Ising models and multiresolution quad-trees

Wilfrid S. Kendall  
Department of Statistics  
University of Warwick

Roland G. Wilson  
Department of Computer Science  
University of Warwick

3 April 2003

Research supported by EPSRC research grant GR/M75785
Introduction

Multiresolution quad-trees are used in image analysis. For example: given a layer of pixels at finest resolution, successively aggregate blocks of pixels to produce coarser layers. A simple model then stipulates the black/white value of a pixel depends on its neighbours in the same layer and (with a different interaction strength) on its parent and daughters.

This talk describes initial steps in understanding the qualitative behaviour of this algorithm, addressing the following question:

“What do phase transitions look like when the finest layer of pixels is unconstrained (“free boundary”)?”

For details, see Kendall and Wilson (2003).
Example of motivating dataset from medical image analysis:

(Warwick Statistics PhD students will go to extreme lengths in order to contribute to scientific progress . . . )
1. **Image analysis**


Multiresolution MAP algorithm, 1.3% misclassification:

Good for the right kind of objects.
Quad-tree formed by successive averaging using “decimation”: a kind of scale-invariance.
Multiresolution: pros and cons

+ FAST since low resolution “steers” high resolution;
+ adapted to some kinds of HIGH-LEVEL objects;
  - can produce “BLOCKY” reconstructions: resolution hierarchy mediates all spatial interactions.

Possible solution

Add further explicitly spatial interactions?

Where are the phase transitions?
2. **Generalized quad-trees**

Define $Q_d$ as graph whose vertices are cells of all dyadic tessellations of $\mathbb{R}^d$, with edges connecting each cell to its $2^d$ neighbours, and also its parent (covering cell in tessellation at next resolution down) and its $2^d$ daughters (cells which it covers in next resolution up).

![Diagram of generalized quad-tree](image)

Case $d = 1$:

Neighbours at same level also are connected.

**Remark:** No spatial symmetry!

**However:** $Q_d$ can be constructed using an iterated function system generated by elements of a skew-product group $\mathbb{R}^+ \otimes_s \mathbb{R}^d$. 
Further define:

- $\mathcal{Q}_{d,r}$ as subgraph of $\mathcal{Q}_d$ at resolution levels of $r$ or higher;

- $\mathcal{Q}_d(o)$ as subgraph formed by $o$ and all its descendants.

- **Remark:** there are many graph-isomorphisms between $\mathcal{Q}_{d,r}$ and $\mathcal{Q}_{d,s}$, with natural $\mathbb{Z}^d$-action;

- **Remark:** there are graph homomorphisms injecting $\mathcal{Q}(o)$ into itself, sending $o$ to $x \in \mathcal{Q}(o)$ (*semi-transitivity*).
Simplistic analysis

Define $J_\lambda$ to be strength of neighbour interaction, $J_\tau$ to be strength of parent interaction. If $S_x = \pm 1$ then probability of configuration is proportional to $\exp(-H)$ where

$$H = -\frac{1}{2} \sum_{(x,y) \in E(G)} J_{\langle x,y \rangle} (S_x S_y - 1),$$

for $J_{\langle x,y \rangle} = J_\lambda, J_\tau$ as appropriate.

If $J_\lambda = 0$ then the free Ising model on $\mathbb{Q}_d(0)$ is a branching process (Preston 1977; Spitzer 1975); if $J_\tau = 0$ then the Ising model on $\mathbb{Q}_d(0)$ decomposes into sequence of $d$-dimensional classical (finite) Ising models. So we know there is a phase change at $(J_\lambda, J_\tau) = (0, \ln(5/3))$ (results from branching processes), and expect one at $(J_\lambda, J_\tau) = (\ln(1+\sqrt{2}), 0+)$ (results from 2-dimensional Ising model).

But is this all that there is to say?
3. Random clusters

A similar problem, concerning Ising models on products of trees with Euclidean lattices, is treated by Newman and Wu (1990). We follow them by exploiting the celebrated Fortuin-Kasteleyn random cluster representation (Fortuin and Kasteleyn 1972; Fortuin 1972a; Fortuin 1972b):

The Ising model is the marginal site process at $q = 2$ of a site/bond process derived from a dependent bond percolation model with configuration probability $P_{q,p}$ proportional to

$$q^C \times \prod_{\langle x,y \rangle \in \mathcal{E}(G)} \left( (p_{\langle x,y \rangle})^{b_{\langle x,y \rangle}} \times (1 - p_{\langle x,y \rangle})^{1 - b_{\langle x,y \rangle}} \right).$$

(where $b_{\langle x,y \rangle}$ indicates whether or not $\langle x, y \rangle$ is closed, and $C$ is the number of connected clusters of vertices). Site spins are chosen to be the same in each cluster independently of other clusters with equal probabilities for $\pm 1$. 


