Generalized Graduated Non-Convexity Algorithm for Maximum A Posteriori Image Estimation[1]

A. Rangarajan and R. Chellappa

Signal and Image Processing Institute

University of Southern California

Abstract

We are interested in restoring a degraded scene while preserving the edges. Edges are represented as line processes and are estimated along with the intensities in a Maximum A *Posteriori* (MAP) framework. Assumptions regarding the prior and degradation distributions reduce the problem to one of energy function minimization. The energy function is highly non-convex and finding the global minimum is a non-trivial problem. When constraints on the interactions between line processes are removed, the deterministic, Graduated Non-Convexity (GNC) algorithm has been shown to find close to optimum solutions.

We have generalized the GNC model. Any number of constraints on the line processes can now be added. This has been achieved by using the *adiabatic approximation*, a well known technique in synergetics. Our resulting algorithm is a combination of the Conjugate Gradient (CG) and the Iterated Conditional Modes (ICM) algorithms and is completely deterministic. Since the GNC algorithm can be obtained as a special case of our approach, we refer to our algorithm as the Generalized GNC or G^{2}NC algorithm. The algorithm was executed on two aerial images. Results are presented along with comparisons to the GNC algorithm.

1 Introduction

Intensity edges are an important source of information in early vision. Even in a degraded scene, the human visual system possesses the ability to quickly extract the principal features of interest. Consequently, it is of interest to us to be able to restore a degraded scene while simultaneously preserving the edges.

We formulate the problem in a Bayesian framework. Edges are represented as discontinuity variables or *line processes*. The line processes are chosen to be in between the intensity sites and are either vertical or horizontal in orientation. The prior distribution is defined jointly over the intensities and the lines. The prior and degradation distributions are chosen to be Gibbs. Then the posterior distribution is also a Gibbs distribution [1]. All that is needed to specify a Gibbs distribution is an *energy function* (*Hamiltonian*). It can be shown that MAP estimation in the case of a Gibbs distribution is equivalent to minimizing an energy function [1].

In this paper, the energy function used is made up of four terms; (i) closeness to data (ii) smoothness constraint except over the discontinuities (iii) penalty for imposing a discontinuity and (iv) penalties on broken contours. Finding the global minimum of this energy function is complicated by the existence of several local minima. When the energy function is highly non-convex, it is difficult to visualize any scheme other than stochastic relaxation (SR) which can help us avoid these local minima, [1], [2]. Deterministic methods have been used [3], [4], [5], [6] but have been shown to converge to near optimum solutions only when there are no constraints on the interactions of the line processes [7], [8], [9], [10]. One of the most popular of these techniques is the Graduated Non-Convexity (GNC) algorithm.

The GNC algorithm as the name suggests starts with a convex approximation of the original energy function. In order to achieve the objective of a convex approximation, several constraints on line process interactions in the form of penalties levied on broken contours etc. are removed. The energy function becomes a function of the intensities and a control parameter. The next step is to successively minimize a sequence of energy functions generated by varying this control parameter with the first stage being convex. This places it in the class of the so-called continuation methods [10]. The actual algorithm is highly successful in achieving its limited objective. We wish to improve the prior distribution to include line process interactions. This is missing in the GNC algorithm. Generalization of the GNC model is complicated by the fact that the GNC energy function is a function of the intensities alone.

We derive a new energy function for MAP estimation. This energy function is indexed by a control parameter which is varied during the relaxation process. A sequence of minima are tracked as this parameter is varied. We show that any general constraint on the interactions of the line processes can be included in this energy function. We have chosen a basic interaction term which proved to be successful in Mean Field theory [4] and demonstrate its efficacy in our set up. The relationship between this Hamiltonian and the GNC energy function is shown by appealing to the *adiabatic approximation*, a basic technique in synergetics [11], [12]. Dynamical systems are generated by gradient descent on a Hamiltonian. The adiabatic approximation states that, in such a dynamical system, those variables that are fast relaxing can be eliminated from the system by setting their time variation to zero. The relationship to the GNC function can now be shown. We first assume an energy function with a general structure. When the line processes are eliminated adiabatically from this energy function, the new Hamiltonian is a function of only the intensities. We equate this Hamiltonian and the dynamical system generated by it to the GNC sequence and its corresponding dynamical system. The structure of the new Hamiltonian gets fixed in this process. This allows us to derive an energy function sequence which is a function of the intensities and the line processes.

