

Validating a Model of Colon Colouration Using an Evolution Strategy with Adaptive Approximations

Džena Hidović and Jonathan E. Rowe

School of Computer Science, The University of Birmingham, Birmingham B15 2TT, UK
{D.Hidovic, J.E.Rowe}@cs.bham.ac.uk

Abstract. The colour of colon tissue, which depends on the tissue structure, its optical properties, and the quantities of the pigments present in it, can be predicted by a physics-based model of colouration. The model, created by analysing light interaction with the tissue, is aimed at correlating the histology of the colon and its colours. This could be of a great diagnostic value, as the development of tissue abnormalities and malignancies is characterised by the rearrangement of underlying histology. Once developed, the model has to be validated for correctness. The validation has been implemented as an optimisation problem, and evolutionary techniques have been applied to solve it. An adaptive approximate optimisation method has been developed and applied in order to speed up the computationally expensive optimisation process. This works by iteratively improving a surrogate model based on an approximate physical theory of light propagation (Kubelka Munk). Good fittings, obtained under the histologically plausible values of model parameters, are presented. The performances of the new method were compared to that of a simple Evolution Strategy which uses an accurate, but expensive, Monte Carlo method. The new method is general and can be applied with any surrogate model for optimisation.

1 Introduction

The colour of tissue is a result of light interacting with the tissue while propagating through it. It therefore directly depends on tissue architecture, structure and composition. Many pathological changes are characterised by rearrangements of tissue structure and composition, which implies that also the colour of normal and abnormal tissue differ from each other. However, the changes in colour may often be very subtle and invisible to the human eye. To overcome this problem, a novel approach to interpreting the medical images of tissue has been developed and successfully implemented on the skin [1,2]. Our current work concentrates on extending the application area of that approach to the colon.

The basic idea of the research lies in understanding the physics of colour image formation: that is, the interaction of incident light with the tissue. White light interacts with the tissue structure and composition and penetrates it to a certain depth. Some of the light is absorbed and scattered forward by particles and macromolecules inside the tissue, and some remitted at its surface. The remitted portion of light gives tissue its colour. The fractions of the light that are remitted, scattered and absorbed depend on the optical properties of the tissue itself, which in turn directly depend on its structure and

the pigments present in it. By simulating the processes of light interaction with tissue for all plausible normal variations of its histological parameters, it is possible to predict all its possible colours. This means that a tissue which has a normal architecture will have colour that can be predicted by the simulations. Any deviation from these predicted colours is a sign of abnormality, since differences in the colour between normal and abnormal tissue are due to the rearrangements of its architecture and composition, as argued by many clinicians and physicians.

The key concept in interpreting the images of tissue in terms of its histology is a *model of colouration* which is a set of correspondences between the parameters characterising the tissue and its colours. The model is constructed by taking the optical properties of tissue and calculating the corresponding spectral reflectances by mathematical models of light propagation. Spectra are then convolved with the curves of the response system and colours, typically RGB, are obtained. The model can then be used to perform the inversion process, i.e. to infer the combination of histological parameters which lead to a particular colour. The model of colon colouration which predicts the light reflected back from the colon tissue, and its parameters, will be briefly described in the next section.

Once developed, the model has to be validated for correctness. Validation, described in more detail in section 3, is normally done by comparing the light reflected from the tissue (spectral reflectance) predicted by our theoretical models, and the measured spectral reflectance from the clinical experiments. This problem is implemented as an optimisation problem with the goal of minimising the distance between the measured and predicted spectra. To solve the optimisation problem, evolutionary algorithms have been used. The high cost of calculating the spectral reflectance accurately resulted in the development of an adaptive approximate optimisation method. The basic idea lies in creating a surrogate model, and iteratively correcting the error between the accurate and approximate models which solve the original problem. More detailed description of the method and the motivation for its development are given in section 4. As shown in section 5, the new method gives satisfactory results in significantly less time than that required using only the accurate but computationally expensive model of light transport in tissue. A discussion of the results and some limitations of the method will conclude the paper.

2 Background: Model of Colon Tissue Colouration

Colon tissue is a layered structure (figure 1) composed of four layers, characterised by different optical properties. Starting from the innermost layer, sequentially they are: mucosa, submucosa, muscularis externa propria and serosa.

