The approximate entropy concept extended to three dimensions for calibrated, single parameter structural complexity interrogation of volumetric images

The approximate entropy concept extended to three dimensions for calibrated, single parameter structural complexity interrogation of volumetric images

Christopher Moore[1] **and Thomas Marchant**

Manchester Academic Health Sciences Centre, The University of Manchester, The Christie NHS Foundation Trust, Christie Medical Physics & Engineering, Withington, Manchester, M20 4BX, UK

[1]E-mail Chris.Moore@christie.nhs.uk

Abstract

Reconstructive volumetric imaging permeates medical practice because of its apparently clear depiction of anatomy. However, the tell tale signs of abnormality and its delineation for treatment demand experts work at the threshold of visibility for hints of structure. Hitherto, a suitable assistive metric that chimes with clinical experience has been absent. This paper develops the complexity measure approximate entropy (ApEn) from its one-dimensional physiological origin into a three-dimensional (3D) algorithm to fill this gap. The first 3D algorithm for this is presented in detail. Validation results for known test arrays are followed by a comparison of fan-beam and cone-beam X-ray computed tomography image volumes used in image guided radiotherapy for cancer. Results show the structural detail down to individual voxel level, the strength of which is calibrated by the ApEn process itself. The potential for application in machine assisted manual interaction and automated image processing and interrogation, including radiomics associated with predictive outcome modeling, is discussed.

1. Introduction

Regularity statistics are a promising way of detecting subtle changes in data where conventional variability statistics (mean, standard deviation, gradient and the like) have failed. They are sensitive not only to the data values but also their ordering in sample sequences. Ordering can be temporally or spatially directed. Hence, they have been formulated for use in the analysis of perturbations to one dimensional (1D) physiological time series (Moore *et. al.* 2006) where a useful heuristic measure has been 1D approximate entropy, $ApEn_{1D}$ (Pincus, 1991*).* By assuming an image is a folded raster sequence of 1D data, $ApEn_{1D}$ has been used to smooth two dimensional (2D) magnetic resonance (MR) images of patients (Parker *et. al.* 1999) and to look for cardiac abnormalities in nuclear medicine MUGA scans (Cullen et. al. 2010). Motion detection from digital video has been considered (Ngan 1997). Each pixel location in a video frame is treated as the source of changing grey-scale values that are temporally windowed and then $ApEn_{1D}$ analyzed. Specific values of $ApEn_{1D}$ are shown to be indicative of target motion rather than random background variations. In both of the imaging applications short vector time series are processed independently of their orthogonal neighbours, even though this is known to be less than realistic. Interestingly, Ngan's indicative values correspond to those described by Moore albeit in a quite different physiological context, which suggests some potential for standardization.

In a theoretical study of randomness, Singer and Pincus (1998) applied the original 1D ApEn measure to small simulated patterns, or 'Latin square' arrays, where values occur exactly once in each row and once in each column i.e. an NxN array that contains N distinct values. Latin squares have inherent linear structural content, a feature that has been exploited by using them as masks for edge detection in correlated noise (Stern *et. al.* 1988). In the study, ApEn was calculated by moving 1D sample vectors, of size m<N taken from the array, along a pre-defined direction through the entire array in the search for a match. The result is a family of $ApEn_{2D}$ parameters whose values do depend on the chosen direction of search. Beyond this we note that the experiment is a useful example of validation, which we extend and develop to prove the multi-dimensional development of approximate entropy in our paper.

Recently, Moore (2016) reported the first ApEn algorithm inherently operating in 2D by introducing divergence as a generalization of direction. This paper extends the algorithm to 3D to support the quantitative analysis not only of planar but also volume data arrays. These are now widespread in imaging applications ranging from non-destructive testing to medical interventions. 2D and 3D ApEn offer statistical measures of pattern regularity associated with image points in the context of their immediate surroundings, which, in a self calibrated manner, can represent simplistic through complex to random structural content. Furthermore, validation testing with simulated patterns of data is presented. The 2D and 3D ApEn extensions are then illustrated by application to wide-angle X-ray cone-beam CT volume images, which are widely used in hospitals for image guided radiation therapy of cancer. These are increasingly interrogated in 'radiomics' in the effort to understand how what was done to a patient influences their treatment outcome. This application reflects our historical motivation, and the fact that clinicians still manually delineate disease boundaries that are often at the threshold of perceptibility on CT images whose 'quality' varies considerably by modality, patient and the constraints that come with the comparatively high frequency of imaging in radiotherapy.

2. Theory

Images represent real world data on a prodigious scale because of their high spatial dimensionality and occurrence in temporal sequences. Medical images in particular are neither random nor simplistically regular but complex, since they carry useful structural information about the human form, perturbing diseases and pathologies that require expert training to discern. Local changes in the behaviour and connectivity of image data can define useful detail. Hence, quantifying these often subtle changes, ideally using a single parameter, assumes considerable importance.

2.1 Approximate Entropy

In the 1950s Kolmogorov and Sinai developed a 1D entropy statistic, KS, for assessing non-linear dynamic systems whose complex behaviour showed changes from regular to irregular states (Kolmogorov 1958). KS entropy is zero for regular systems and positive-finite for irregular systems. These are often termed chaotic to distinguish them from ones that are random and characterised by infinite entropy. Latora and Baranger (1999) describe KS entropy as an entropy rate, differentially related to but not to be confused with the more widely known physical and Shannon entropies. To obtain convergence when calculating the KS entropy for real world signals required an impractically large amount of data. Subsequently, Pincus (1991) developed a more pragmatic algorithm for calculating the entropy from heuristics basics, which can give meaningful results from practical levels of real world data. Due to the lack of full mathematical rigour the term 'approximate' entropy, ApEn, is used to describe his measure.

In a single value, ApEn quantifies the repeatability of the elemental local patterns within a 1D sample series. An attractive feature is that no assumptions need to be made about the shape or functional basis of the patterns being sought; these are directly represented by short contiguous sections systematically sourced from the series itself. Each source section is compared with all similar length target sections of data along the entire length of the series. Where there is a source-target section 'match', the size of the sampled section is incremented to see if the matching condition endures with a similarly extended target section.

Exact matching is not required, rather matches within a given tolerance 'r' are considered. In practice 'r' must not be so low as to unduly restrict the potential for matching (Pincus et. al. 1994), since it is the probabilities of pre and post incremental matching that indicate the regularity of the data and are used to compute ApEn. In particular, the order in which the data values appear assumes importance; a feature that distinguishes ApEn as a regularity rather than a variability statistic.