CH2898-5/90/0000/0127$01.00 © 1990 IEEE

Any number of constraints can now be added to the new sequence. Adding these constraints directly makes it difficult to minimize the new energy function. The interactions are transferred to a new domain; the gradient-magnitude or (GMAG) domain. It is argued that this domain is more natural for the introduction of further constraints on the discontinuity field. A transformation is derived from the GMAG process to the line process. Minimization of this new sequence of energy functions is now straightforward. We have chosen to run the Conjugate-Gradient (CG) algorithm [13] on the intensities and the Iterated Conditional Modes (ICM) algorithm [14] on the line processes. Since the GNC algorithm can be obtained as a special case of our algorithm, we refer to this as the Generalized GNC (G^2NC) algorithm.

2 MAP estimation via the adiabatic approximation

We derive new energy functions for MAP estimation and show that the GNC algorithm can be brought under our framework. The adiabatic approximation, a technique popularized by the field of synergetics [11] is crucial to our framework.

2.1 New energy functions for MAP estimation

In the case of image estimation, the posterior distribution and the Hamiltonian are linked by the following Gibbs distribution.

$$\mathcal{P}(\mathbf{F}=\mathbf{f},\mathbf{V}=\mathbf{v},\mathbf{H}=\mathbf{h}/\mathbf{D}=\mathbf{d}) = \frac{1}{Z}e^{-\beta\mathcal{H}(\mathbf{f},\mathbf{v},\mathbf{h})} \quad (1)$$

where $\mathbf{d}$ is the observed, noisy image, $\mathbf{f}$, $\mathbf{v}$, and $\mathbf{h}$ are the intensity, vertical and horizontal line processes respectively, Z is the *partition function*, and β is a constant greater than zero. We assume that the original image is corrupted by additive noise to yield the degraded image. The line processes are located in between adjacent intensity variables and denote the presence (or absence) of a discontinuity. We seek the minimum of the Hamiltonian ($\mathcal{H}$) or energy function in the phase space of $\mathbf{f}$, $\mathbf{v}$, and $\mathbf{h}$.

$$\begin{aligned}\mathcal{H}(\mathbf{f},\mathbf{v},\mathbf{h}) &= \sum_{\{i,j\}}(f(i,j)-d(i,j))^2 \\ &+\sum_{\{i,j\}}\left(\lambda^2 f_x^2(i,j)(1-v(i,j))+\alpha\, v(i,j)\right) \\ &+\sum_{\{i,j\}}\left(\lambda^2 f_y^2(i,j)(1-h(i,j))+\alpha\, h(i,j)\right) \\ &+\mathcal{H}_c(\mathbf{v},\mathbf{h}),\ v(i,j),h(i,j)\ \in\ \{0,1\} \end{aligned} \quad (2)$$

where $f_x(i,j) \stackrel{\text{def}}{=} f(i+1,j)-f(i,j)$ and $f_y(i,j) \stackrel{\text{def}}{=} f(i,j+1)-f(i,j)$.

The first term in (2) attempts to keep the restoration of $\mathbf{f}$ close to the observed data in a least squares sense. The second term encodes our assumption that the data is smooth everywhere except at the discontinuities. The third term enforces a penalty for incorporating a discontinuity in the restoration. The last term ($\mathcal{H}_c(.)$) can be used to incorporate prior assumptions on the nature of discontinuities. Edges in images tend to be connected and form closed contours. This term can reflect these beliefs. When the $\mathcal{H}_c(.)$ term is absent, this energy function reduces to the popular *weak membrane* [7].

The weak membrane energy function is highly non-convex due to the inclusion of the line processes. Deterministic schemes have generally relied on continuation methods to solve this problem. Here, a control parameter is varied during the relaxation. Initially, the problem solved is convex or nearly convex and as the parameter is varied, the successive energy functions approach the original. The advantage of these deterministic methods over SR is that they avoid shallow local minima without incurring the computational burden of SR.

We have modified the Hamiltonian $\mathcal{H}$. A new sequence of energy functions is obtained to which interactions can be added. The line processes now vary over $[0,1]$ and deterministic algorithms can find "good" suboptimal solutions. This new energy function depends on a control parameter c and a sequence of minima can be tracked while varying this parameter. We first present this energy function $\mathcal{H}_a$.

$$\begin{aligned}\mathcal{H}_a(\mathbf{f},\mathbf{v},\mathbf{h}) &= \sum_{\{i,j\}}(f(i,j)-d(i,j))^2 \\ &+\sum_{\{i,j\}}k(v(i,j))\left(\lambda^2 f_x^2(i,j)(1-v(i,j))+\alpha\, v(i,j)\right) \\ &+\sum_{\{i,j\}}k(h(i,j))\left(\lambda^2 f_y^2(i,j)(1-h(i,j))+\alpha\, h(i,j)\right) \end{aligned} \quad (3)$$

where

$$k(z) = \frac{c}{4\lambda^2 z(1-z)}\left[\sqrt{1+\frac{8\lambda^2 z(1-z)}{c}}-1\right],\ z\ \in\ [0,1] \quad (4)$$

Most of the material in this section is devoted to deriving this new energy function from the GNC energy function.

