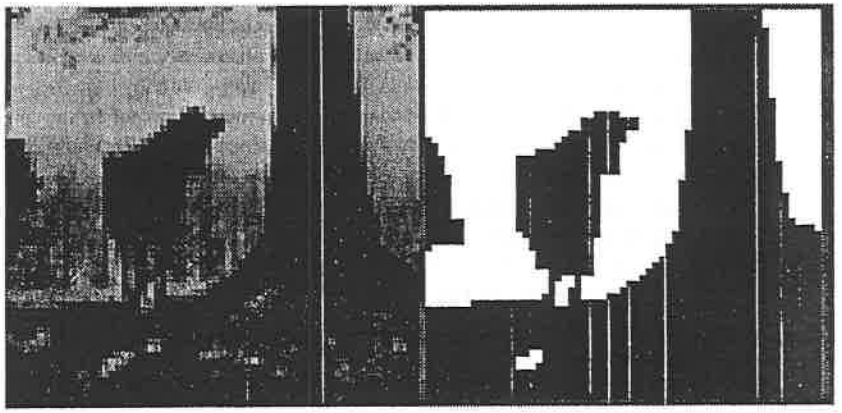

Figure 5.1. A scene with a cow and trees, and its Ising model results.

Instead of assuming that there exists a binary segmentation of the image into contrasting light and dark regions, in this model we assume that the real world scene consists of a set of *shaded* regions within which the intensity changes slowly, but across the boundaries between them, the intensity changes, in general abruptly. Thus what we want to infer is not the set S of light (or dark) foreground regions, but a *cartoon* consisting of a simplified noiseless version J of I. The cartoon has a curve Γ of discontinuities, but everywhere else is assumed to have small gradient $\|\nabla J\|$. The prior model can be built up by starting with the Ising model prior of Γ: $E_p^{(1)}(\Gamma) = \nu|\Gamma|$. We then put a prior on J by asking that its gradient be small: $E_p^{(2)}(J,\Gamma) = \iint_{D-\Gamma} \psi(\|\nabla J\|) d\vec{x}$, where ψ is some convex even function. The standard choice is $\psi(x) = x^2$, but $\psi(x) = |x|$ is also very interesting and should be more investigated. The full prior is $E_p = E_p^{(1)} + E_p^{(2)}$. The data model is again essentially the same as for the Ising model, except that instead of assuming $I = \chi_S + n$, where n is white noise, we assume $I = J + n$. This gives $E_d = \iint_D (I-J)^2 d\vec{x}$. We may summarize this model by saying that we seek to approximate an arbitrary function I by a piecewise smooth function J so that three things are kept as small as possible: i) the difference between I and J, ii) the gradient of J where it is smooth and iii) the length of the curve Γ where J has discontinuities.
In the case where D is discrete, it is always obvious that any energy functional like our E has a minimum, because we are minimizing over a finite set of possible Γ and the functional is continuous in the values of J, and goes to infinity if any of the values of J goes to infinity (i.e. it is 'proper' as a function of J). However, if D is continuous, it is not at all clear

that E has a minimum in any sense. This is referred to as asking whether the variational problem associated to E is 'well-posed' or not. This was an open question in [264]. A so-called 'weak solution' was given by Ambrosio [10] where the cartoon J was allowed to be a very nasty sort of function, however: a so-called 'Special Bounded Variation' function. A first step in showing this weak solution was not horrendous consisted in proving that Γ was closed, see [83]. Shah and I conjecture that at minima of E, the curve Γ consists in a finite set of C^1 arcs (possibly with cusps at the ends of branches which terminate) but this is unproven: a survey of the theory of this functional is given in the chapter by Leaci and Solimini in this book. The computer scientist might be tempted to say: what do I care about the continuous case and the subtle estimates mathematicians require to prove that problems like $\min E$ are well-posed. The remarkable thing is that the estimates used in establishing the *existence* of well behaved solutions are exactly the same as the estimates which enable the engineer who wants to use the discrete model to be sure that his minima *behave predictably*: that this finite model doesn't produce artifacts or perceptually meaningless zigs and zags dependent on small details of how the problem is discretized.

