Statistical Restoration of Images
Using a Hybrid Bayesian Approach

D. Hudson and M. Razaz

School of Information Systems,
University of East Anglia, Norwich, UK.
Email: dmh@sys.uea.ac.uk, mr@sys.uea.ac.uk

ABSTRACT

The desire to view smaller and smaller attributes within biological specimens means that confocal microscopes are often used at the limit of their resolution. For quantitative analysis of smaller sized attributes, and as a necessary pre-processing stage for automatic recognition and classification of objects it is essential that confocal images are restored. A fast new hybrid statistical restoration algorithm is presented which makes use of deterministic methodology to speed up optimisation of the posterior probability. Additionally, a prior probability model based on the bayesian classifier is proposed. Restorations of real confocal image data using the above technique and prior are presented and discussed. Quantitative analysis of the improvement gained through our hybrid approach is also presented.

1. INTRODUCTION

Three dimensional confocal laser scanning microscopy (CLSM) provides excellent 2D or 3D images of fluoresently labelled cellular structures within 3D specimens. CLSM has many applications in different fields of science including biological, medical and biomedical sciences, see for example[13]. The confocal microscope achieves its 3D imaging ability by being able to focus at different depths through a specimen. Pinholes in the optical path of the microscope effectively block out light from focal depths other than the one being imaged. As we might expect, this process is not perfect and blurring occurs. Blurring also occurs as a result of the diffraction caused by the finite size of the microscope aperture. Noise is present, mainly due to the photon counts experienced in some samples [9]. This results in the observed image being noisy, despite taking repeat images and averaging them.

Our experience shows that the traditional linear restoration techniques[10] although fast are unsuitable for optical microscopy imaging. They produce artefacts and hence result in poor restored confocal images. The main cause of the difficulty is the incompleteness of the data set (limited number of the focal sections) and noise. The confocal restoration process, like other real imaging systems, is an ill-posed inverse problem with many possible solutions. That is to say that a number of possible ideal (source) images are capable of producing the same observed image. In order to overcome this difficulty, prior knowledge of the characteristics of the imaging system and ideal image are required.

A fast nonlinear algorithm incorporating a priori constraints into the restoration process, called the iterative deconvolution algorithm (IDA), has been developed which overcomes these difficulties and produces high quality confocal restorations [7, 12]. IDA is however, less successful under highly noisy conditions and tends to be not so flexible to changes in the restoration criteria.

We have therefore, for several years, been investigating statistical restoration techniques such as maximum entropy, maximum likelihood and Bayesian techniques. Statistical image processing, especially in a Bayesian framework, provides a powerful tool well suited to the task of image restoration. Such techniques have been well exploited by researchers such as Besag and Green[2-4], Geman and Geman[5]. The work carried out by these researchers has uncovered a whole new methodology for image restoration, one which is very well suited to a wide range of applications and probability models. The Bayesian-based approach, is much more powerful and versatile since it copes well with the inclusion of sources of information related to the images and imaging system. Additionally, we can have greater confidence in the results compared to those achieved by other techniques since it provides not only an estimate of the ideal image but also an estimate of each pixel’s probability of occurrence. If we are unhappy with some feature within our restored image, we can study the probabilities involved in an attempt to get a measure of how confident we should be of the feature’s existence. Statistical image processing does however have its own problems. A purely Bayesian approach is extremely slow, and hence not practical for our vast confocal image data sets. There are also difficulties when we try to formulate statistical models of our images and the imaging system. Accurately estimating the parameters and correct form of these models, especially the prior probability model, is a subject of great discussion.

The restoration technique proposed in this paper makes feasible the use of statistical restoration on moderately sized images with 256 grey levels, as opposed to much of the previous research which has concentrated on small images with few grey levels. This has been achieved using a hybrid approach which exploits deterministic methodology as well as statistical. Also, our algorithm makes use of a number of novel optimisations which greatly reduce processing time. Previous research has tended to deal mainly with the task of noise removal. We show that our algorithm is capable of dealing with highly blurred images. We present here a flexible and robust hybrid methodology combining the Bayesian approach
and the nonlinear IDA approach. The application of the Bayesian restoration to confocal microscopy is new, and has not been reported before in the literature. Our hybrid method is considerably faster than pure Bayesian restoration. We have also developed a new form of prior probability model which rewards image-like characteristics rather than just smoothness.