In the beginning ($c=0.25$), the lines are forced to the center of the interval $[0,1]$ ($z=0.5$) except for extremely low or high values of the intensity gradient. Then as the control parameter c is increased, this condition is slowly relaxed. For high values of c, we recover the weak membrane energy function. The only difference between the weak membrane and our function is $k(z)$ and the domain of z, which is $[0,1]$. In the weak membrane, the line processes could only take values in the set $\{0,1\}$.

We use the generic symbol z for the line process when distinctions between vertical and horizontal are unnecessary. The pixel location is suppressed whenever it is convenient.

One method of obtaining a fixed point for the Hamiltonian in (3) is gradient descent. This defines a dynamical system which is shown below.

$$\frac{d\mathbf{f}}{dt} = -\nabla_{\mathbf{f}}\mathcal{H}_a,\ \frac{d\mathbf{v}}{dt} = -\nabla_{\mathbf{v}}\mathcal{H}_a,\ \frac{d\mathbf{h}}{dt} = -\nabla_{\mathbf{h}}\mathcal{H}_a \quad (5)$$

where $\nabla_{\mathbf{z}}V$ is the vector equivalent of the scalar gradient $\frac{\partial V}{\partial z(i,j)}$. We have mentioned that z lies in the interval $[0,1]$. Unconstrained gradient descent may violate this constraint. A modified dynamical system which overcomes this problem is presented in [15]. The essential features of our approach do not change. We have chosen not to present the modified equations due to lack of space.

2.2 The adiabatic approximation

The system of equations in (5) defines a dynamical system. The variables in the dynamical system may possess vastly different

time constants. These can be roughly grouped into "fast" and "slow" modes. Typically, the fast modes quickly relax to their equilibrium state and are stable whereas the slow modes may be unstable. The simplest procedure of elimination of fast relaxing variables is the adiabatic approximation which is a well known technique in the field of synergetics.

Consider the dynamical system shown below, which is the 1-D version of (5). It can be obtained by applying gradient descent on the energy function $\mathcal{H}(\mathbf{f},\mathbf{z})$.

$$\frac{d\mathbf{f}}{dt} = -\nabla_{\mathbf{f}}\mathcal{H}(\mathbf{f},\mathbf{z}) \stackrel{\text{def}}{=} \mathbf{K}_1(\mathbf{f},\mathbf{z}),\ \frac{d\mathbf{z}}{dt} = -\nabla_{\mathbf{z}}\mathcal{H}(\mathbf{f},\mathbf{z}) \stackrel{\text{def}}{=} \mathbf{K}_2(\mathbf{f},\mathbf{z}) \tag{6}$$

We assume that the line processes can be eliminated adiabatically; i.e. they relax much faster than the intensity processes. At present, we only have empirical evidence in favor of this conjecture and are in the process of developing analytical support. The adiabatic technique is as follows [11].

1. Set the line processes to their equilibrium state, given the intensities. This eliminates the dynamics over the line processes since $\frac{d\mathbf{z}}{dt} = 0$.
$$\mathbf{K}_2(\mathbf{f},\mathbf{z}) = -\nabla_{\mathbf{z}}\mathcal{H}(\mathbf{f},\mathbf{z}) = 0. \tag{7}$$
2. Solve for $\mathbf{z}$ in (7) yielding $\mathbf{z} = \mathbf{z}(\mathbf{f})$.
3. Substitute $\mathbf{z} = \mathbf{z}(\mathbf{f})$ in (6) giving a new dynamical system which is a function of only the intensities.

When the solution for the line processes is substituted in (6), we get

$$\frac{d\mathbf{f}}{dt} = \mathbf{K}_1(\mathbf{f},\mathbf{z}(\mathbf{f})) \tag{8}$$

However, (8) can be obtained from a different energy function.

$$\frac{d\mathcal{H}(\mathbf{f},\mathbf{z}(\mathbf{f}))}{df(i)} = \frac{\partial\mathcal{H}(\mathbf{f},\mathbf{z})}{\partial f(i)}\Big|_{\mathbf{z}=\mathbf{z}(\mathbf{f})} + \sum_{i'}\left(\frac{\partial\mathcal{H}(\mathbf{f},\mathbf{z})}{\partial z(i)}\,\frac{\partial z(i)}{\partial f(i')}\right)\Big|_{z(i)=z(\mathbf{f})} \tag{9}$$

Since the adiabatic approximation implies $\nabla_{\mathbf{z}}\mathcal{H}(\mathbf{f},\mathbf{z})|_{\mathbf{z}=\mathbf{z}(\mathbf{f})} = 0$, the second term on the right hand side of (9) reduces to zero, leaving

$$\nabla_{\mathbf{f}}\mathcal{H}(\mathbf{f},\mathbf{z}(\mathbf{f})) = \nabla_{\mathbf{f}}\mathcal{H}(\mathbf{f},\mathbf{z})|_{\mathbf{z}=\mathbf{z}(\mathbf{f})} \tag{10}$$

We have shown that (8) can be directly obtained by gradient descent on the energy function $\mathcal{H}(\mathbf{f},\mathbf{z}(\mathbf{f}))$.