The above can be formally described as follows. Given N data points $\{u(i)\}=u(1),u(2),\ldots,u(N)$ and commencing with the ith point, vector sequences $\tilde{x}(1)$ to $\tilde{x}(N-m+1)$ are formed consisting of m consecutive u values $\tilde{x}(i)=[u(i),\ldots,u(i+m-1)]$. The sequence, $\tilde{x}(1),\tilde{x}(2),\ldots,\tilde{x}(N-m+1)$ is used to construct $C_i^m(r)$ values for each $i \leq (N-m+1)$, where;

$$C_i^m(r) = (\text{ number of } j \leq (N-m+1) \text{ such that } d[\tilde{x}(i),\tilde{x}(j)] \leq r \text{ })/(N-m+1) \qquad (1)$$

Note that $d[\tilde{x}(i), \tilde{x}(j)]$ in equation (1) is the distance between vectors, defined as the maximum difference in their scalar components, compared on an element by element basis i.e. calculated by taking the modulus of differences and then taking the maximum of these values. The $C_i^m(r)$ values measure, within a tolerance 'r', the regularity or frequency of sequences occurring in the data set $\{u(i)\}$ which are similar to the given sequence, $\tilde{x}(i)$ of length m. The Pincus approximate entropy statistic is then defined by;

$$ApEn = (N-m+1)^{-1} \sum_{i=1}^{N-m+1} \ln[C_i^m] - (N-m)^{-1} \sum_{i=1}^{N-m} \ln[C_i^{m+1}] \qquad (2)$$

$$= \Phi_m - \Phi_{m+1}$$

Equation (2) is interpreted heuristically as a measure of the average logarithmic likelihood, over all sequences $\tilde{x}(1)$ to $\tilde{x}(N-m+1)$, such that any sequence in the data series $\{u(i)\}$, which is within a tolerance r of the given sequence $\tilde{x}(i)$ of length m, remains within the same tolerance when the length of both sequences is increased by one data point. Tolerance r can be defined as proportional to the measured series' standard deviation σ, i.e. r=k σ where k is a constant. It is necessary to determine the optimum k empirically so that the widest range of complexity values is achieved.

In summary, ApEn encapsulates the degree of irregularity by measuring the frequency with which patterns of length *m* selected from a sequence then appear in the sequence. Note that changing the range of a sequence makes no difference to ApEn, which is a particularly useful property. In a highly irregular sequence, each possible pattern appears with roughly equal frequency, and so gives rise to a high ApEn. In contrast a highly regular sequence tends to contain a predominance of a few patterns and a scarcity of other patterns, thus yielding a low ApEn. Finally, a uniform sequence contains just one pattern, the simplest being flat, such that Φ^m and Φ^n would be equally high. Therefore, when subtracted they give zero ApEn.

2.2 Extension of Approximate Entropy to Higher Spatial Dimensions: Key Observations

When considering a single point or datum in an array the question arises how much data does one need to establish ApEn? Ngan considers 1D ApEn with m=1 for sliding vector 'windows' of grey scale data. Ngan's choice of m=1 for an N=16 sample vector appears to be contextually pragmatic, with individual pixel locations considered to be point sources of independent time series. Despite being small, the m and N values are consistent with the Pincus (1994) paper on physiological time series, where he notes that the choices of N and 'm' depend on the available data in the manner predicted by Wolf *et. al.* (1995). Namely, N should be $\approx 10^m < N < \approx 30^m$. This provides a lower limit on sample size for a given value of m, which ensures that enough pattern matches occur in the sample to produce a statistically valid estimate of ApEn. In extremis, the single case of matching the source to itself as the target would obviously introduce biasing through under estimation of the conditional matching probabilities.

Observe here that stationary, real world images tend to have short correlation lengths. In fact, the adjacent pixels in first order finite differenced 2D CT cross-sectional images remain correlated but others, at larger displacements, much less so (Moore 1988). This is why the use of finite differencing is an effective, information preserving method for CT image data compression, albeit suboptimal in terms of achievable compression ratios. Clearly then, so far as these 2D image data are concerned, the choice of m=1 for ApEn calculation will result in source vectors whose elements are likely to be in some way optimally attuned to the matching process, because of residual adjacent element correlation, thus helping to reduce ApEn unreliability. By extension, the same is expected to hold true for the 3D CBCT image volumes. Nevertheless, where flat areas of an image are dominated by for example photon noise, ApEn estimates may be high but inaccurate, not only because of bias but also because of the inherent instability of ApEn calculation to low signal to noise ratios (Pincus 1994). For imaging this may be of little consequence, since the absolute accuracy of ApEn is less important than its utility as a structural regularity indicator.

Recall that an interesting feature of the calculation of ApEn for 1D data is that the source is simply a short sample of the vector itself; a *template* that is compared to all other similar length target sections in the vector to determine whether or not a match endures when their lengths are incremented. In fact a tolerable match is

commonly sought based on the statistics of the *entire* set of data in the vector. The process is iterated, with all possible source sections constructed from the vector used in the comparative process. Computationally, for 1D data series, this is done most efficiently by an incremental 'shift and compare' methodology. However, as long as comparisons are performed along the entire vector it is observed that there is no need for a contiguous approach. Indeed, the comparison of source and target sections could be accomplished by addressing all vector locations once but in random order.

Finally, note that no data connectivity is implied in ApEn calculation *except* in the source section itself, the target sections being compared and incrementing the length of both whenever a match is found. To extend the concepts underlying ApEn computation to higher spatial dimensions, for example the 2D and 3D imaging applications that are the focus of this paper, such incrementing is treated as 'divergence' from a point. The 2D ApEn algorithm has already been described in detail by Moore (2016). The first 3D ApEn algorithm is presented below. Both 2D and 3D ApEn are then validated and compared using planar and volumetric test data arrays and medical CT reconstructions.

2.3 The 3D ApEn Algorithm

For volume images composed of regularly spaced voxels, and bearing in mind the suggested lower limits for N are $\approx 10^m < N < \approx 30^m$ (Pincus 1994), the computation of a 3D approximate entropy analogue, $ApEn_{3D}$ could comfortably operate using m=1 with matching operations based on the 27 points in a 3x3x3 minimal matrix with the point of interest at the centre.
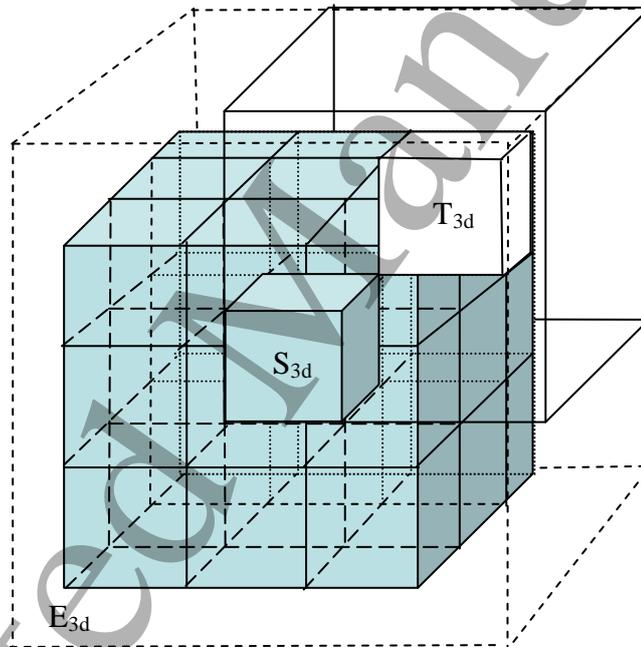
Figure 1. For a 3D image volume the minimal 3x3x3 cell consists of a central source point, S_{3d}, (solid, shaded cube) and 26 nearest neighbours (solid lattice). $ApEn_{3D}$ with m=1, is required for S_{3d}. In addition to the source point itself, each of the nearest neighbours is taken as a target point, T_{3d}, and compared to the source point for a tolerable match (e.g. the unshaded cube, top right, is a target point with lower left vertex contacting the source point). Given a source-target match, the absolute differences between corresponding pairs of data in the extended envelopes around the source and target S_{3d} and T_{3d} are calculated, and their mean $m_{abs.dif}$ taken. Where $m_{abs.dif} < k.SD_{ecell}$ the match is said to endure where SD_{ecell} is the standard deviation of the values in the extended, 5x5x5 cell, E_{3d}, (the union of S_{3d}, all possible T_{3d} and their nearest neighbours (dashed shell).

