Texture Synthesis using Hidden Markov Measure Fields

Mike Rogers* and Chris Taylor

Imaging Science and Biomedical Engineering, The University of Manchester,
Stopford Building, Oxford Road, Manchester, M13 9PT.

1 Introduction

Statistical models of shape and appearance are well-known to provide a powerful constraints in many areas of medical image analysis. However, current modelling techniques still have some important shortcomings that limit their applicability to some medical image analysis problems. Specifically, no adequate representation of image features lying somewhere between consistent structure and fine textural detail has yet been developed. The aim of our work is to develop models capable of representing this *strextural* data. Advances in this area would benefit medical applications such as breast cancer screening using digital mammography, diabetic retinopathy and CT chest scans for lung cancer screening, all of which produce images that contain highly detailed textures with complex structure.

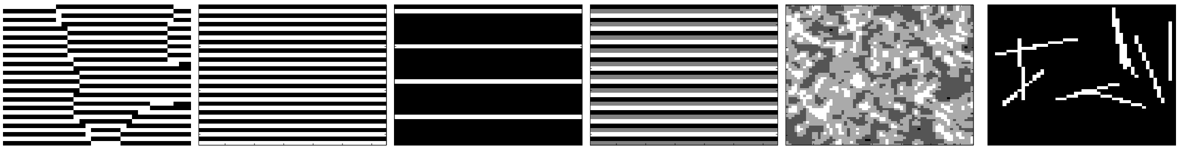
In this preliminary work, we have investigated methods of generating ergodic texture. Our aim is to continue the work of Rose and Taylor [1] on representing mammographic textures. Their method is based on an MRF formulation of texture. We have applied the Hidden Markov Measure Field (HMMF) method of Marroquin [2] to texture synthesis. The HMMF method sets the problem of MRF segmentation within a formal optimisation framework. We are able to use the method to avoid the slightly ad-hoc pixel or region-wise texture generation scheme used by Rose and others. By using HMMFs, we also hope to avoid some of the other problems associated with MRF texture synthesis process, such as texture boundary inconsistencies and invalid texture growing.

2 Method

We have made several modifications to the HMMF segmentation method of Marroquin *et al* [2] in order to synthesise texture images. Very briefly, the goal of the original algorithm is to find a segmentation label field, f, and label model parameters, θ, which maximise the segmentation posterior given the image I: $P(f, \theta|I) = 1/Z \exp[-U(f, \theta)]$, where $U(f, \theta) = -\sum_r \log v_{f(r)}(r) + \sum_C V_C(f)$. Here, $v_{f(r)}(r)$ represents the goodness of fit of model class $f(r)$ to the image data at pixel r (for example a gaussian noise model), and $V_C(f)$ is a *potential function* ranging over all cliques of a given neighbourhood system, and is used to enforce prior constraints on the spatial structure of the label field. Z is a normalising constant. Marroquin avoided the need to use computationally expensive random sampling or approximation methods to solve this maximisation by the introduction of an additional field, p, which represents the probability of each label at each pixel location. This reformulation results in a minimisation of the form: $U(p, \theta) = -\sum_r \log(v(r) \cdot p(r)) + \sum_C W_C(p)$. The image term, $\log(v(r) \cdot p(r))$, and field prior term, $W_C(p)$, are now written in terms of the label probability field, p, and are differentiable. This enables optimisation using any constrained general purpose gradient descent algorithm. The final label field can then be simply taken as the label with maximum probability in p for each pixel in a decoupled manner. Marroquin showed the method can produce accurate segmentations extremely quickly in a variety of applications.