White light penetrating into the colon tissue is mainly scattered by a network of collagen fibres whose different sizes in colon layers, imply different scattering properties. Most of the light remitted at the tissue surface is the result of backward light scattering which occurs in the mucosal layer. Besides being scattered, the incident light gets absorbed on its way inside the colon tissue. The absorption is mainly due to oxy and deoxy hemoglobin which form the main part of the red blood cells. Light transmitted through the muscle layer forms just a small fraction of the incident light and

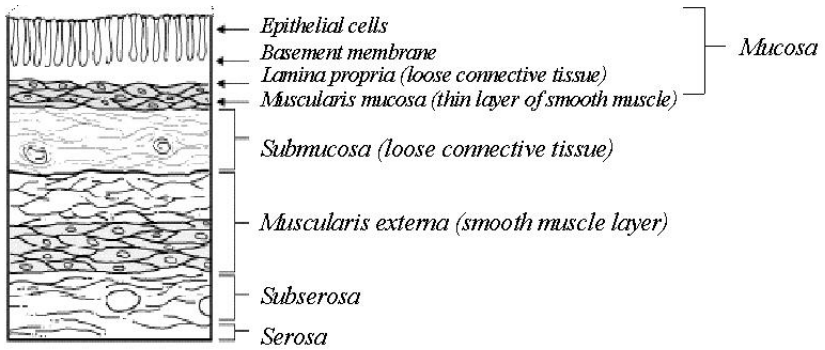


Fig. 1. Colon architecture

is strongly forward directed. Hence, it does not influence the light remitted from the colon. Therefore, our model simulates the light interaction with only the first three layers.

For calculating the spectra of the colon we use the Monte Carlo method [3,4], which is considered to give an accurate solution to the radiative transfer equation, describing light transport in tissue. Given the information about the absorption and scattering coefficients of each of the layers composing the model, their respective thicknesses, refractive indexes and anisotropy factors, the Monte Carlo method calculates tissue spectral reflectance, by statistically analysing the interaction of photons with the tissue. Each photon is traced independently on its way through the tissue, until it is either completely absorbed by the pigments of tissue; transmitted into deeper layers; or it escapes at the surface. In order to get accurate analysis, a large number of photons, usually $10^5 - 10^6$, have to be simulated. That gives a huge computational complexity to the Monte Carlo method.

In order to apply the above method to calculating the light remitted from the colon, scattering and absorption coefficients, and thicknesses of the first three layers of the colon tissue have to be known. Given that the previous literature reports little on the quantitative optical properties of the colon, we have modelled scattering and absorption properties ourselves. Scattering coefficients were calculated using Mie theory [5] starting from collagen fibre size and density. Absorption coefficients on the other hand are expressed as a function of haemoglobin concentration and saturation of each layer.

The need to calculate the optical coefficients prior to calculating the spectral reflectance, resulted in the following parameters of our model:

- *haemoglobin concentration*, i.e. the amount of hemoglobin per unit volume of tissue. This parameter can be expressed as a product of hemoglobin concentration per unit volume of blood and the volume fraction of blood in a particular tissue.
- *size of collagen fibres*, i.e. diameter of collagen fibres.
- *density of collagen fibres*, i.e. number of collagen fibres per unit volume of tissue.
- *thickness of tissue layer*

Each of the above parameters was defined for both mucosa and submucosa resulting in eight model parameters. For each parameter a range of all values corresponding to

normal colon tissue was defined. All the ranges have been confirmed as histologically plausible by a pathologist with a special interest in colon tissue. Optical properties of muscularis externa were fixed and hence no additional parameters regarding that layer were considered. There were therefore a total of eight model parameters.

3 Model Validation

The main objective of the validation process was to verify the correctness of the model of colon colouration. This has been done by comparing spectra computed from the model by Monte Carlo simulations with spectra measured from the real tissue. If the model is correct, then it must be able to generate spectra which approximate well the measured spectral reflectances. However, this is only the first step, as the inverse is not necessarily true.

The simulated spectra were compared against the spectra obtained by diffuse reflectance spectroscopy *in vivo* [6], where a bundle of fibre optics which deliver and collect the incident and reflected light respectively is passed through a working channel of an endoscope and placed in contact with the colon wall of the patients during ordinary colonoscopy procedures. In particular, 84 spectra of normal colon, kindly provided by Kevin Schomacker, have been used in the validation. No spectra of abnormal colon were used, as the model developed so far simulates the light interaction with only normal tissue. The measured spectra was recorded every 2 nm in the range from 300 nm to 800 nm. However, given the huge computational complexity of Monte Carlo simulations, which is proportional to the number of wavelengths at which the reflected spectrum is calculated, the measured and simulated spectra were compared at only the following set of 17 wavelengths (chosen to give a sufficiently accurate characterisation of the spectra): {450, 480, 506, 514, 522, 540, 548, 560, 564, 568, 574, 586, 594, 610, 640, 676}. Most wavelengths were selected in the green region of the visible light, because that part of the spectra is heavily changed by the light absorption of haemoglobin.