As in the 2D case, each available point becomes a source section for matching across the minimal cell. If S_{3d} is the current source point being compared to a target point such as T_{3d} in figure 1, and a tolerable source-target match is found, both S_{3d} and T_{3d} must then be incremented divergently in 3D, i.e. orthogonally, to test the enduring nature of the match. This amounts to a comparison of the two nearest neighbour shells, each with 26 values. When all possible S_{3d} and T_{3d} locations have been considered, the data within a 5x5x5 lattice will have been considered i.e. an extended cell, E_{3d}, consisting of the minimal cell plus a shell of *next* nearest neighbours to the central test point. Hence, the basis for the matching tolerance is $r = k.SD_{ecell}$, where SD_{ecell}

is the standard deviation of all the data in E_{3d}, rather than the data in the minimal cell alone. Calculating $ApEn_{3D}$ for higher values of m would require cell comparisons based on larger lattices.

Consider current point T_{3d} again. Ideally, the extended matching process in 3D requires the values of *each* of the nearest neighbours to T_{3d} to be within 'r' of the *corresponding* nearest neighbour to S_{3d}. This would be a true $ApEn_{3D}$. However, the computational burden is potentially severe in the context of application to large volume image analysis. A more pragmatic approach might again be to consider the nearest neighbours collectively. As before we compare neighbourhoods surrounding T_{3d} and S_{3d} using the mean of absolute differences between corresponding elements. Then if the mean value of the absolute differences between each of the corresponding nearest neighbours, m_{diff}, is less than $k.SD_{ecell}$ then an extended and tolerable match has been identified, i.e. $m_{diff} < k.SD_{ecell}$, where k is a constant empirically determined in the usual way. The point and extended match frequencies yield $ApEn_{3D}$ in the established fashion albeit with the denominator in equation-2 being fixed at N. The algorithm for calculating $ApEn_{3D}$ is as follows:

1. Choose *D*, an odd integer, *m*, an integer, and *r*, a positive real number. *D* specifies the size or dimension of the image region to be compared, *m* defines the length of compared runs of data within the region, and *r* specifies a filtering level.

2. $\left\{ u(i,j,k): x - \frac{(D-1)}{2} - m \le i \le x + \frac{(D-1)}{2} + m, \; y - \frac{(D-1)}{2} - m \le j \le y + \frac{(D-1)}{2} + m, \right.$

$\left. z - \frac{(D-1)}{2} - m \le k \le z + \frac{(D-1)}{2} + m \right\}$ forms a segment of 3D image data of dimension $(D+2m) \times (D+2m) \times (D+2m)$, centred on the point (x,y,z).

3. Form an $D \times D \times D$ array of vectors

$$q_{i,j,k} : \left(x - \frac{(D-1)}{2} \le i \le x + \frac{(D-1)}{2}, \; y - \frac{(D-1)}{2} \le j \le y + \frac{(D-1)}{2}, \; z - \frac{(D-1)}{2} \le k \le z + \frac{(D-1)}{2} \right)$$

defined by $q_{i,j,k} = [u(i,j,k), u_1^{rim}(i,j,k), u_2^{rim}(i,j,k), \ldots, u_{m-1}^{rim}(i,j,k)]$, where $u_m^{rim}(i,j,k)$ is itself a vector consisting of the $(24m^2 + 2)$ elements forming the mth rim around $u(i,j,k)$.

4. Use vectors $q_{i,j,k}$ to construct for each

$$(i,j,k) : \left(x - \frac{(D-1)}{2} \le i \le x + \frac{(D-1)}{2}, \; y - \frac{(D-1)}{2} \le j \le y + \frac{(D-1)}{2}, \; z - \frac{(D-1)}{2} \le k \le z + \frac{(D-1)}{2} \right),$$

the values $C_{i,j,k}^m(r) = (\text{number of } q_{a,b,c} \text{ such that } d[q_{i,j,k}, q_{a,b,c}] \le r) / D^3$.

Define $d[q,q^*] = \max_l [diff(u_l, u_l *)]$, where $u_l : (0 \le l \le m-1)$ are the *m* components of *q*, and *diff*() represents the mean absolute difference between the corresponding elements of u_l and u_l* i.e.

$$diff(u_0, u_0 *) = |u_0 - u_0 *| \text{ if } l=0 \qquad \text{otherwise } diff(u_l, u_l *) = \frac{1}{24l^2 + 2} \sum_{a=1}^{24l^2+2} |u_l(a) - u_l *(a)|$$

Here $u_l(a)$ represents the *a*th element of u_l (recall that u_l has $24l^2 + 2$ elements). Hence *d* represents the distance between the vectors $q_{i,j,k}$ and $q_{a,b,c}$, which is given by the maximum difference in their respective components.

5. Next define $\Phi^m(r) = D^{-3} \sum_{i=x-(D-1)/2}^{x+(D-1)/2} \sum_{j=y-(D-1)/2}^{y+(D-1)/2} \sum_{k=z-(D-1)/2}^{z+(D-1)/2} \ln C_{i,j,k}^m(r),$

where ln is the natural logarithm.

6. Define 3D approximate entropy of the region around point (x,y,z) by $ApEn_{3D} = \Phi^m(r) - \Phi^{m+1}(r)$, for m and r fixed as in step 1.

3. Methods

A multi-stage methodology for proving 3D ApEn was adopted. Consistency with the original heuristic 1D ApEn is first demonstrated with random then progressively degraded regularly structured data series. This approach is then extended to 2D and 3D multi-dimensional test patterns. The 2D and 3D ApEn algorithms are then applied to a range of synthetic test images before final application to real-world planar and volume images, which in this instance are medical computed tomograms.

3.1 Functional consistency between the original heuristic $ApEn_{1D}$ and $ApEn_{3D}$

This is important to establish. Hence, a dimensionally downgraded version of $ApEn_{1D}$, was created simply by modifying the extension of the sample template in the matching process to backward-and-forward i.e. diverging from a datum in a 1D series; henceforth this is termed $mApEn_{3D}$, which should show comparable behaviour tot the Pincus ApEn. Purely random sample sequences, uniformly distributed with mean = 0.0 and SD = 1.0, hence r=k, were then processed using both algorithms. In order to establish their optimum ApEn ranges, calculations were performed for m = 1, N = 1000, with values of k varied from 0.001 to 1.0. Subsequently, with the two optimum k values and m=1, the value of N was varied from 30 to 10000 to establish how these ranges behave with sample size.