We have taken this scheme and applied it to texture synthesis. Our basic idea is to use the current estimate of label probability to estimate a complete label field during the optimisation process. The field prior term W_C is used to represent the spatial relationships that characterise the texture. We start with an image with a fixed number of discrete gray-levels, $I(r) \in \{1, \ldots, M\}$, and a set of pixels to be synthesised. The label probability field, p, is initialised randomly and normalised subject to the constraints $p(r) \in S_M$, $S_M = \{u \in \Re^M : \sum_{l=1}^{M} u_l = 1, u_l \geq 0, l = 1, \ldots M\}$. A pseudo-random label, $f(r)$, can be drawn for each pixel using the label probabilities of $p(r)$. The label field forms the current estimate of the generated texture. We have included this label generation process in the optimisation procedure in a simple way. At each iteration, the derivative $\frac{\partial U}{\partial p}$ is used to perform a line search. After each line search, pixel values are synthesised into the original image using the current estimate of p. Optimisation w.r.t θ is not required in this formulation. The method converges to a label image that satisfies the spatial constraints encoded in W_C. We represent the spatial relationships characterising a texture using a mixture of gaussians $W_C(p) = \sum_{k=1}^{N_k} Z_k \exp(-0.5(p_C - \mu_k)\Sigma_k^{-1}(p_C - \mu_k))$, where p_C are the values of p within the clique C, μ_k and Σ_k are the parameters of the kth gaussian, and Z_k is a normalising constant. Cliques are composed of

*Email: mike.rogers@manchester.ac.uk, WWW: http://www.isbe.man.ac.uk/∼mdr



(a) (b) (c) (d) (e) (f)

Figure 1. Images (a)-(d) show generated simple textures, images (e) and (f) show data for future work.

neighbourhoods of $n \times n$ pixels and therefore only represent the local structure of the texture in the same way as a standard MRF formulation. The parameters N_k, μ_k and Σ_k can be obtained from training images. For simple textures we represent every distinct neighbourhood with it own gaussian and set $\Sigma_k = I_{n^2}$. For more complex textures a clustering approach would be more appropriate.

When applied to an entire image, this scheme can form regions of texture in separate parts of the image simultaneously. These regions grow independently during the optimisation process and meet one another, potentially producing textural incompatibilities. This is illustrated in figure 1(a). To overcome this, we have introduced a spatial annealing process. Normalisation can be adjusted to force values of p to lie close to their class priors, effectively introducing some forced label uncertainty. We start with a high level of forced uncertainty across the entire image except for the centre (or some other arbitrary subset of pixels). After each optimisation iteration, we use a diffusion process to allow more of the image to become fixed. This process effectively removes boundaries between incompatible texture regions.

We have also found that the mixture model's normalising constants Z_k can be adjusted during optimisation to better meet global prior class probabilities, even though the optimisation process is a local one. This is achieved by increasing Z_k for under-represented clusters and vice-versa.

3 Results and Discussion

In this section we present some preliminary results obtained by our method followed by discussion of the technique. Figure 1(a)-(d) shows examples of simple synthesised textures. Each figure was produced using a neighbourhood clique of 5×5 pixels. The white stripes in the training data used to generate figure 1(c) were spaced 8 pixels apart, a distance exceeding the neighbourhood size. Despite this, the result has approximately the correct number of stripes, with no stripes within 5 pixels of one another, showing the global prior class probabilities have been respected.

The framework combines segmentation and synthesis allowing consistent texture in-filling with no modifications. At present, in this simple formulation, the method has high computational complexity; depending on the synthesised image size and the number of clusters representing the texture. Also, the HMMF algorithm's storage requirements increase with the number of pixel classes. Appropriate implementational and theoretical solutions, such as hierarchical mixture models and dimensionality reduction, will have to be applied to address these issues. These improvements will allow the synthesis of natural ergodic textures such as the stone pattern in figure 1(e).

In the longer-term, our goal is to integrate this texture synthesis scheme into a generative strexture model. In working towards this goal, we hope to formulate W_C to represent linear structures, such as straight lines with arbitrary length distributions (see figure 1(f) for example). We will also investigate using more descriptive image features, such as the DT-CWT, to represent texture regions and apply model-selection techniques to develop near-optimal representations of structures in training images.

References

1. C. J. Rose & C. J. Taylor. "A generative statistical model of mammographic appearance." In D. Rueckert, J. Hajnal & G.-Z. Yang (editors), *Proceedings of Medical Image Understanding and Analysis 2004*, pp. 89–92. 2004.
2. J. L. Marroquin, E. A. Santana & S. Botello. "Hidden markov measure field models for image segmentation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25(11)**, pp. 1380–1387, 2003.