The validation was implemented as an optimisation problem, with the aim of minimising the distance between the simulated and measured spectra. The distance between the two curves was calculated using the following measure of distance between two sets of points:

$$d(y, z) = \frac{1}{m} \sum_{i=1}^m |y_i - z_i|$$

where y_i is a value of the spectral reflectance measured on real colon tissue *in vivo* at the wavelength w_i , z_i is a value of reflectance simulated using our model of colon colouration (starting from a particular set of parameter values) at wavelength w_i , and m is the total number of wavelengths (in our case $m = 17$).

In addition to the eight model parameters described in the previous section, a scale factor was introduced to account for the adjustments of the normalisation process, where the measured reflectance spectra were normalised to a spectrum remitted from a reflectance standard. This is a standard procedure in experimental spectroscopy to account

for the characteristics of the illumination used in the measurements. There were therefore nine parameters to be optimised.

Originally, a novel (1+1)-ES was used to solve to optimisation problem [7]. At each iteration, a set of parameters was chosen as described in [7], and the corresponding reflectance spectra was calculated. The high computational complexity of the Monte Carlo method (in our simulations 60000 photons were used), resulted in an optimisation procedure that was very slow. The execution time of each Monte Carlo simulation was typically several minutes. For (1+1)-ES to find a good approximation, typically a few hundred of iterations had to be executed. Consequently, we developed a method that makes use of a fast approximate model of light transport, as explained in the next section.

4 Approximate Models

The problem of model validation, from an optimisation point of view, may be summarised as follows. We have a set of parameters, each lying in some bounded interval, describing a physical model. For any particular choice of parameter values, we can calculate the resulting optical spectrum using an expensive Monte Carlo algorithm. Given a particular spectrum (measured during colonoscopy), we need to find the associated parameter values. The fitness function is the distance between the spectrum derived from the Monte Carlo algorithm and the target spectrum (as described above).

This situation, in which the fitness evaluation is computationally expensive, is quite common in real-world optimisation problems. A typical strategy for dealing with this situation is to use approximate models, or *surrogates*, that are faster to evaluate. These may either come from fitting a model to sampled data, or may have derived from a theoretical analysis of the problem. In our case, the Kubelka Munk theory [8,9], which is an approximate analytical solution to the radiative transfer equation, provides a fast approximate method. The Kubelka Munk method describes the two flux theory of light transport in a seminfinte slab of particular thickness. Propagation of light is limited to being only directly forward and backward oriented, which makes this method essentially a one dimensional approximation.

The derivation of surrogates and their management within the optimisation process has been considered by a number of authors (for example, see [10] for a useful survey). If the surrogate comes from fitting a model to sample data (e.g. by training a neural network [11], or fitting a smooth function over a mesh of points [12]), then it can be periodically improved by incorporating more data points as the search progresses. The idea is that one interleaves periods of search with updates to the surrogate (e.g. by retraining the neural network on the new data), in such a way that the surrogate becomes more accurate as the search converges to an optimum.

In our situation, we have a surrogate model derived from an approximate physical theory of light transport. It does not make sense to “re-train” this model on the basis of sampled data. However, we can build a model of the *error* between the Kubelka Munk method and Monte Carlo. One could, perhaps, train a neural network to model this error, but we propose a much simpler and faster method.

The overall management scheme is as follows:

1. Initially, set the error model to be zero.
2. Optimise using Kubelka Munk, plus the error model.
3. Run Monte Carlo on some points near the “optimum” found in step 2.
4. Use these points to update the error model.
5. Go to step 2.

This cycle is repeated as many times as necessary (or can be afforded). Note that each cycle requires the optimisation of an approximate model (at step 2). We use a novel (1+1)-ES to perform this optimisation, which it does efficiently and accurately. This Evolution Strategy is based on a new mutation probability distribution and is described in detail elsewhere [7]. In this paper, we concentrate on the method for updating the error model, which we describe in the next section. At each cycle, we produce an improved surrogate model, that estimates the error function within the vicinity of the optimum point found using the surrogate. When a new cycle begins, with a new surrogate, we start the next run of the (1+1)-ES from that point, ensuring that the search takes place that is both near the true optimum, and in a region for which the new surrogate is an accurate model.