2.3 Relationship to the GNC algorithm

In this section, we show that our general Hamiltonian ($\mathcal{H}_a$) (3) can be derived from the GNC energy function by applying the adiabatic approximation.

The new energy function (3) presented in Section 2.1 has close ties to the GNC algorithm. To show this, we begin with an energy function that is a generalized version of (3).

$$\mathcal{H}_b(\mathbf{f},\mathbf{v},\mathbf{h}) = \sum_{\{i,j\}}(f(i,j)-d(i,j))^2 + \sum_{\{i,j\}}\left(\lambda^2 f_x^2(i,j)\; l_1(v(i,j)) + \alpha\; l_2(v(i,j))\right)$$
$$+\sum_{\{i,j\}}\left(\lambda^2 f_y^2(i,j)\; l_1(h(i,j)) + \alpha\; l_2(h(i,j))\right) \tag{11}$$

where $l_1(z)$ and $l_2(z)$ are unspecified functions at this point.

The intuition behind this structure can be readily seen. If $l_1(z) = (1-z)$ and $l_2(z) = z$, we get back our original energy function (2) without the interaction term. If $l_1(z) = k(z)(1-z)$ and $l_2(z) = k(z)z$, we obtain (3). Once again, a dynamical system can be obtained through gradient descent on (11).

$$\begin{aligned}
\frac{df(i,j)}{dt} &= -\frac{\partial\mathcal{H}_b}{\partial f(i,j)} = -2(f(i,j)-d(i,j)) \\
&\quad +2\lambda^2 f_x(i,j)\; l_1(v(i,j)) - 2\lambda^2 f_x(i-1,j)\; l_1(v(i-1,j)) \\
&\quad +2\lambda^2 f_y(i,j)\; l_1(h(i,j)) - 2\lambda^2 f_y(i,j-1)\; l_1(h(i,j-1)) \\
\frac{dv(i,j)}{dt} &= -\frac{\partial\mathcal{H}_b}{\partial v(i,j)} = -\lambda^2 f_x^2(i,j)\; l_1'(v(i,j)) - \alpha\; l_2'(v(i,j)) \\
\frac{dh(i,j)}{dt} &= -\frac{\partial\mathcal{H}_b}{\partial h(i,j)} = -\lambda^2 f_y^2(i,j)\; l_1'(h(i,j)) - \alpha\; l_2'(h(i,j))
\end{aligned} \tag{12}$$

where $l_{1(2)}'(z) = \frac{dl_{1(2)}(z)}{dz}$.

In the GNC setup, a sequence of energy functions can be constructed that start out convex and end as the weak membrane energy. Formally, in Blake and Zisserman's scheme, the final energy function and the sequence leading up to it can be written as follows.

$$\mathcal{H}_{GNC}^k(\mathbf{f}) = \sum_{\{i,j\}}(f(i,j)-d(i,j))^2 + \sum_{\{i,j\}}\left(g^{(k)}(f_x(i,j)) + g^{(k)}(f_y(i,j))\right),\ k = 0,1,.. \tag{13}$$

The sequence of energy functions has been translated into a sequence of $g(.)$ functions. In [7], the authors show that for an initial $g^*(.)$, the Hamiltonian is convex, thus providing a reasonable heuristic for this procedure. They derived a closed form expression for the $g(.)$ functions and this is reproduced below.

$$g^{(k)}(f_p) = \begin{cases} \lambda^2 f_p^2 & 0 \le |f_p| \le q \\ \alpha - \frac{c}{2}(r - |f_p|)^2 & q < |f_p| < r \\ \alpha & |f_p| > r \end{cases} \tag{14}$$

where $c = 2^k c^*$, $k = 0,1,..$, $r^2 = \alpha(\frac{2}{c} + \frac{1}{\lambda^2})$, $q = \frac{\alpha}{\lambda^2 r}$ and f_p is the generic symbol used for the intensity gradient in this paper. c^* is the initial value of the control parameter and for this value, $\mathcal{H}_{GNC}$ is convex. The other parameters q and r arise as a consequence of creating a convex Hamiltonian. They can be interpreted as thresholds on the line process. The line processes are undecided only in the interval $[q,r]$.