2. METHOD

We assume multiplicative Poisson distributed noise, the conventional additive Gaussian distribution used in the literature is not appropriate to this imaging modality. We use the Maximum A Posteriori (MAP) approach based on optimisation of the Bayesian formulation:

\[ P(f | g) = P(g | f) P(f) \]  \hspace{1cm} (1)

where \( f \) is the ‘ideal’ (source) image and \( g \) the observed image. Note that \( P(g) \) is not included since it is constant and therefore irrelevant to our optimisation. As seen in (1), two probabilities need to be estimated. These are, the joint probability \( P(g|f) \) and the prior probability \( P(f) \).

2.1 Joint Probability Estimation

The Joint Probability describes the probability of occurrence of the observed image given knowledge of the ideal image. Assuming multiplicative noise, the joint probability can be formulated as follows:

\[ P(g|f) = P\left(n = \frac{g}{f \otimes h}\right) \]  \hspace{1cm} (2)

Where \( n \) is a value from the noise distribution, \( \otimes \) the convolution operation and \( h \) the point spread function (PSF).

We use a highly optimised space domain convolution for calculation of the joint probability. This is based on the knowledge that when optimising the posterior probability using our GICM technique, we are only interested in the change in probability caused by a change in a single pixel's value. Assume that the pixel at \((i,j)\) is the one which has been modified. If \( g'(x,y) \) is the value of the \((x,y)\)th pixel of the \( f \otimes h \) convolution, we can define the joint probability as follows:

\[ P(g(x,y)|f) = P\left(n = \frac{g(x,y)}{g'(x,y)}\right) \]  \hspace{1cm} (3)

If we now consider a pixel at \((x,y)\) in the ideal image whose posterior probability is affected by a single pixel change close by at \((i,j)\), we can calculate its posterior by simply calculating the change that single pixel has on its posterior rather than recalculating \( f \otimes h \) in its entirety. This is expressed by the following equation:

\[ g'_{n+1}(x,y) = g'_n(x,y) + (\Delta f(i,j,n).h(x-i,y-j))g'_n \]  \hspace{1cm} (4)

where \( \Delta f(i,j,n) = f_{n+1}(i,j) - f_n(i,j) \)

2.2 Prior Probability Estimation

The prior probability of occurrence of the ideal image estimate attempts to include \textit{a priori} knowledge of the images we are restoring. Smoothness is an important example of such criteria. We assume that the ideal image can be described by a Markov random field (MRF). Therefore:

\[ P(f|\delta) = P(f_i|f_{\delta i}) \]  \hspace{1cm} (5)

where \( f_{\delta i} \) is the set of values of the pixels which are ‘neighbours’ of \( f_i \) [14].

Estimation of the joint probability is well understood, however the authors of this paper are not convinced by current techniques used for prior probability estimation since, whilst mathematically sound, their exact implementation and choice of parameters seem heuristic. Gibbs priors [5,8] are a prime example of this. They are elegantly formulated however, there is no recognised technique for selecting appropriate potential functions. Some researchers seem to have resorted to using the sum of the difference between the current pixel’s value and the values of each of its neighbours as a basis for potential functions [14]. Ultimately this prior favours completely constant intensity images. Our aim on the other hand has been to design a prior which favours truely image-like characteristics. This has been achieved by the use of a carefully chosen training set. This training set contains confocal images of real samples taken under low magnification conditions. Hence, the images are of high resolution and have low noise content. This data set is used to train a Bayesian Classifier. A general block diagram for the classifier is shown below:

![Fig. 1](image)

Four ‘features’, \( v_1, v_2, v_3, v_4 \) are calculated and the following probability is computed:

\[ P(f_i|f_{\delta i}) = P(f_i|v_1). P(f_i|v_2). P(f_i|v_3). P(f_i|v_4) \]  \hspace{1cm} (6)

The pixels that are neighbours of pixel \((i,j)\) are split into four disjoint sets, and the four features are each a function of the pixel values of one of these sets. The four conditional probabilities in (6) are calculated based on probability density functions generated in advance from the training set. An assumption of independence between features is made. This is in fact not the case, although the use of disjoint sets does increase the degree of independence to a level where useful results can be achieved.
2.3 Optimisation Technique

A frequently discussed technique for probability optimisation in image processing is that of simulated annealing [1]. Given an infinite running time, simulated annealing is capable of finding the global maxima of any function. Unfortunately, the time required to produce good optimal solutions is very high. For this reason, other methods have been employed. We have developed a hybrid algorithm based on a mix of the iterative conditional mode (ICM) approach[2] and IDA[7,12] which combines random attempts at changes in pixel intensity with deterministic ones. This algorithm, which we refer to as guided iterative conditional mode (GICM) optimisation, leads to very fast convergence whilst retaining the extra flexibility of the statistical approach.