Random cluster facts

- (Representation of Ising model when \( q = 2 \).) The marginal site process is Ising with

\[
p_{\langle x,y \rangle} = 1 - \exp(-J_{\langle x,y \rangle}); \tag{2}
\]

- (FK-comparison inequalities.) If \( q \geq 1 \) and \( A \) is an increasing event then

\[
\mathbb{P}_{q,p}(A) \leq \mathbb{P}_{1,p}(A) \tag{3}
\]

\[
\mathbb{P}_{q,p}(A) \geq \mathbb{P}_{1,p'}(A) \tag{4}
\]

where

\[
p_{\langle x,y \rangle}' = \frac{p_{\langle x,y \rangle}}{p_{\langle x,y \rangle} + (1 - p_{\langle x,y \rangle})q} = \frac{p_{\langle x,y \rangle}}{q - (q - 1)p_{\langle x,y \rangle}}.
\]

Since \( \mathbb{P}_{1,p} \) is bond percolation (bonds open or not independently of each other), we can find out about phase transitions by studying independent bond percolation.
4. Percolation

Independent bond percolation on products of trees with Euclidean lattices have been studied by Grimmett and Newman (1990), and these results were used in the Newman and Wu work on the Ising model. So we can make good progress by studying independent bond percolation on $\mathbb{Q}_d$, using $p_\tau$ for parental bonds, $p_\lambda$ for neighbour bonds.

**Theorem 1** There is almost surely no infinite cluster in $\mathbb{Q}_{d;0}$ (and consequently in $\mathbb{Q}_{d}(o)$) if

$$2^d \tau \mathcal{X}_\lambda \left( 1 + \sqrt{1 - \mathcal{X}_\lambda^{-1}} \right) < 1,$$

where $\mathcal{X}_\lambda$ is the mean size of the percolation cluster at the origin for $\lambda$-percolation in $\mathbb{Z}^d$.

Modelled on Grimmett and Newman (1990, §3 and §5).

Get $\left( 1 + \sqrt{1 - \mathcal{X}_\lambda^{-1}} \right)$ from matrix spectral asymptotics.
The story so far: small $\lambda$, small to moderate $\tau$. 
Case of small $\tau$

Need $d = 2$ for mathematical convenience. Use Borel-Cantelli argument and planar duality to show, for supercritical $\lambda > 1/2$ (that is, supercritical with respect to planar bond percolation!), all but finitely many of the resolution layers $L_n = [1, 2^n] \times [1, 2^n]$ of $\mathbb{Q}_2(o)$ have just one large cluster each of diameter larger than constant $\times n$.

Hence . . .

**Theorem 2** When $\lambda > 1/2$ and $\tau$ is positive there is one and only one infinite cluster in $\mathbb{Q}_2(o)$. 
The story so far: adds small $\tau$ for case $d = 2$. 
Uniqueness of infinite clusters

The Grimmett and Newman (1990) work was remarkable in pointing out that as \( \tau \) increases so there is a further phase change, from many to just one infinite cluster for \( \lambda > 0 \). The work of Grimmett and Newman carries through for \( Q_d(0) \). However the relevant bound is improved by a factor of \( \sqrt{2} \) if we take into account the hyperbolic structure of \( Q_d(0) \)!

**Theorem 3** If \( \tau < 2^{-(d-1)/2} \) and \( \lambda > 0 \) then there cannot be just one infinite cluster in \( Q_{d;0} \).

**Method:** sum weights of “up-paths” in \( Q_{d;0} \) starting, ending at level 0. For fixed \( s \) and start point there are infinitely many such up-paths containing \( s \lambda \)-bonds; but no more than \((1 + 2d + 2^d)^s\) which cannot be reduced by “shrinking” excursions. Hence control the mean number of open up-paths stretching more than a given distance at level 0.
Contribution to upper bound on second phase transition:

**Theorem 4**  If $\tau > \sqrt{2/3}$ then the infinite cluster of $Q_{2:0}$ is almost surely unique for all positive $\lambda$.