In the discrete case, the functional E can be re-written in a suggestive way. The simplest form of the term $E_p^{(2)}$ in the discrete case is

$$E_p^{(2)} = \sum_{\text{adj. pixels } \alpha,\beta} \psi(J(\alpha) - J(\beta)).$$

Then it is easily seen that at a minimum of $E(J,\Gamma)$, Γ cuts an adjacent pair of pixels α,β if and only if $\psi(J(\alpha) - J(\beta)) > \nu$. Thus if we define

$$\psi'(x) = \min(\psi(x), \nu), \tag{5.1}$$

$$E_p'^{(2)}(J) = \sum_{\text{adj.pixels}} \psi'(J(\alpha) - J(\beta)) \quad \text{and} \tag{5.2}$$

$$E'(J) = E_d(J) + E_p'^{(2)}(J) \tag{5.3}$$

we find that

$$\min_{\Gamma} E(J,\Gamma) = E'(J).$$

This form enabled Blake and Zisserman [34] to analyze many properties of E, and to approximate E' by a third functional in which ψ' was replaced by a smooth ψ'' which, even though not convex itself, made E'' convex! They use this as a basis for a continuation method of getting good *local* minima of the original functional E.

To carry over this approach in the continuous case, however, requires that we modify the variable Γ, replacing it by a 'line process' $\ell(x,y)$ which is a

smooth function with values in $[0,1]$, mostly zero but climbing to one along Γ. One then replaces $E(J,\Gamma)$ with

$$E(J,\ell) = E_d(J) + \iint_D (1-\ell)\psi(\|\nabla J\|)d\vec{x} + \iint_D \phi_c(\ell)d\vec{x}$$

for suitable ϕ_c. Such ϕ_c are described in the chapter by Mitter and Richardson in this book, where they also give theorems on when these functionals approximate the original $E(J,\Gamma)$. This approach seems very useful because it offers a way of taming the wildness inherent in having Γ itself as a variable.

In a nutshell, here is the cartoon model:

$$\begin{aligned} w &= (J(\vec{x}),\Gamma) \\ E_p &= \int\int_{D-\Gamma} \psi(\|\nabla J\|)d\vec{x} + \nu|\Gamma| \\ E_d &= \int\int_D (I-J)^2 d\vec{x} \end{aligned}$$

In figure 5.2, we give an illustration of this model applied to a close-up of Marilyn Monroe's eye. The original eye is shown on the left, then the cartoon J and finally the contours Γ.

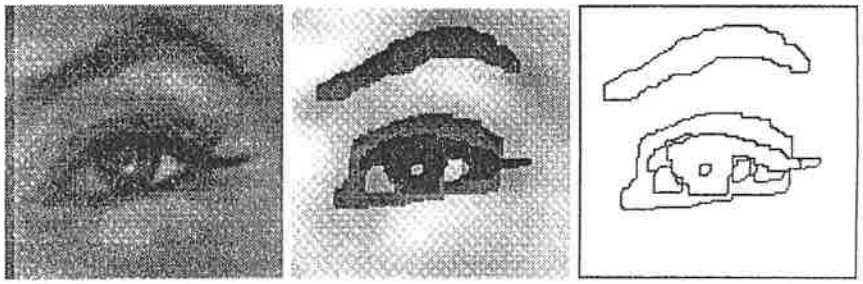

Figure 5.2. An image of the eye and its final cartoon and boundary, produced by graduated nonconvexity algorithm.

2.3. The Theater Wing Model

Although the model in the previous section is attractive and interesting from a mathematical point of view, it unfortunately is very crude as a model of image segmentations. One of its problems is that where 3 domains D_i meet at a point P, so that the boundary Γ has a singular point with 3 branches

meeting at P, the branches will meet at 120° angles. This is because locally the effect of I is very weak and Γ behaves like soap bubbles do when 3 sheets meet: they form 120° angles. In images, on the contrary, 3 domains meeting usually means that the edge of a foreground object cuts across the edge of a more distant object. This gives instead a "T-junction" on Γ: Γ consists locally of a smooth curve Γ_1 through P with a second smooth branch Γ_2 ending at P. These singularities are not only typical of real world images, but they are very powerful psychological clues to depth relations, as the Gestalt school of psychology and especially Kanisza [172] discovered.

The problem is that we have not included even qualitative depth information in our model variables $\{w\}$ and occlusion edges in the real world are inherently asymmetrical, having a nearer and a farther side. (Nakayama refers to this by saying that an edge 'belongs' to one of its sides.) In the previous model, no region is considered foreground nor any background. Working with Mark Nitzberg [270], we sought a model which was a reasonable first step in modeling sets of regions occluding each other. This model goes back to the Ising model assumption that the individual regions have more or less uniform brightness, but it now assumes they are ordered in depth, and one region can vanish behind another, only to reappear elsewhere. The regions cannot change their depth relations however, interweaving like wicker chairs, nor can a region circle round and occlude itself (like your palm when you bend your thumb over it). This is why we called it the 'theatre wing model'.