5 The Error Correction Algorithm

Let us re-state the problem in more general terms. We have n parameters and a genotype-phenotype map $M : \mathbb{R}^n \rightarrow \mathbb{R}^m$. There is a function $\phi : \mathbb{R}^m \rightarrow \mathbb{R}$ which evaluates the phenotype. The fitness function is given by $\phi \circ M$. The map M is computationally expensive. We also have an alternative map $K : \mathbb{R}^n \rightarrow \mathbb{R}^m$ which is an approximation to M . The map K is computationally efficient.

In our example, we have $n = 9$ parameters, and M is the Monte Carlo algorithm for generating the corresponding spectra evaluated at $m = 17$ wavelengths. ϕ is the distance function $\phi(s) = d(s, \tau)$, where s is the simulated spectrum, and τ is the target spectrum. K is the Kubelka Munk method.

Define the *error function* $E : \mathbb{R}^n \rightarrow \mathbb{R}^m$ to be $E(x) = M(x) - K(x)$. This is the function we will try to model more and more accurately at each cycle. Let us suppose that at cycle t the approximate error model is E_t . Initially, we have $E_0(x) = 0$ for all x . The surrogate that is used in cycle t is $S_t = K + E_t$. We suppose that running the optimisation algorithm with surrogate S_t has produced the parameter vector $z \in \mathbb{R}^n$. This means that $S_t(z) = K(z) + E_t(z)$ is very close to the target spectrum τ .

We seek to construct our next approximation to the error function, E_{t+1} . We will approximate the error function E at z by estimating its differential dE at z . The differential is the best linear approximation to the function at a given point. So we wish to find a matrix A_{t+1} which is an estimate of dE at z , and then set

$$E_{t+1}(x) = A_{t+1}(x - z) + E(z)$$

which we can see agrees exactly with E at the point z . Define a vector

$$b_{t+1} = E(z) - A_{t+1}(z)$$

then our new error model is

$$E_{t+1}(x) = A_{t+1}(x) + b_{t+1}$$

The problem is to find the matrix A_{t+1} and, from there, the vector b_{t+1} .

Notice that $E(z)$ is the difference between the accurate model M and the initial surrogate $S_0 = K$. We would like to incrementally improve our approximation in terms of the behaviour of the *current* surrogate S_t . We can do this as follows. Define a *surrogate error function*

$$F_t(x) = M(x) - S_t(x) = M(x) - K(x) - E_t(x) = E(x) - E_t(x)$$

This implies that $E(x) = F_t(x) + E_t(x)$. We now find an approximation to $F_t(x)$ of the form $B_t(x - z) + F_t(z)$, where B_t is an estimate of the differential of F_t at z . Then we set

$$\begin{aligned} E_{t+1}(x) &= B_t(x - z) + F_t(z) + E_t(x) \\ &= B_t(x) - B_t(z) + F_t(z) + A_t(x) + b_t \\ &= (A_t + B_t)(x) + b_t + F_t(z) - B_t(z) \end{aligned}$$

This equation gives us an algorithm for updating the surrogate model at each cycle.

1. Initially set $A_0 = 0$ and $b_0 = 0$.
2. At cycle t get the error, $F_t(z)$, between the accurate model M and the current surrogate S_t at the point z found by the optimisation procedure.
3. Estimate the differential of F_t at the point z by the matrix B_t .
4. Set $A_{t+1} = A_t + B_t$
5. Set $b_{t+1} = b_t + F_t(z) - B_t(z)$

The surrogate model at each cycle is $S_t(x) = K(x) + E_t(x) = K(x) + A_t(x) + b_t$.

It remains to show how to calculate B_t . Let e_k be the vector with zeros everywhere except at position k , where there is a one. Then $B_t(e_k)$ tells us the k th column of the matrix B_t . To find this, we simply make the approximation to F_t agree with F_t at a small perturbation $z + \delta e_k$ from z (where δ is small compared to the defining bounds of parameter k). That is, we require

$$F_t(z + \delta e_k) = B_t(z + \delta e_k - z) + F_t(z)$$

Therefore

$$\begin{aligned} B_t(e_k) &= \frac{1}{\delta}(F_t(z + \delta e_k) - F_t(z)) \\ &= \frac{1}{\delta}(M(z + \delta e_k) - S_t(z + \delta e_k) - F_t(z)) \end{aligned}$$

which we can calculate by running the accurate model M and the surrogate S_t on the perturbed vector $z + \delta e_k$. We repeat this for each parameter k to get the whole of the matrix B_t . We see that each cycle therefore requires $n + 1$ runs of the accurate model M .