The manner in which the two 1D algorithms vary when a simply structured signal is progressively compromised by randomness was investigated using a 1D sinusoidal signal modified by incorporating elements of randomness via the so called MIX(p) process (Pincus 1991). This consists of a regular 1D sine wave component with a random selection of points replaced by a uniformly distributed random number between $-\sqrt{3}$ and $+\sqrt{3}$. The parameter p defines the probability that each point will be replaced by a random number. Hence, as p increases from 0 to 1, the process makes a transition from being wholly regular to wholly random. MIX(p) sequences were generated for values of p between 0 and 1 and calculated for both the conventional $ApEn_{1D}$ and modified $mApEn_{3D}$ definitions.

3.2 Functional behaviour and range optimization of $ApEn_{2D}$ and $ApEn_{3D}$

2D and 3D approximate entropies, $ApEn_{2D}$ and $ApEn_{3D}$, were coded in the IDL language for planar and volume array processing respectively. 3D arrays corresponding to the seven 2D test patterns shown in figure 2 were computed. However, we are of course only able to reproduce the 2D versions in print.

To find the optimum ranges of $ApEn_{2D}$ and $ApEn_{3D}$ for uniformly distributed random number patterns-7 the threshold parameter, k, was varied for D = 3 and D = 5. The value of k resulting in the largest value of ApEn is preferred, since it gives the largest separation between uniform (where ApEn = 0) and random patterns 1 and 7 respectively. The mean ApEn value for 1000 random patterns was used at each value of k.

p=0.0   p=0.2   p=0.4   p=0.6   p=0.8   p=1.0

1   2   3   4   5   6   7
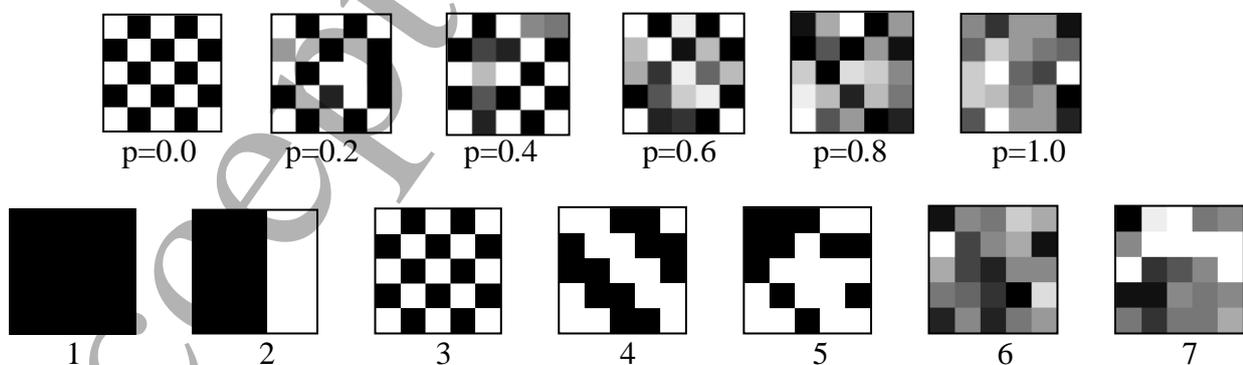
Figure 2: (upper) Examples of synthetic chequer board with increasing randomization parameter p
(lower) Examples of synthetic test arrays, left to right: (1) uniform zeroes, (2) step, (3) chequerboard,
(4) diagonal stripes, (5) random zeroes & ones, (6) Gaussian random numbers (mean 1000, SD=50) ,
(7) uniform random numbers (range 0 to 10).

2D and 3D binary chequer board patterns-3 with 1 voxel squares then had a random selection of voxels replaced by a uniformly distributed random number between 0 and 1. The parameter p defines the probability that each voxel is replaced. Patterns were generated with values of p between 0 and 1. Figure 2(upper) illustrates the modifications seen in the chequer board. $ApEn_{2D}$ and $ApEn_{3D}$ were calculated for m = 1, and image region sizes of D = 3 and D = 5 (see algorithms in section 2.3 and Moore (2016)). For each value of p 1000 patterns were generated and the average and standard deviation of the resulting $ApEn_{2D}$ and $ApEn_{3D}$ values found.

Multi-dimensional $ApEn_{2D}$ and $ApEn_{3D}$ algorithms were used to calculate ApEn values for 2D and 3D patterns ranging from regular through to random as shown in figure 2 (lower): (1) uniform zeros (2) step function (3) chequerboard pattern with 1 voxel squares (4) diagonal stripes 2 voxels wide (5) binary random distribution (6) normally distributed random numbers (7) uniformly distributed random numbers. For the random patterns, the mean and standard deviation of the ApEn values for 10,000 different random selections was calculated.

3.3 Processing of Real World Volumetric Data

Multi-dimensional ApEn processing using the $ApEn_{2D}$ and $ApEn_{3D}$ algorithms was applied to real world data in the form of fully anonymized reconstruction images sourced from the local IGRT programme. These are diagnostic quality fan-beam CT images acquired using a GE high-speed CT/i scanner (GE Medical Systems) at 120 kV with 5 mm slice width and 0.9 mm pixel size within each slice. Additionally, therapy quality cone beam CT images (CBCT) were acquired with the Elekta XVI imaging system (XVI 3.5, Elekta, Crawley, UK), at 120 kV. The latter images were reconstructed with 1 mm^3 voxels covering a field of view of $41 \times 41 \times 12$ cm^3 in the lateral, vertical and longitudinal directions respectively. CBCT images were analyzed at both 1 mm slice width and after averaging to produce 5 mm slices to match the slice width of the fan beam CT images. Use of $ApEn_{2D}$ and $ApEn_{3D}$ gave some insight into the impact of extending processing from conventional planar cross-sections to volume images, where the influence of out of plane data is currently more difficult for experts to fathom let alone quantify.

To a first approximation the tomographic image data can be assumed to be locally stationary. However, field bias is clearly visible in CBCT reconstructions in the form of slowly changing shading irregularities. Second order discrete differential processing will remove bias in the form of both linear and quadratic trends. For 2D images this is readily achieved by applying a digital Laplace filter, which is a 3x3 kernel with central value 8 and peripheral values set to −1. In 3D this becomes a 3x3x3 kernel with central value 26 and peripheral values of −1.

$ApEn_{2D}$ and $ApEn_{3D}$ were calculated for the raw tomographic images and also after applying this differential processing. Here it is noted that finite differencing of this nature is a discrete analogue of differentiation that sharpens detail. However, it can reduce signal to noise ratio, especially where reconstruction images already have high frequency noise content that is characteristic of filtered back projection. Consequently, after stationarization one would expect higher ApEn values that reflect noise amplification in both CT and CBCT.

4. Results and Discussion

The following sub-sections correspond to the staged approach described in the methodology section-3.

4.1 Functional consistency between the original heuristic $ApEn_{1D}$ and $ApEn_{3D}$

Optimum k was computed for $ApEn_{1D}$ and $mApEn_{3D}$ using sequences of uniformly distributed random numbers. For m=1 these are revealed by the peaks in figure 3(upper left) at 0.06 and 0.10 respectively. It can be seen that both definitions of approximate entropy show reassuringly similar behaviour with respect to k although the divergently modified $mApEn_{3D}$ exhibits usefully higher values than the conventional forward-only $ApEn_{1D}$.