The basic variable for this model is a sequence of regions $D_1, D_2, \cdots, D_k$ in D, which is *ordered*, where D_i represents the parts of the domain D where object i would be visible *if all closer objects were removed.* The ordering represents depth: object 1 is nearest, object 2 is behind it, object 3 behind both, etc., while object k, called the background and assumed equal to D, is most distant. What is the prior model here? Since we assume all singularities of Γ come from T-junctions where one occluding boundary interrupts another, we can now assume all D_i have smooth boundaries and include not only the length but the curvature of ∂D_i into the prior. Thus our model asks for a small number of regions D_i with short smooth boundaries. In some cases, it turned out that we also needed to ask that their areas not be too big either, so the final prior we chose was $E_p = \sum(\int_{\partial D_i} \psi(\kappa_i)ds + \epsilon\text{Area}(D_i))$. Here a typical choice is $\psi(x) = a + bx^2$.

The data model assumes that the intensity of each region is more or less uniform, but that its mean intensity may be anything. This leads us to $E_d = \sum\int\int_{D_i'}(I - m_i)^2$, where D_i' is the *visible* part of D_i, i.e. D_i minus the parts $D_i \cap D_j$ occluded by nearer parts $j < i$, and where m_i is the mean value of I on this visible part (we have no way of knowing what is the brightness of the occluded parts of D_i!). In summary:

$$
\begin{aligned}
w &= \{D_1, D_2, \cdots, D_k\} = D, \\
E_p &= \sum\int_{\partial D_i}\psi(\kappa_i)ds + \epsilon\sum\text{Area}(D_i), \qquad \kappa_i = \text{curvature}(\partial D_i) \\
E_d &= \sum\int\int_{D_i'}(I - \text{mean}_{D_i'}(I))^2 d\vec{x}, \qquad D_i' = D_i - \cup_{j<i} D_j
\end{aligned}
$$

In figure 5.3 we show an example of this model. The image I depicts a beer bottle, an orange and a potato occluding each other. The figure shows how the theory correctly reconstructs their occlusion relations and gives its best shot at guessing how their contours continue behind each other. It might be thought that this kind of wild reconstruction of occluded contours is irrelevant to the interpretation of the scene. But curiously, extensive experiments by the Gestalt school of psychologists and more recently by Nakayama and his collaborators have shown that people very frequently make exactly such reconstructions, sometimes choosing one of several reasonable 'amodal' contours for seemingly unaccountable reasons. If people do it, it may not be absurd for computers to do it too.

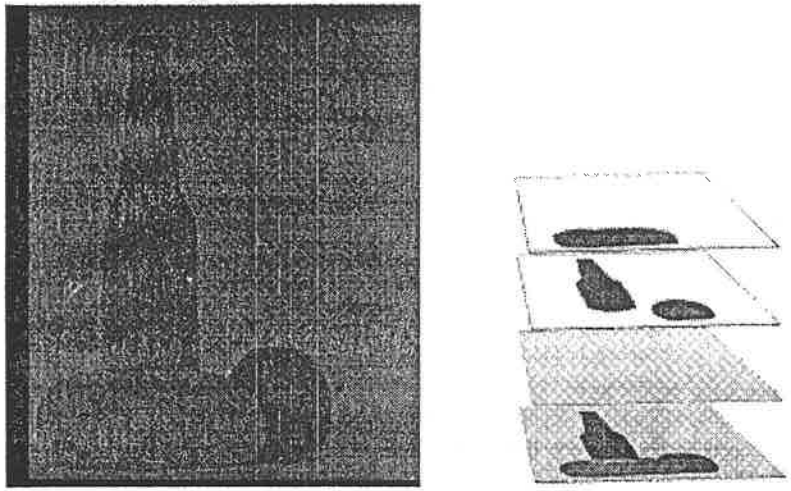

Figure 5.3. Still Life and its 2.1-D sketch.

2.4. The Spectrogram Model

So far our models have had one glaring omission: they have assumed that all visible surface patches have intensities which are slowly varying plus

noise. In fact, this is wrong for a majority of surfaces: much more often, surfaces have a natural texture in which their intensity has strong systematic variation. To segment in real world scenes, we need to model not merely nearly uniform intensity but nearly uniform or slowly varying texture. There are many puzzles which are still unsolved about what constitutes a coherent, perceptually distinguishable texture, and we don't want to get involved with this. As in the previous section, working with Tai Sing Lee [207], we sought to model the simplest types of texture segmentation. Our approach here is to assume that each texture has, locally, a *spectral signature* which characterizes it.