6 Experimental Results

We tested the adaptive error correction algorithm on a sample of 20 spectra drawn randomly from our collection of normal colon tissue images. Since running the Monte Carlo algorithm takes several minutes, we allowed a maximum of 200 runs of this algorithm for each spectrum. Since there are 9 parameters, each cycle of the error correction algorithm requires 10 runs of Monte Carlo. We therefore ran 20 cycles of the algorithm. The optimisation of the surrogate at each cycle was done with our (1+1)-ES with novel mutation probability distribution [7]. This was allowed 2000 evaluations of the surrogate at each cycle. This was usually sufficient to optimise the surrogate within a small tolerance.

In comparison, we also tried to solve the optimisation problem by just using the (1+1)-ES with 200 function evaluations using Monte Carlo. Six of the results are illustrated in figure 2 which shows the original spectra, and the best achieved by the error correction algorithm and the (1+1)-ES (by itself). We also compared the two algorithms after 50, 100, 150 and 200 runs of Monte Carlo (that is, after 5, 10, 15 and 20 cycles of the error correction algorithm). The results are presented in table 1. These show clearly that the new algorithm significantly outperforms the plain (1+1)-ES. It also appears to be more robust, in that the results show a smaller standard deviation for the new algorithm.

Table 1. Results of running the error correction algorithm versus (1+1)-ES on 20 spectra, after 50, 100, 150 and 200 iterations of the Monte Carlo algorithm. The average fitness (distance from target spectrum) is shown with standard deviation in brackets. Significance is measured by a paired t-test.

Iterations	Error Correction algorithm	(1+1)-ES	Significance
50	0.0176 (0.0102)	0.0269 (0.0219)	89.13%
100	0.0130 (0.0063)	0.0183 (0.0099)	97.74%
150	0.0111 (0.0055)	0.0161 (0.0080)	99.07%
200	0.0096 (0.0044)	0.0151 (0.0075)	99.93%

7 Discussion

The first observation to make about the experimental results, is that they confirm that the physical model of the colon tissue can account for the spectra observed in normal tissue. In addition to the 20 spectra from the sample data set used in this paper, further testing on the remaining 52 spectra confirm these findings. Good fittings obtained under the assumption of histological plausibility of the model parameters, which was confirmed by a pathologist with special interest in colon tissue, are a step forward in proving the correctness of the model. However, the validation process was done only on the normal spectra as the model developed so far predicts the light interaction with the normal colon