The functional variation of ApEn with k is due to two competing effects in the source-target matching process. At small k the matching tolerance is severe and so very few matches are found. Consequently, the bias due to self matching is significant and reduces ApEn. At large k the sensitivity of the statistic is lost as large variations fall within the comparison threshold. This means the number of extended matches increases until it is similar to the number of initial matches. Hence ApEn is decreased as the sequence appears more

regular. Figure 3(upper right) shows how the two approximate entropies vary with N, as calculated for the optimum values of k of 0.06 and 0.1. Convergence is approached for small N. The reduction of ApEn values as N decreases is a result of increasing bias due to counting of self matches in the calculation. Counting self matches has the effect of adding 1 to the number of matches both of length m and length m+1 for each source point. See Richman *et. al.* (2000) for discussion of this bias. These results illustrate the wisdom of choosing r to maximize the range of computed ApEn. Furthermore, they show that with divergence incorporated into $mApEn_{3D}$ there is increased potential to do this.
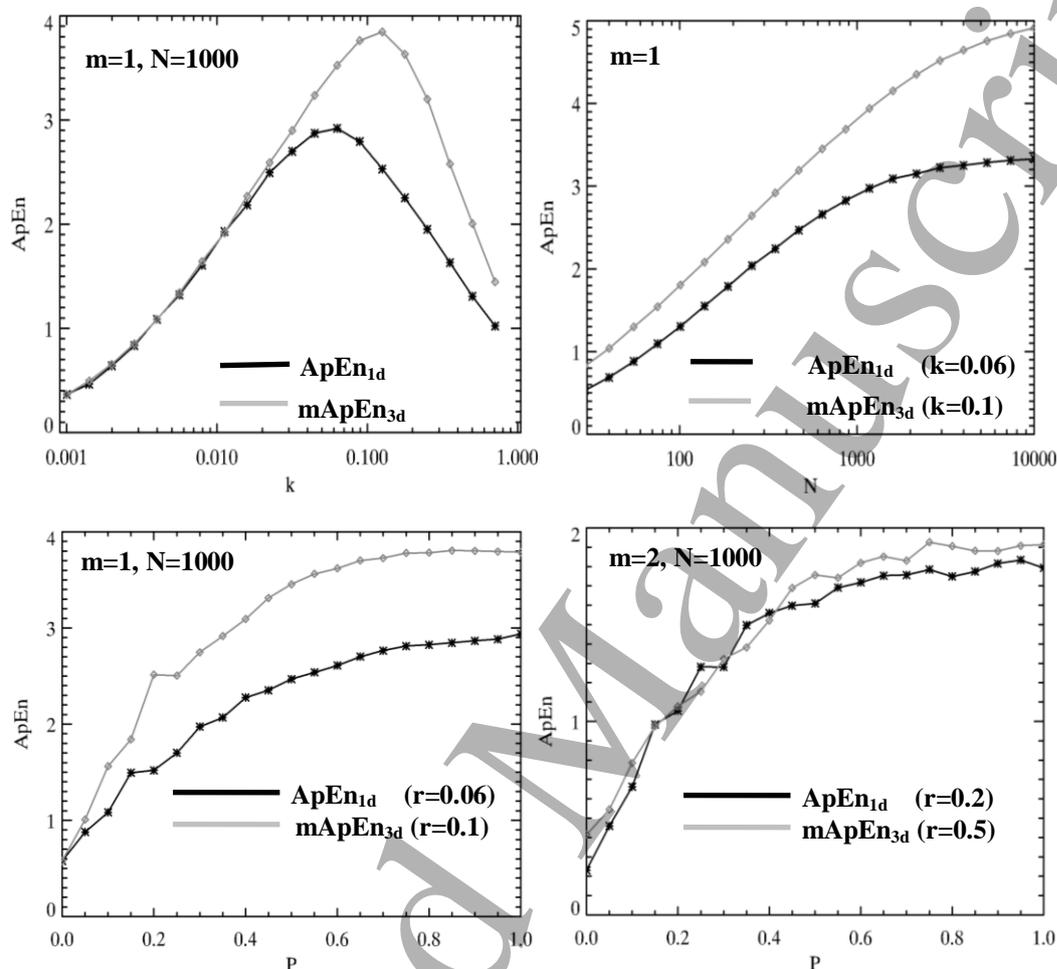
Figure 3. (upper) Comparison of conventional forward only and divergently modified approximate entropy
analogues for uniformly distributed random number sequences:
(left) variation with k (right) variation with N.
(lower) Comparison of conventional and modified ApEn for MIX(*p*) process:
(left) m=1 (right) m=2**.**

The results of degrading a structured series, in the form of a regular sinusoidal series by progressively adding random elements via MIX(*p*) and computing $ApEn_{1D}$ and $mApEn_{3D}$, are shown in figure 3(Lower). This is done for $N = 1000$ and $m = 1$ and $m = 2$, where r was selected to give the maximum value of ApEn for random number sequences, as above. Both versions of ApEn show similar steady increase as the proportion of randomness increases and leveling out above about *p*=0.6. This result and those beforehand show that the divergently modified definition of ApEn behaves similarly in 1D to conventional ApEn for both $m = 1$ and $m = 2$, with almost identical values calculated for $m = 2$. In particular the ability to distinguish the degree of regularity in a sequence is preserved.

4.2 Functional behaviour and range optimization of $ApEn_{2D}$ and $ApEn_{3D}$
Optimum k values for $ApEn_{2D}$ and $ApEn_{3D}$ were calculated for 2D and 3D uniformly distributed random number patterns by varying threshold parameter, k, for $D = 3$ and $D = 5$. The results are shown in figure

4(Upper). Optimum k gives maximum ApEn and these values occur at around k = 0.6-0.7 for $ApEn_{2D}$ and around k = 0.7-0.8 for $ApEn_{3D}$.
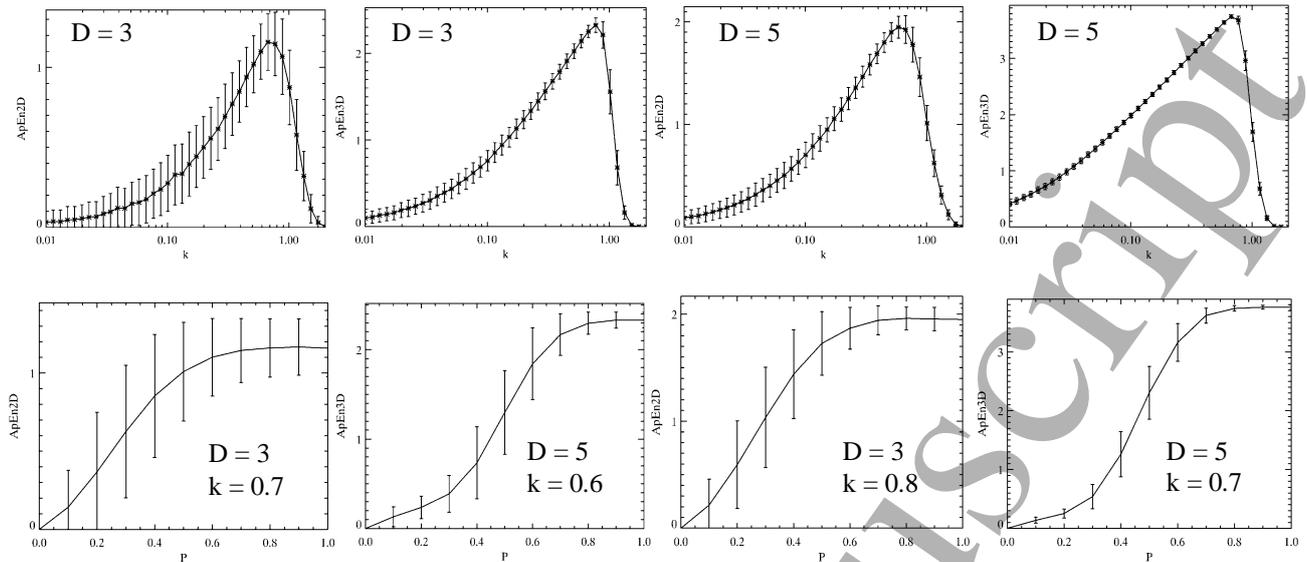
Figure 4. (upper) $ApEn_{2D}$ and $ApEn_{3D}$ for uniformly distributed random number patterns as a function of threshold parameter k for both D=3 and D=5.
(lower) $ApEn_{2D}$ and $ApEn_{3D}$ for a regular pattern-3 with increasing proportion of random noise p for both D=3 and D=5 at optimum k.