Thus we start by taking a local Fourier analysis of the image I. There are two ways of doing this. The more traditional way is to form the windowed spatial (2-dimensional) Fourier transform of $I(x,y)$:

$$\mathcal{F}(I)(\vec{x_0},\vec{\xi}) = \iint I(\vec{x}+\vec{x_0})w(\vec{x})e^{2\pi i\vec{x}\cdot\vec{\xi}}d\vec{x}$$

where w is the window function. If we note that we may rewrite this using the scaling and rotation matrix

$$A_\xi = \begin{pmatrix} \xi_1 & \xi_2 \\ -\xi_2 & \xi_1 \end{pmatrix}$$

noting that $A_\xi \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \vec{\xi}$, we get:

$$\mathcal{F}(I)(\vec{x_0},\vec{\xi}) = \iint I(\vec{x})w(\vec{x}-\vec{x_0})e^{2\pi i A_\xi(\vec{x}-\vec{x_0})\cdot\vec{\xi}}d\vec{x}.$$

Then all we need to do is to change the windowing function with the frequency to get a new transform $\mathcal{W}$:

$$\mathcal{W}(I)(\vec{x_0},\vec{\xi}) = \iint I(\vec{x})w(A_\xi(\vec{x}-\vec{x_0}))e^{2\pi i A_\xi(\vec{x}-\vec{x_0})\cdot\vec{\xi}}d\vec{x}.$$

This is exactly the *wavelet* transform of I associated to the wavelet $\psi(\vec{x}) = w(\vec{x})e^{2\pi i\vec{x}\cdot\vec{\xi}}$. Either way, what we use to model the texture is its local spectral power:

$$\mathcal{P}(I) = |\mathcal{F}(I)|^2 \text{ or } |\mathcal{W}(I)|^2.$$

The model is based on the idea of finding a cartoon for $\mathcal{P}(I)$. More precisely, we seek firstly a set of boundary curves Γ in the spatial domain D (n.b. *not* boundaries in the 4-dimensional $D\times\hat{D}$ space of the variables $(\vec{x},\vec{\xi})$ where $\hat{D}$ is a part of the dual two-dimensional vector space of the spatial frequency variables (ξ_1,ξ_2) with suitable high and low frequency cut-offs). Secondly,

we seek a smooth spatial frequency description $J(\vec{x},\vec{\xi})$ of the signal on $(D-\Gamma)\times\hat{D}$. The prior model is just like that of the cartoon: terms for the length of Γ and for the gradient (now 4-dimensional) of J. The data model has an important subtlety: one of the problems which bedevil texture segmentation is that spectral filters whose support overlaps the correct boundary between 2 quite distinct textures give erratic responses as a result of the partial fields visible in each texture. Thus J should not be compared with $\mathcal{P}(I)$ for such fields. For the windowed Fourier transform, this will be a strip around Γ of fixed size; for the wavelet transform, this is rather a shadow cast by Γ from high frequencies, i.e. at high frequencies, the window is smaller and the data can be modeled very close to Γ, while at low frequencies, the window grows and the strip of mixed responses grows. In either case, call $\mathcal{S}(\Gamma)$ the set of points of $D\times\hat{D}$ for which the corresponding filter overlaps Γ. Then the model can be summarized by:

$$\begin{aligned} w &= (J(\vec{x},\vec{\xi}),\Gamma) \\ E_p &= \int\int\int\int_{(D-\Gamma)\times\hat{D}} \psi(\|\nabla_{\vec{x}}J\|,\|\nabla_{\vec{\xi}}J\|)d\vec{x}d\vec{\xi} + \nu|\Gamma| \\ E_d &= \int\int\int\int_{(D\times\hat{D})-\mathcal{S}(\Gamma)} (\mathcal{P}(I)-J)^2 d\vec{x}d\vec{\xi}, \quad \mathcal{P}=\text{local spectral power} \end{aligned}$$

In figure 5.4, we show an example of this model for a lady with scarf: note how it finds most of the edges where the scarf has folds causing a break in texture statistics, but that it treats as uniform the slow changes where the scarf bends around her head. It should be noted that the figure probably does not present the minimizing Γ for this figure: rather it shows the best Γ found by either the Geman's annealing algorithm or Blake and Zisserman's continuation algorithm. I expect that the model will give even better results when a more effective optimizing technique is devised for energy functionals like this one.