Once again, we can obtain a dynamical system through gradient descent on $\mathcal{H}_{GNC}$.

$$\begin{aligned}
\frac{df(i,j)}{dt} &= -\frac{\partial\mathcal{H}_{GNC}}{\partial f(i,j)} = -2(f(i,j)-d(i,j)) \\
&\quad +g^{(k)'}(f_x(i,j)) - g^{(k)'}(f_x(i-1,j)) \\
&\quad +g^{(k)'}(f_y(i,j)) - g^{(k)'}(f_y(i,j-1))
\end{aligned} \tag{15}$$

where $g^{(k)'}(f_p) = \frac{dg^{(k)}(f_p)}{df_p}$.

We wish to equate the two Hamiltonians, $\mathcal{H}_b$ and $\mathcal{H}_{GNC}$, after the line processes have been eliminated adiabatically. Concisely speaking,

$$\mathcal{H}_b(\mathbf{f}, \mathbf{v}(\mathbf{f}), \mathbf{h}(\mathbf{f})) = \mathcal{H}_{GNC}(\mathbf{f}) \tag{16}$$

The dynamical systems generated by gradient descent on the two Hamiltonians are then equated after applying the adiabatic approximation. These two conditions can then be used to solve for the unknown functions $l_1(z)$ and $l_2(z)$. We use (10), which allows us to equate the dynamical systems generated by $\mathcal{H}_b$ and by $\mathcal{H}_{GNC}$.

$$\frac{df(i,j)}{dt} = -\frac{\partial \mathcal{H}_b(\mathbf{f}, \mathbf{v}(\mathbf{f}), \mathbf{h}(\mathbf{f}))}{\partial f(i,j)} = -\frac{\partial \mathcal{H}_{GNC}(\mathbf{f})}{\partial f(i,j)} \tag{17}$$

This is a compact way of expressing equivalence between (12) and (15) with the time variation of the lines set to zero. We now make an assumption that simplifies this derivation immensely. The adiabatic condition theoretically makes the line process at each site a function of all the intensities. We restrict this dependence only to the intensity gradient. For the vertical (horizontal) line process, this assumption simplifies the relation $v(i,j) = v_{ij}(\mathbf{f})$ $(h(i,j) = h_{ij}(\mathbf{f}))$, to $v(i,j) = v(f_x(i,j))$ $(h(i,j) = h(f_y(i,j)))$. From (12) and (15), we see that (17) now reduces to

$$2\lambda^2 f_p \; l_1(z(f_p)) = g^{(k)'}(f_p) \tag{18}$$

From the GNC formulation (14), and [7], $g^{(k)'}(f_p)$ can be written as follows.

$$g^{(k)'}(f_p) = \begin{cases} 2\lambda^2 f_p & 0 \le f_p \le q \\ c(r - f_p) & q < f_p < r \\ 0 & f_p \ge r \\ -g^{(k)'}(-f_p) & f_p < 0 \end{cases} \tag{19}$$

$l_1(z(f_p))$ can be obtained from (18) and (19).

$$l_1(z(f_p)) = \begin{cases} 1 & 0 \le f_p \le q \\ \frac{c(r-f_p)}{2\lambda^2 f_p} & q < f_p < r \\ 0 & f_p \ge r \\ l_1(z(-f_p)) & f_p < 0 \end{cases} \tag{20}$$

We have obtained $l_1(z(f_p))$ by assuming a specific structure of the Hamiltonian $\mathcal{H}_b$. $l_2(z(f_p))$ can be obtained, once $l_1(z(f_p))$ is known since from (16), the Hamiltonians $\mathcal{H}_b$ and the GNC energy function are identical when the line processes are in their equilibrium state.

$$l_2(z(f_p)) = \begin{cases} 0 & 0 \le f_p \le q \\ \frac{c(f_p-q)}{2\lambda^2 q} & q < f_p < r \\ 1 & f_p \ge r \\ l_2(z(-f_p)) & f_p < 0 \end{cases} \tag{21}$$

After the line processes have been eliminated adiabatically, the new energy function is equal to the GNC function. The GNC energy function is obtained when the line processes are set to their equilibrium state in $\mathcal{H}_b$. An obvious step is to extend (20) and (21) when the line processes are in a non-equilibrium state. We can replace the two functions $l_1(z(f_p))$ and $l_2(z(f_p))$ by $l_1(z)$ and $l_2(z)$.