The results of degrading structured pattern-3 in the form of 2D and 3D chequer boards by increasing random content are shown in figure 4(lower). $ApEn_{2D}$ and $ApEn_{3D}$ increase steadily as the proportion of irregularity in the test patterns increases. The standard deviation of ApEn values is relatively large for D=3 and smaller for D=5.

Table 1. $ApEn_{2D}$ and $ApEn_{3D}$ (m = 1, k = 0.8, D = 3) for regular and irregular test patterns.

Input	$ApEn_{2D}$	$ApEn_{3D}$
1, Uniform array of zeros	0.0	0.0
2. Binary Step	0.0	0.0
3. Binary Chequer Board, 1 pixel squares	0.0	0.0
4. Binary Diagonal Stripes, 2 pixels wide	0.68	0.69
5. Binary Random 0s and 1s*	1.14 ± 0.24	2.13 ± 0.18
6. Gaussian Random, mean 1000, SD 50*	1.11 ± 0.21	2.27 ± 0.12
7. Uniform Random (range 0-10)* * changing range makes no difference to ApEn	1.13 ± 0.21	2.33 ± 0.09

$ApEn_{2D}$ and $ApEn_{3D}$ were calculated for simple, synthetic test images whose content ranged from regular through to randomised (examples in figure 2). The results of processing with dimension D = 3, m = 1 and k = 0.8 are shown in table 1. The uniform, chequerboard and step patterns have $ApEn_{2D}$ and $ApEn_{3D}$ both at zero. The random patterns have higher ApEn, in the range 1.11 to 2.33. The binary greyscale diagonal stripe pattern has intermediate ApEn of 0.68-0.69. Hence, the 2D and 3D algorithms can distinguish regular from irregular image patterns regardless of greyscale range and content.

4.3 Processing of Real World Volumetric Data
$ApEn_{2D}$ and $ApEn_{3D}$ were computed for clinical CT and CBCT image volumes of the pelvis acquired for image guided radiotherapy (IGRT). An ApEn value was calculated at each image voxel for D+2, as

described in section 2.3. Point by point results were also recalculated for the images after stationarization by Laplace filtering a.k.a. finite differencing. Finally, ApEn results were analyzed collectively in histogram form to test the ability to distinguish scanning modalities.
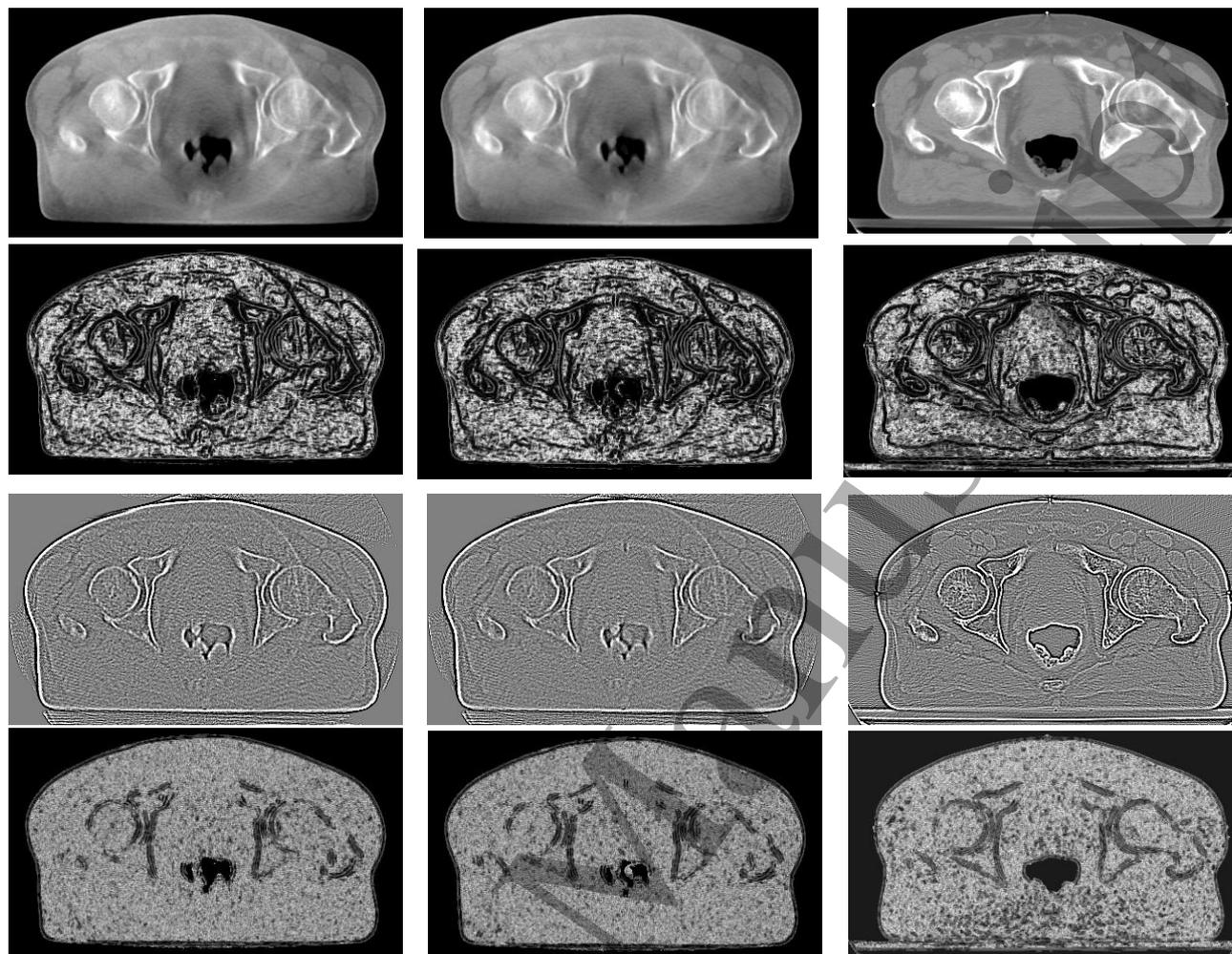
Figure 5.
(1st Row) (left) native section 1mm thick from CBCT   (centre) 5mm equiv' CBCT   (right) 5mm thick CT.
(2nd Row) $ApEn_{2D}$ for 1st row images, calculated for m=1, k=0.7, D=3, display window 0.0-1.5.
(3rd Row)   Finite difference results for 1st row images.
(4th Row) $ApEn_{2D}$ for 3rd row differenced images, calculated for m=1, k=0.7, D=3, display window 0.0-1.5.