I would also like to mention that this model should be quite interesting for *speech*. In speech, we have a function $I(t)$ of time, which gives us a time-frequency local power descriptor $\mathcal{P}(I)(t,\omega)$. Speech naturally breaks up into phonemes each with a typical power spectrum. This power spectrum changes slowly during a phoneme and changes rapidly when one phoneme succeeds another. Thus a model like that just given provides a method of segmenting speech without the detailed modeling of each individual phoneme required by the standard 'HMM' approach to speech.

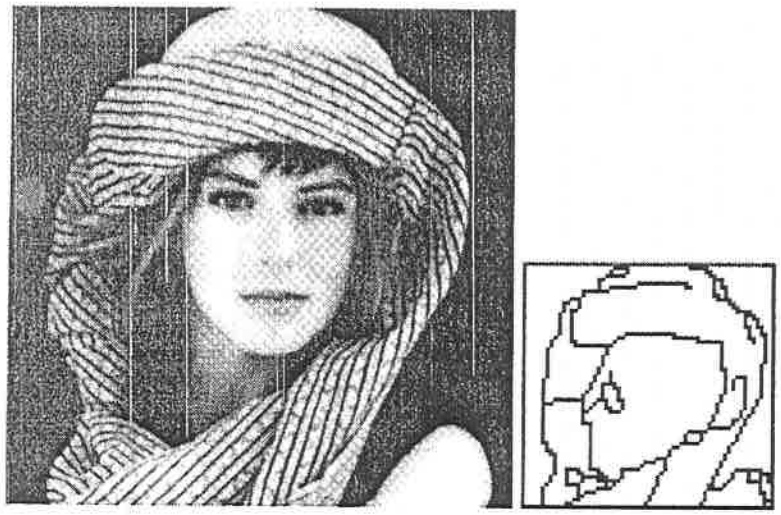

Figure 5.4. Lady with a scarf and the segmentation boundary.

A Low-Dimensional Representation of Human Faces For Arbitrary Lighting Conditions

Peter W. Hallinan
Division of Applied Sciences
Harvard University
Cambridge, MA 02138

Abstract

When recognizing a fixed object from a fixed viewpoint, the dominant source of variation in image intensity is lighting changes. We propose a low-dimensional model for human faces that can both synthesize a face image when given lighting conditions and can estimate lighting conditions when given a face image. The model can handle non-Lambertian and self-shadowing surfaces such as faces because it does not make any assumptions about either the surface's geometry or bidirectional reflectance function. The model can be adapted to handle any arbitrary lighting condition, and is easily extendable to any other viewpoint or to any other object.

1 Introduction

To date, almost all research in object recognition has either attempted to discount lighting effects by employing such allegedly lighting-invariant features as edges, or has assumed that the features to be used could be chosen according to known or previously estimated lighting conditions. Unfortunately, for any specific intensity-based feature, there exists some lighting condition that will remove it or otherwise alter it, and methods to analyze lighting conditions before feature extraction is performed typically make strong assumptions about the surface geometry or bidirectional reflectance function.

In this paper, we propose a model for object recognition and scene analysis that permits a recognition system to both estimate lighting conditions given an image and synthesize an image given the lighting conditions. The model is tested on faces, but the model should work for any object. Given a face viewed from a fixed direction, our approach uses principal components analysis (PCA) to find an approximation to the basis for images of the face. In its use of PCA, the model is related to past work by Sirovitch and Kirby [0] and Turk and Pentland [0]. More closely related work is that by Shashua [0] and Moses [0] who independently proposed that lighting effects could be discounted using a 3 dimensional basis for face space with basis vectors that were actual images. However, the model we propose is designed to *explain* lighting conditions and not to discount them. Thus the model is evaluated not by a recognition rate for a set of faces, as is the case for previous eigenimage-based systems, but by its success in the synthesis and analysis of face images under different lighting conditions.

The new contributions in this paper are (1) an explanation of why arbitrary lighting conditions can be modeled using an image basis, (2) a method for constructing such a model, and (3) a set of results showing both that five eigenfaces suffice to analyze and synthesize images under a wide range of lighting conditions, and that the eigenface model performs much more successfully than a model employing actual images as basis vectors. These results hold both for models of individual faces and for a model of a generic male face. Details are given in [0].

2 Constructing the Lighting Model

Since any given set of lighting conditions can be exactly decomposed as a sum of point light sources, a surface patch's radiance when illuminated by two light sources is the sum of the radiances for the light sources applied separately. Thus for any given object and viewing projection, there is, up to a scale factor $c(\theta, \phi)$, just one image $\tilde{I}(\theta, \phi)$, called a *boundary* image, that is associated with each point light source $\tilde{L}(\theta, \phi)$, where θ is longitude and ϕ is latitude. Of