Firstly let us consider the ICM approach. An initial estimate of the ideal image is made. A raster scan is then made over the image. At each pixel, a random change is made to that pixel's value and the appropriate probabilities are re-calculated. If the change made increases the conditional probability of occurrence of the ideal image, \( P(f|g) \), then the change is adopted. Otherwise, the original pixel value before the change is restored to that location. The principle behind the ICM approach is twofold. Firstly, it is quicker than a straight simulated annealing based restoration since it starts off closer to the ideal image and does not spend time 'cooling' in the way that simulated annealing does. Secondly, it avoids the large scale characteristics of prior probabilities based on the MRF assumption. Problems with the ICM approach include the fact that the initial estimate has a very large effect on the restored image. ICM can therefore only restore the image to a locally maximal probability in the immediate region of the initial estimate. In addition to this, for reasons of speed, the initial estimate is often gained using deterministic techniques which do not perform well under highly noisy conditions and are unaware of probabilistic implications. GICM on the other hand uses the observed image as its initial estimate and then makes use of deterministic methods for a number of iterations in order to guide it as to how to modify the ideal image estimate. Every modification is made in light of its effect on \( P(f|g) \). Only pixel changes which increase the posterior probability are accepted. An important aspect of an ICM or GICM optimisation is the choice of method for generating new pixel values in the ideal image estimate. ICM is based on the concept of using a random deviation from the current value at the pixel location in question. GICM on the other hand seeks to make more intelligent guesses at pixel values in the ideal image whilst still testing for their effect on \( P(f|g) \) and therefore remaining within the statistical framework. It is in this way that GICM exploits both deterministic and statistical techniques to form a hybrid approach.

If we define \( f_n(x,y) \) as the value of \( f(x,y) \) at the \( n \)th iteration, we can describe the method of pixel modification as follows:

\[
\begin{align*}
    f_{n+1}(x,y) &= f_{\text{det}}(x,y) \quad n < N_1 \\
    &= f_{\text{det}}(x,y) + \gamma(sd_1) \quad N_1 \leq n \leq N_2 \\
    &= f_n(x,y) + \gamma(sd_2) \quad n > N_2
\end{align*}
\]

\( f_{\text{det}}(x,y) \) is an estimate of ideal image according to an iterative, deterministic technique and \( \gamma(sd_1) \) is a random value from a Gaussian distribution with zero mean and standard deviation \( sd_1 \). The iterations at which the behaviour of the algorithm changes are denoted by \( N_1 \) and \( N_2 \), typically taking values of 5 and 10 respectively.

We make use of IDA in order to generate values for \( f_{\text{det}} \) however, any fast, high quality, iterative, deterministic algorithm could be used.

3. RESULTS

Our hybrid Bayesian methodology was successfully applied to a variety of synthetic and confocal images. An accurate 3D PSF, modelled from images of sub-resolution beads using bi-cubic splines [11], was used. We have also successfully implemented the theoretical PSFs proposed by Hell et. al. [6] and achieved good quality restorations. Typical restorations of biological images are presented here (see Figures 2-5). The major advantages of this hybrid methodology are that it is comparatively very fast, and produces high quality restoration under strong blur and noise conditions. Figure 6 shows optimisation of the MAP probability in our algorithm using both the ICM and GICM approaches. The image being restored is an artificially blurred image corrupted by multiplicative noise to give a signal to noise ratio of 10 dB.
4. CONCLUSIONS

In making use of both statistical and deterministic methodology within our algorithm, we have been able to formulate a statistical algorithm, based on GICM, which is far faster than a straight statistical approach. The algorithm is very flexible and capable of restoring images for which deterministic methods alone would fail. As a result of being space domain based, specific areas within an image can be restored and this reduces processing time even further. GICM is shown to be better suited to our optimisation problem than ICM. It achieves better optimisations in less iterations and can easily be adopted for use in other applications. Finally, our work on the prior probability points towards a different form of model which relies much less on trial and error in its formulation than previous efforts.

ACKNOWLEDGEMENTS

The support of Bio-Rad Microsciences Ltd and the UK Biotechnology and Biological Sciences Research Council, BBSRC, for this project is acknowledged.

REFERENCES