Since recovery of the actual functions $l_1(z)$ and $l_2(z)$ is complicated by the fact that only the functions $l_1(z(f_p))$ and $l_2(z(f_p))$ are known, we define a new process u and an unknown transformation (at this point) $z(u)$. Then we define $l_3(u) \stackrel{\text{def}}{=} l_1(z(u))$ and $l_4(u) \stackrel{\text{def}}{=} l_2(z(u))$. The equilibrium condition now demands, $l_{1(2)}(z(u)) = l_{1(2)}(z(f_p))$. The u process obviously has the flavor of an intensity gradient but arises through a transformation on the line process.

$$\begin{aligned} \mathcal{H}_b(\mathbf{f}, \mathbf{u}_v, \mathbf{u}_h) = & \sum_{\{i,j\}} (f(i,j) - d(i,j))^2 \\ & + \sum_{\{i,j\}} \left(\lambda^2 f_x^2(i,j) \; l_3(u_v(i,j)) + \alpha \; l_4(u_v(i,j)) \right) \\ & + \sum_{\{i,j\}} \left(\lambda^2 f_y^2(i,j) \; l_3(u_h(i,j)) + \alpha \; l_4(u_h(i,j)) \right) \end{aligned} \tag{22}$$

where u_v and u_h are the u processes in the vertical and horizontal directions, analogous to the line process.

Consider a likely candidate for $l_3(u)$ and $l_4(u)$ shown below.

$$l_3(u) = \frac{c(r-u)}{2\lambda^2 u}, \; q \le u \le r \tag{23}$$

and

$$l_4(u) = \frac{c(u-q)}{2\lambda^2 q}, \; q \le u \le r \tag{24}$$

The new process u is allowed to assume values only in the set $[q, r]$.

The only criterion for the choice of $z(u)$ is that $z = 0$ for $u = q$, and $z = 1$ for $u = r$. This is the original GNC condition, reformulated in terms of the new variable u. We can come up with many transformations that satisfy this condition. For the sake of symmetry, we add one more constraint; $z = 0.5$ when $\lambda^2 u^2 = \alpha$. One transformation obeying these conditions is $\frac{z}{1-z} = \frac{l_4(u)}{l_3(u)}$ which yields

$$z = \frac{u(u-q)}{u^2 - 2qu + qr} \tag{25}$$

Now we can rewrite the Hamiltonian $\mathcal{H}_b$ as an explicit function of the intensities and line processes by substituting z back into $\mathcal{H}_b$. The inverse transformation for this purpose is

$$u = u(z) = q \left[\frac{(1-2z) + \sqrt{1 + \frac{8\lambda^2 z(1-z)}{c}}}{2(1-z)} \right] \tag{26}$$

Substituting this back in (23), (24), we get (3). We have shown that starting from $\mathcal{H}_b$ and $\mathcal{H}_{GNC}$, we can obtain $\mathcal{H}_a$ by invoking the adiabatic approximation. This clearly indicates that $\mathcal{H}_a$ can be viewed as the general energy function with the GNC function as a special case. The purpose behind deriving $\mathcal{H}_a$ is to be able to add further constraints on the line process.

3 The G^{2}NC Algorithm

3.1 Incorporation of line interactions

The new Hamiltonian $(\mathcal{H}_a)$ (3) can now be used as a springboard for introducing a more sophisticated line process. Several problems remain. Introducing line process interactions is complicated by the multiplying factor $k(z)$. If interactions are introduced in the traditional way, gradient descent on the lines would become necessary. We have tried to avoid this due to the different time constants involved and because of the complex nature of $k(z)$. In order to overcome this problem, we examine the Hamiltonian as a function of the intensities and $u(i,j)$ instead of the line process.

$$\mathcal{H}_a(\mathbf{f}, \mathbf{u}_v, \mathbf{u}_h) = \sum_{\{i,j\}} (f(i,j) - d(i,j))^2$$

$$+\frac{c}{2}\sum_{\{i,j\}}\left(f_x^2(i,j)\frac{(r-u_v(i,j))}{u_v(i,j)}+r(u_v(i,j)-q)\right)$$
$$+\frac{c}{2}\sum_{\{i,j\}}\left(f_y^2(i,j)\frac{(r-u_h(i,j))}{u_h(i,j)}+r(u_h(i,j)-q)\right) \quad (27)$$

Our principal contribution in this paper is to suggest introducing a new process which has a monotonic transformation to the lines, namely the u process that we call the gradient-magnitude (GMAG) process. All line interactions can be transferred to the GMAG domain.