4.3.1 Planar Processing of Clinical Images and $ApEn_{2D}$. Figure 5 shows native CT and CBCT pelvis images and the results of $ApEn_{2D}$ processing in rows 1 and 2 respectively. In comparison to the conventional CT image, it can be seen that the CBCT contains greater noise, reduced contrast and reduced uniformity. This is a result of increased X-ray scatter in the CBCT image, narrower slice width in the full resolution CBCT (1 mm vs 5 mm) and detector non-ideality. The histogram of $ApEn_{2D}$ values for the three images is shown in figure 6(left). Note that the histogram is a global representation of 2D ApEn distribution, which in the raw image state provides no differentiation when the entire image field is considered.

The result of finite differencing the three native tomographic images and the resultant 2D ApEn mappings is shown in rows 3 and 4 of figure 5. The corresponding global histograms of 2D ApEn are shown in figure 6(right). For the latter, the CBCT images now have a distinctive peak at a high ApEn of 1.2, which is indicative of structural paucity. The CT peak has much more power at lower ApEn values, especially those associated with structure: table-1 indicates stripe structures registering as 0.7. Additionally, CT and to a lesser extent CBCT have a secondary spike at ApEn=0.5 for both raw and finite differenced ApEn histograms. Collectively, this is signalling that CT has quantifiably higher meaningful structural content than CBCT though the more obvious boundaries are prominent.
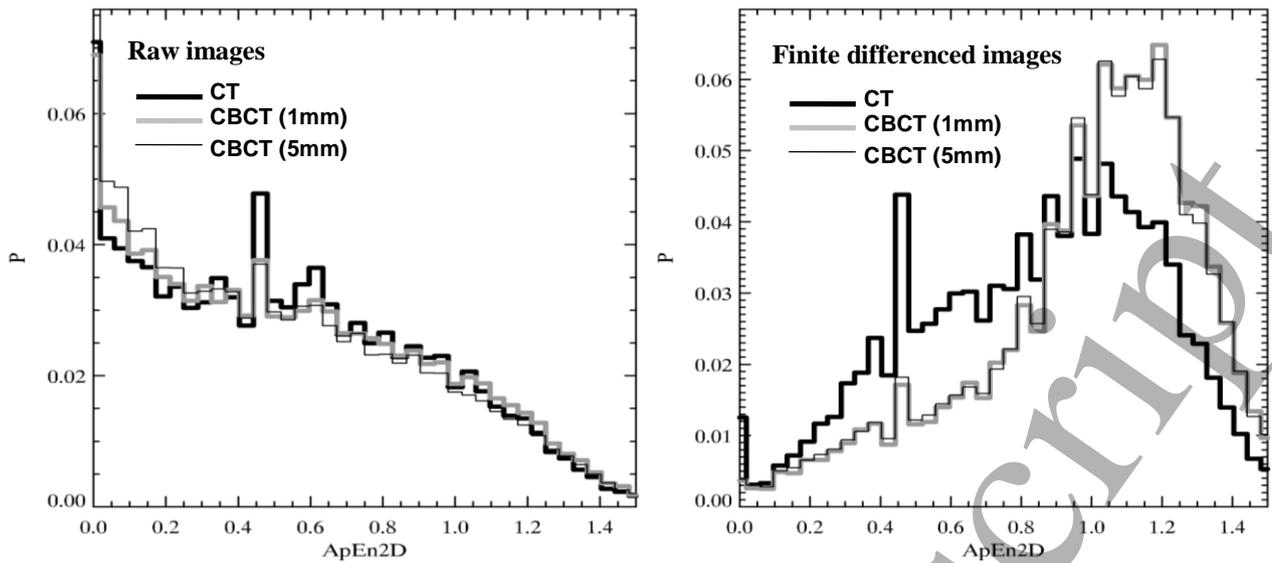
Figure 6. Histograms showing distribution of $ApEn_{2D}$ values for CT, CBCT (1mm slices), and CBCT (5mm slices): (left) raw image data   (right) 2D finite differenced.
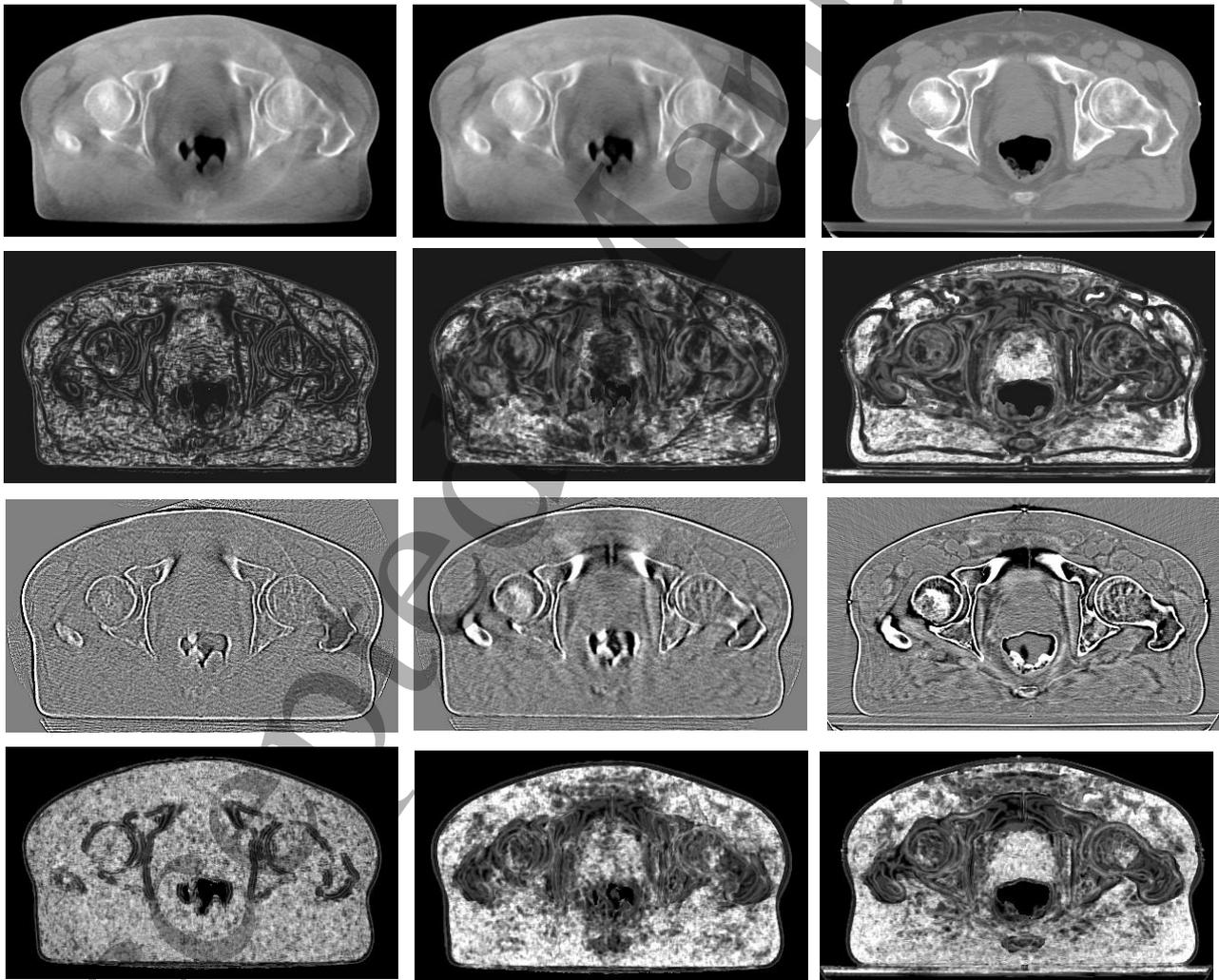
Figure 7.