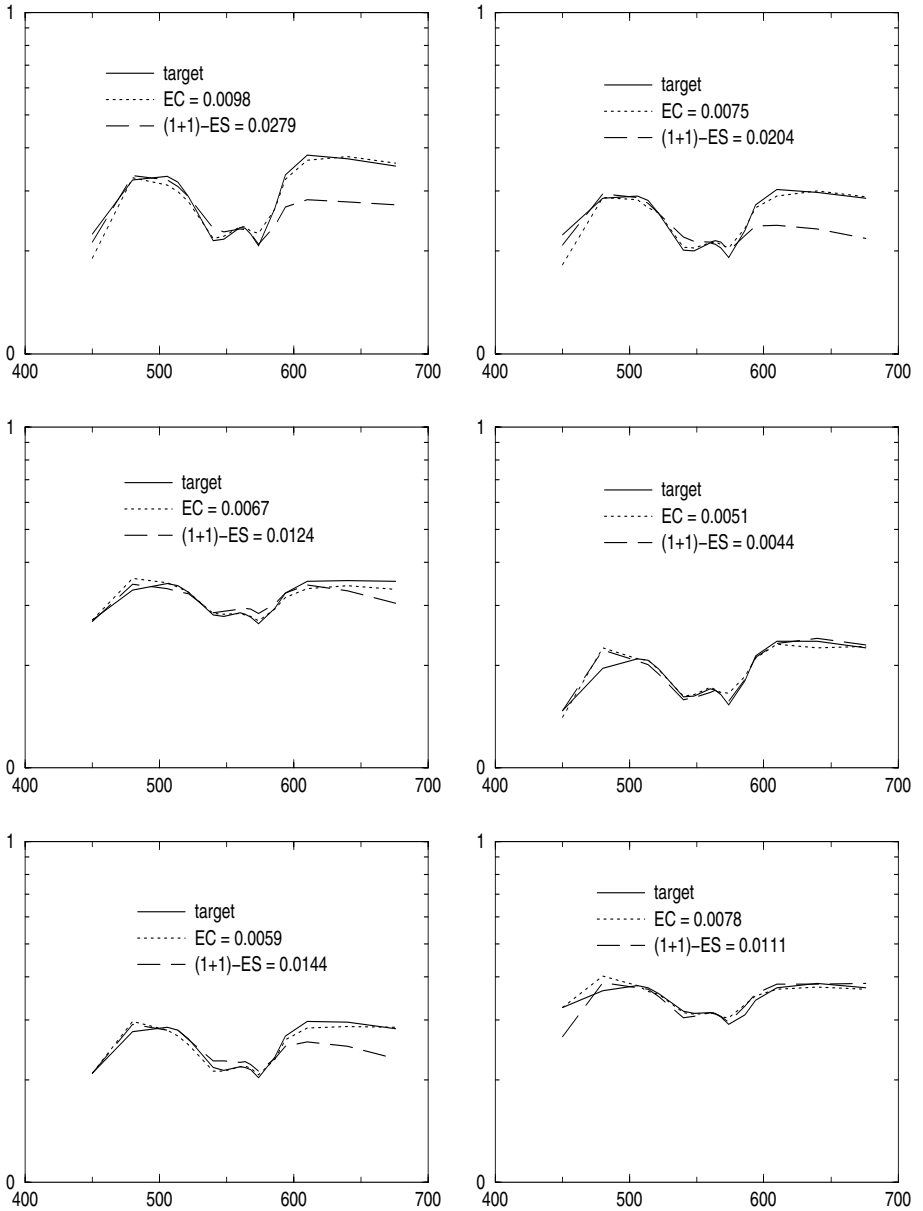


Fig. 2. Six examples of evolved spectra using the error-correcting algorithm (EC) and plain (1+1)-ES. Each algorithm uses the Monte Carlo algorithm 200 times for each spectrum. The final fitness is shown in each case, given by the distance to the target spectrum (solid line)

tissue. Further work is being done to analyse the effects of abnormal tissue structure (e.g. cancer). First findings suggest that due to the structural and compositional differences between the two types of tissue, the spectra of abnormal colon differs from that of the normal tissue. This is promising that our approach could provide diagnostically significant information. Detailed studies, however, have still to be done.

The second observation is that the error-correcting algorithm is generally more efficient at finding the optima than the plain (1+1)-ES, although this algorithm is itself very good (getting even close to the optimum in only 200 function evaluations is no mean feat!). The fact that the (1+1)-ES is a good algorithm directly helps the error-correcting algorithm. This is because we can be fairly sure that the *surrogate* optimisation problem can be solved efficiently at each cycle. Clearly, if we could not find the surrogate optimum, then we would not be looking in the right area to correct the surrogate.

This leads on to another observation about how the error-correcting algorithm is working. We need to be confident that the optima that are found for the surrogate are reasonably close to the optimum of the accurate model. At least, the closer it is, the fewer cycles will be necessary to find it. And obviously, if the surrogate is completely wrong, then it will be misleading rather than helpful to use it. In our case, the initial surrogate, the Kubelka Munk model, is actually not a great approximation. Apart from the physical assumptions it makes about the way light scatters in a tissue, it also calculates the total diffuse reflectance from the surface. What we actually need is the amount of reflectance within the given area of the collection probe, which is considerably smaller (especially at the red end of the spectrum). Nevertheless, after 20 cycles, the corrected surrogate is acting as a very good approximation in the vicinity of the optimum — which is also where the search is taking place.

A more technical issue is the question of how accurately we can model the error function $E(x) = M(x) - K(x)$ by an affine function. This is equivalent to asking how good an approximation is the differential of E to E itself at a given point. We know that (assuming E is differentiable) it is arbitrarily good within a sufficiently small radius of the point. So if the error function varies very slowly, then we should be able to get a good approximation. If it varies very rapidly (or indeed is not differentiable) then we could have a problem. But note that, if the surrogate has an optimum near the genuine optimum, we only require that the error function varies smoothly in that region of the search space.