We have chosen the most basic kind of interaction. This has been inspired by the work of Geiger and Girosi [4] who have clearly shown the effect of this interaction using mean field theory. Consider the following modification of $\mathcal{H}_a$;

$$\mathcal{H}_{G^2NC}(\mathbf{f},\mathbf{u}_v,\mathbf{u}_h) = \mathcal{H}_a(\mathbf{f},\mathbf{u}_v,\mathbf{u}_h) - \frac{c}{2}\epsilon\sum_{\{i,j\}}(u_v(i,j)-q)(u_v(i,j+1)-q) - \frac{c}{2}\epsilon\sum_{\{i,j\}}(u_h(i,j)-q)(u_h(i+1,j)-q) \quad (28)$$

where all the indices are analogous to the earlier line process case. Physically, we are trying to decrease the penalty on a vertical (horizontal) line if its vertical (horizontal) neighbors are "on". The penalty is controlled by the parameter ϵ.

We now proceed with the algorithm itself. The improved line process in the GNC set up prompted us to call this algorithm the Generalized GNC or G^{2}NC algorithm. We first solve for u keeping the intensities fixed. Now, u becomes a function of the neighbors as well as the intensity gradient. We choose to run Besag's ICM algorithm [14] on the GMAG process which essentially involves minimizing the energy function with respect to the GMAG process at each point keeping all other processes fixed. This process is repeated until the GMAG processes have converged.

In order to run ICM on the GMAG process, we have to solve for u when interactions are present. The criterion for a minimum is :

$$\frac{\partial\mathcal{H}_{G^2NC}}{\partial u}=0 \quad (29)$$

From this we get

$$u=\frac{|f_p|}{\sqrt{1-\frac{2\epsilon(u_{av}-q)}{r}}} \quad (30)$$

where u_{av} is $\frac{u_v(i,j+1)+u_v(i,j-1)}{2}$ when the minimum w.r.t $u_v(i,j)$ is sought. A similar relationship exists for $u_h(i,j)$.

This solution for u is not guaranteed to be within $[q,r]$. If the solution lies outside, the end point energies are again compared and the solution with the lower energy is chosen. The solution depends heavily on ϵ.

A natural choice of ϵ can now be made. We would like the influence of the interaction term to be significant when the neighbors are "decided" (neighboring u's equal r) and the current u equals q. From the solution for u, it can be seen that when $f_p=q$, $u=r$ if $\epsilon=\frac{(r+q)}{2r}$. This result is very pleasing since ϵ roughly varies from 0.5 to 1. The sweep of the parameter ϵ is satisfying, as no estimation is necessary. The range of interactions can be extended well beyond this initial scheme.

3.2 Description of the algorithm

We have chosen to run the Conjugate Gradient (CG) algorithm on the intensities. We wait until the intensities have converged before updating the GMAG processes. This follows naturally from our intuition. The intensities relax more slowly than the lines. We have noticed that ICM takes very few (one to five) iterations to converge (for our case studies).

- The algorithm:
  1. Set $c=c^*$ (usually $c^*=0.25$)
  2. Run the CG algorithm on the intensities until convergence
  3. Update GMAG processes at the end of the CG iterations using Besag's ICM algorithm till they converge
  4. Return to Step 2 until intensities and GMAG processes converge
  5. Increase c, usually $c=2^kc^*, k=0,1,..$
  6. Return to Step 2 until the final $c=cstop$ is reached

The algorithm is guaranteed to converge for each value of k.

4 Experimental Results

We executed the G^{2}NC algorithm on two noisy, aerial images. The two aerial images are an airport and Mission Bay (San Diego).

The original airport image is shown in Figure 1a and the corresponding noisy image (5dB, $\sigma^2=147$) in Figure 1b. The estimated intensity and line process images are displayed in Figure 2a and 2b. We experimented with a few heuristics for the parameter settings [15].

We see from Figures 1 and 2 that all the "relevant" details are captured by the algorithm. The line process image contains all the important edges without arbitrary gaps in contours. The estimated intensity image clearly exhibits the effect of feedback from the discontinuities. The improvement in the line process prior is showcased in Figures 3, 4 and 5. Figures 3a and 3b contain the aerial view of Mission Bay and the corresponding degraded image. Figures 4 and 5 are obtained from the GNC and the G^{2}NC algorithms with the same parameters. The line process image using G^{2}NC is vastly improved. Better results can be obtained for the GNC case when the parameters are chosen differently. Our view is that, introducing interactions makes this technique more "robust" with respect to the parameters.

5 Conclusions

We highlight below some of the new ideas and results in this paper. We have

- Presented a new energy function for MAP estimation
- Shown the importance of the adiabatic approximation in deriving this new energy function from the GNC energy function.