**Method:** prune bonds, branching processes, 2-dim comparison . . .
The story so far: includes uniqueness transition for case $d = 2$. 
5. Comparison

We need to apply the Fortuin-Kasteleyn comparison inequalities (3) and (4). The event “just one infinite cluster” is not increasing, so we need more. Newman and Wu (1990) show it suffices to establish a finite island property for the site percolation derived under adjacency when all infinite clusters are removed. Thus:
Comparison arguments then show the following schematic phase diagram for the Ising model on $\mathbb{Q}_2(o)$:

- All nodes substantially correlated with each other in case of free boundary conditions.
- Unique Gibbs state.
- Root influenced by wired boundary.
- Free boundary condition is mixture of the two extreme Gibbs states (spin 1 at boundary, spin −1 at boundary).

Parameters:
- $J_\lambda$
- $J_\tau$
- Transition points:
  - $0.511$ for $J_\tau$
  - $0.288$ for $J_\lambda$
  - $1.228$ for $J_\tau$
  - $1.099$ for $J_\lambda$
  - $1.228$ for $J_\lambda$
  - $2.292$ for $J_\tau$
  - $2.292$ for $J_\lambda$
6. **Simulation**

Approximate simulations confirm the general story:

http://www.dcs.warwick.ac.uk/rgw/sira/sim.html

(1) Only 200 resolution levels;

(2) At each level, 1000 sweeps in scan order;

(3) At each level, simulate square sub-region of $128 \times 128$ pixels conditioned by mother $64 \times 64$ pixel region;

(4) Impose periodic boundary conditions on $128 \times 128$ square region;

(5) At the coarsest resolution, all pixels set white. At subsequent resolutions, ‘all black’ initial state.
Future work

(a) $J_\lambda = 1, J_\tau = 0.5$
(b) $J_\lambda = 1, J_\tau = 1$
(c) $J_\lambda = 1, J_\tau = 2$

(d) $J_\lambda = 0.5, J_\tau = 0.5$
(e) $J_\lambda = 0.5, J_\tau = 1$
(f) $J_\lambda = 0.5, J_\tau = 2$

(g) $J_\lambda = 0.25, J_\tau = 0.5$
(h) $J_\lambda = 0.25, J_\tau = 1$
(i) $J_\lambda = 0.25, J_\tau = 2$
7. Future work

This is about the free Ising model on $\mathbb{Q}_2(o)$. Image analysis more naturally concerns the case of prescribed boundary conditions (say, image at finest resolution level . . .).

**Question:** will boundary conditions at “infinite fineness” propagate back to finite resolution?

_Series and Sinař (1990)_ show answer is yes for analogous problem on hyperbolic disk (2-dim, all bond probabilities the same).

_Gielis and Grimmett (2002)_ point out (eg, in $\mathbb{Z}^3$ case) these boundary conditions translate to a _conditioning_ for random cluster model, and investigate using large deviations.

**Project:** do same for $\mathbb{Q}_2(o)$ . . . and get quantitative bounds? **Project:** generalize from $\mathbb{Q}_d$ to graphs with appropriate hyperbolic structure.
References


Thönnnes, E., A. Bhalerao, W. S. Kendall, and R. G. Wilson [2002]. A Bayesian ap-

Wilson, R. and C. Li [2002]. A class of discrete multiresolution random fields and its application to image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(11), Accepted for publication.
Notes on proof of Theorem 1

Mean size of cluster at $0$ bounded above by

\[
\sum_{n=0}^{\infty} \sum_{t:|t|=n} \mathcal{X}_{\lambda} \tau^n (\mathcal{X}_{\lambda} \tau - 1)^{T(t)} \mathcal{X}_{\lambda}^{n-T(t)} \\
\leq \sum_{n=0}^{\infty} \mathcal{X}_{\lambda} (\tau \mathcal{X}_{\lambda})^n \sum_{t:|t|=n} (1 - \mathcal{X}_{\lambda}^{-1})^{T(t)} \\
\leq \sum_{n=0}^{\infty} \mathcal{X}_{\lambda} (2^d \tau \mathcal{X}_{\lambda})^n \sum_{j:|j|=n} (1 - \mathcal{X}_{\lambda}^{-1})^{T(j)} \\
\approx \sum_{n=0}^{\infty} \mathcal{X}_{\lambda} (2^d \tau \mathcal{X}_{\lambda})^n \left(1 + \sqrt{1 - \mathcal{X}_{\lambda}^{-1}}\right)^n.
\]