A related issue is: even if the differential actually does give a good approximation to the error function, how good is our estimate of the differential? Recall that we estimate this by sampling in each parameter direction and then making our estimate agree with the error function at each of the sampled points. This is determined by the choice of the parameter δ . If this is too large (compared to the rate at which the error function is varying) then the estimate will be poor. On the other hand it can't be made arbitrarily small — we need to pick up the variation of $E(X)$ in the vicinity of z to a significant degree. In our experiments we scaled all our parameters to lie in the range $[-1, +1]$ and took $\delta = 0.025$. Other values also work. It is possible that this parameter could be adapted in response to the size of the error found, but we have not investigated this.

8 Conclusions

In this paper, we have presented a method for validating a model of colon colouration. The model allows us to calculate the spectral reflectance of colon tissue for a range of histological parameters using Monte Carlo simulation. The validation process was formulated as an optimisation problem. Given that the Monte Carlo method is an expensive (yet accurate) method of light propagation in tissue, the optimisation procedure was very slow. In order to speed up the optimisation, we have developed a method using a surrogate approximate model, based on Kubelka Munk theory. Our method is adaptive: the error between the surrogate and the accurate model is estimated and refined at each cycle. This leads to a more efficient use of the accurate model and, in general, leads to better results, faster. The method is general, and can be used with any surrogate model for optimisation.

The validation process of the model of colon colouration has shown that this model is able to generate the spectral reflectance of colon tissue. Moreover, the ranges of its parameters have been confirmed as histologically plausible, which means that the model can be used to predict the colours occurring in the normal colon tissue.

The next step in our approach is the inversion of the model described above in order to derive histological parameters for given tissue colours. The inverse mapping will then be used to create parametric maps, one for each parameter, which, at each pixel, show the magnitude of the relative histological parameter.

Acknowledgements. Jonathan Rowe did some of the work while visiting Darrell Whitley, funded in part from National Science Foundation grant number IIS-0117209. The colonoscopy spectra were kindly given to us by Kevin Schomacker of MediSpectra, Lexington MA.

References

1. Claridge, E., Cotton, S.D.: Developing a predictive model of human skin coloring. *Proceedings of SPIE* **2708** (1996) 814–825
2. Claridge, E., Cotton, S.D., Hall, P., Moncrieff, M.: From colour to tissue histology: Physics based interpretation of images of pigmented skin lesions. *Medical Image Analysis* **7** (2003) 489–502
3. Prah, S.A., Keijzer, M., Jacques, S.L., Welch, A.J.: Monte carlo model of light propagation in tissue. In Mueller, G., Sliney, D., eds.: *SPIE Proceedings of Dosimetry of Laser Radiation in Medicine and Biology*. Volume IS 5. (1989) 102–111
4. Wang, L., Jacques, S.L.: Monte Carlo modelling of light transport in multi-layered tissues in standard C. Univ of Texas, MD Anderson cancer center (1998)
5. Prah, S.A.: Mie theory. Oregon Medical Laser Centre, <http://omlc.ogi.edu/software/mie/index.html> (2000)
6. Ge, Z., Schomacker, K.T., Nishioka, N.S.: Identification of colonic dysplasia and neoplasia by diffuse reflectance spectroscopy and pattern recognition techniques. *Applied Spectroscopy* **52** (1998) 833–839
7. Rowe, J.E., Hidović, D.: An evolution strategy using a continuous version of the Gray-code neighbourhood distribution. To appear in *Proceedings of GECCO* (2004)

8. Kubelka, P., Munk, F.: Ein beitrag zur optik der farbanstriche. *Zeitschrift für Technischen Physik* **12** (1931) 593–601
9. Egan, W.G., Hilgeman, T.W.: *Optical Properties of Inhomogeneous Materials*. Academic Press (1979)
10. Jin, Y.: A comprehensive survey of fitness approximation in evolutionary computation. *Soft Computing* (2003) In press.
11. Jin, Y., Olhofer, M., Sendhoff, B.: A framework for evolutionary optimization with approximate fitness function. *IEEE Transactions on Evolutionary Computation* **6** (2002) 481–494
12. Brooker, A.J., Dennis, J., Frank, P.D., Serani, D.B., Torczon, V., Trosset, M.: A rigorous framework for optimization of expensive functions by surrogates. *Structural Optimization* **17** (1998) 1–13