- Transferred line interactions to the GMAG domain
- Introduced a transformation from this process to the line process
- Set up an interaction term and shown how to vary the interaction parameter ϵ
- Suggested an algorithm to minimize the new energy function using a combination of CG and ICM

Acknowledgments

We wish to thank B. S. Manjunath, Dr. Josiane Zerubia, Ze'ev Lichtenstein, Dr. Tal Simchony and Doug Wiley for their many insightful comments and criticisms.

References

[1] S. Geman and D. Geman, "Stochastic relaxation, Gibbs Distributions and the Bayesian restoration of Images", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741, November 1984.

[2] F. C. Jeng and J. W. Woods, "Image estimation by stochastic relaxation in the compound Gaussian case", *Proc. ICASSP '88, IEEE Conf. on Acoust., Speech, and Signal Proc.*, April 1988.

[3] C. Koch, J. L. Marroquin, and A. L. Yuille, "Analog neuronal networks in early vision", *Proc. Natl. Acad. Sci.*, pp. 4263–4267, 1986.

[4] D. Geiger and F. Girosi, "Parallel and deterministic algorithms for MRFs: surface reconstruction and integration", Technical Report A. I. Memo, No. 1114, Artificial Intelligence Lab, M. I. T, June 1989.

[5] D. Terzopoulos, "The Computation of Visible-Surface Representations", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 10, pp. 417–437, July 1988.

[6] J. Zerubia and R. Chellappa, "Mean Field approximation using Compound Gauss-Markov Random Field for edge detection and image estimation", In *Proc. ICASSP '90, IEEE Conf. on Acoust., Speech, and Signal Processing*, April 1990.

[7] A. Blake and A. Zisserman, *"Visual Reconstruction"*, M. I. T. Press, Cambridge, 1987.

[8] T. Simchony, R. Chellappa, and Z. Lichtenstein, "Graduated Non-Convexity Algorithm for image estimation using Compound Gauss-Markov Random fields", *Proc. ICASSP '89, IEEE Conf. on Acoust., Speech, and Signal Proc.*, May 1989.

[9] A. L. Yuille, "Energy functions for Early Vision and Analog Networks", *Biological Cybernetics*, vol. 61, pp. 115–123, 1989.

[10] Y. G. Leclerc, "Constructing Simple Stable Descriptions for Image Partitioning", *International Journal of Computer Vision*, vol. 3, pp. 73–102, 1989.

[11] H. Haken, *"Synergetics-An Introduction"*, Springer-Verlag, New York, 3rd. edition, 1983.

[12] H. Haken, *"Advanced Synergetics"*, Springer-Verlag, New York, 1983.

[13] D. Luenberger, *"Linear and Nonlinear Programming"*, Addison-Wesley Inc., Menlo Park, CA, 1984.

[14] J. Besag, "On the statistical analysis of dirty pictures", *Journal of the Royal Statistical Society B*, vol. 48(3), pp. 259–302, 1986.

[15] A. Rangarajan and R. Chellappa, "Generalized Graduated Non-Convexity Algorithm for Maximum A Posteriori Image Estimation", Technical Report USC-SIPI, No. 149, Univ. of South. Calif., December 1989.

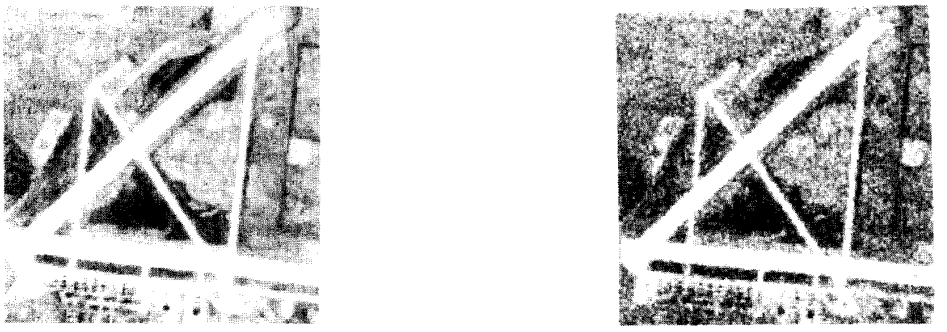

Figure 1: (a) Left: Original Airport (b) Right: Noisy Airport

Figure 2: (a) Left: Restored Image using G^{2}NC algorithm (b) Right: Line process image using G^{2}NC algorithm

Figure 3: (a) Left: Original Mission Bay (b) Right: Noisy Mission Bay

Figure 4: (a) Left: Restored image using the GNC algorithm (b) Right: Restored image using G^{2}NC algorithm

Figure 5: (a) Left: Line process image using the GNC algorithm (b) Right: Line process image using G^{2}NC algorithm

[1]Partially supported by the Joint Services electronics program under the contract F49620-88-C-0067 and by the NSF grant MIP-84-51010