For last step, use spectral analysis of matrix representation

\[
\sum_{j:|j|=n} (1 - \mathcal{X}_{\lambda}^{-1})^{T(j)} = [1 \ 1] \left[ \begin{array}{cc} 1 & 1 \\
1 - \mathcal{X}_{\lambda}^{-1} & 1 \end{array} \right]^n \left[ \begin{array}{c} 1 \\
1 \end{array} \right].
\]
Notes on proof of Theorem 2

Uniqueness: For negative exponent $\xi (1 - \lambda)$ of dual connectivity function, set

$$\ell_n = (n \log 4 + (2 + \epsilon) \log n) \xi (1 - \lambda).$$

More than one “$\ell_n$-large” cluster in $L_n$ forces existence of open path in dual lattice longer than $\ell_n$. Now use Borel-Cantelli . . . .

On the other hand super-criticality will mean some distant points in $L_n$ are inter-connected.

Existence: consider $4^{n-[n/2]}$ points in $L_{n-1}$ and specified daughters in $L_n$. Study probability that

(a) parent percolates more than $\ell_{n-1}$,

(b) parent and child are connected,

(c) child percolates more than $\ell_n$.

Now use Borel-Cantelli again . . . .
Notes on proof of Theorem 3

Two relevant lemmas:

**Lemma 1** Consider $u \in L_{s+1} \subset \mathbb{Q}_d$ and $v = M(u) \in L_s \subset \mathbb{Q}_d$. There are exactly $2^d$ solutions in $L_{s+1}$ of

$$M(x) = S_{u,v}(x).$$

One is $x = u$. The others are the remaining $2^d - 1$ vertices $y$ such that the closure of the cell representing $y$ intersects the vertex shared by the closures of the cells representing $u$ and $M(u)$. Finally, if $x \in L_{s+1}$ does not solve $M(x) = S_{u,v}(x)$ then

$$\|S_{u,v}(x) - S_{u,v}(u)\|_{s,\infty} > \|M(x) - M(u)\|_{s,\infty}. \quad (5)$$

**Lemma 2** Given distinct $v$ and $y$ in the same resolution level. Count pairs of vertices $u$, $x$ in the resolution level one step higher, such that

(a) $M(u) = v$; (b) $M(x) = y$; (c) $S_{u,v}(x) = y$.

There are at most $2^{d-1}$ such vertices.
Notes on proof of Theorem 4

Prune! Then a direct connection is certainly established across the boundary between the cells corresponding to two neighbouring vertices $u, v$ in $L_0$ if

(a) the $\tau$-bond leading from $u$ to the relevant boundary is open;

(b) a $\tau$-branching process (formed by using $\tau$-bonds mirrored across the boundary) survives indefinitely, where this branching process has family-size distribution Binomial($2, \tau^2$);

(c) the $\tau$-bond leading from $v$ to the relevant boundary is open.

Then there are infinitely many chances of making a connection across the cell boundary.
Notes on proof of 
infinite island property

Notion of “cone boundary” $\partial_c(S)$ of finite subset $S$ of vertices: collection of daughters $v$ of $S$ such that $Q_d(v) \cap S = \emptyset$.

Use induction on $S$, building it layer $L_n$ on layer $L_{n-1}$ to obtain an isoperimetric bound: $\#(\partial_c(S)) \geq (2^d - 1)\#(S)$. Hence deduce

$$\mathbb{P}[S \text{ in island at } u] \leq (1 - p_\tau(1 - \eta))^{(2^d - 1)n}$$

where $\#(S) = n$ and $\eta = \mathbb{P}[u \text{ not in infinite cluster of } Q_d(u)]$.

Upper bound on number $N(n)$ of self-avoiding paths $S$ of length $n$ beginning at $u_0$:

$$N(n) \leq (1 + 2d + 2^d)(2d + 2^d)^n.$$

Hence upper bound on the mean size of the island:

$$\sum_{n=0}^{\infty} (1 + 2d + 2^d)(2d + 2^d)^n \eta_{br}^{n(1-2^{-d})},$$

where $\eta_{br}$ is extinction probability for branching process based on Binomial($2^d, p_\tau$) family distribution.