(1st Row)   (left) native section 1mm thick from CBCT   (centre) 5mm equiv' CBCT   (right) 5mm thick CT

(2nd Row) $ApEn_{3D}$ for 1st row images, calculated for m=1, k=0.8, D=3, display window 0.0-1.5.

(3rd Row)   Finite difference results for 1st row images.

(4th Row) $ApEn_{3D}$ finite difference volumes displayed with window 0.0-2.0 calculated for m=1, k=0.7, D=3**.**

4.3.2 Volume Processing of Clinical Images and ApEn$_{3D}$. Figure 7 shows native CT and CBCT pelvis images and the results of ApEn$_{3D}$ processing in rows 1 and 2 respectively. Processing in 3D brings information in the immediate foreground and background planes of the image being viewed into play, with 3D ApEn clearly able to provide a quantitative map of structural content. In particular note the wealth of structure in the CT volume, with a clearly calibrated grey scale range. Similarly note the clear advantage of matching the CBCT slice width to the underlying CT stack, which is composed of consecutive 5mm voxels. The histogram of ApEn$_{3D}$ values for each of the three images is shown in figure 8(left). Unlike the 2D ApEn case, CT is clearly distinguishable from both CBCT examples.

The result of finite differencing the three native tomographic images and the resultant 3D ApEn mappings is shown in rows 3 and 4 of figure 7. The corresponding global histograms of 3D ApEn are shown in figure 8(right). The impact of stationarization on global 3D ApEn histograms is clearly to shift content in favour of the lower values. This is pronounced for the 5mm CT and CBCT voxel examples, which reach a maximum at a structurally significant 0.5 (see table 1). For the 1mm voxel CBCT example the peak at 0.5 is secondary to that at 1.2 but low ApEn content is clearly higher than that for the planar 2D case.
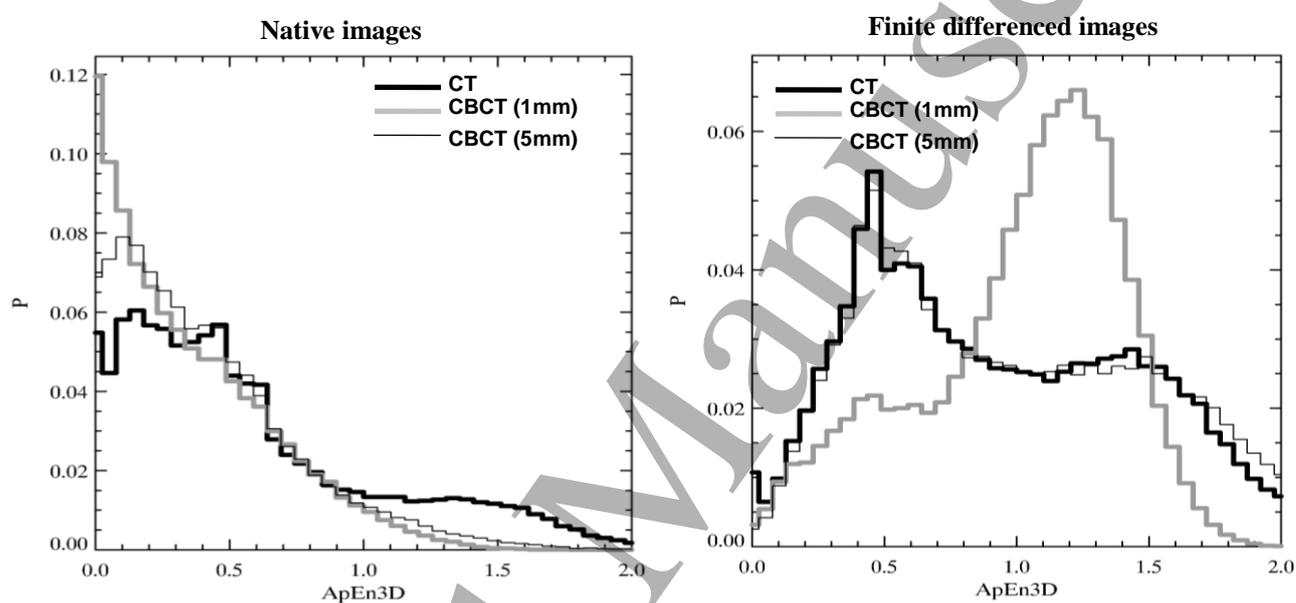
Figure 8. Histograms showing distribution of ApEn$_{3D}$ values for CT (volumes comprising 5mm slices), CBCT where
volumes comprise 1mm slices, and CBCT, where volumes comprise 5mm equivalent slices:
(left) raw image data   (right) 3D finite differenced.

For convenience we have shown 3D ApEn numerical results rendered collectively in a manner that visually indicates the weight of structural content, point by point by across a CT reconstruction field. With this form of presentation the correspondence between myriad ApEn numerical and CT attenuation values can be conveniently appreciated across the entire image field, albeit in a subjective manner. Each form of presentation displays an entirely different attribute, structure and attenuation respectively. Unlike the latter, which has exclusively in-plane relevance, 3D-ApEn shows the strength of structural connectivity to the underlying and overlying axial planes in addition to that in-plane. Their complementary use could be favourable if enhanced ways of CT presentation are being sought. Figures 5 and 7 can be viewed in this light. Being 2D, the former solely reflects in-plane image attributes, a cut plane of attenuation and cropped structure, whereas the latter strongly reveals the 3D weave of structure as it moves into and out of the axial attenuation section.

Following from this, we anticipate it may be possible to provide a quantitative indication of the inherent structural content of a patient image and establish its fitness for clinical purpose. In addition to the conventional QA testing with phantoms, this could have useful implications for clinical practice where it is dependent on image quality. That includes increasing confidence that we are indeed generating images that can support effective delineation and targeting in radiotherapy.

An immediate application will be to assess the improved cone-beam CT images we have developed to make them fit for aggressive targeting & adaptive dosimetric re-planning of photon and proton radiotherapy from in-treatment wide angle X-ray cone beam CT image volumes (Marchant, 2014). Thereafter, application to predictive outcome modeling in radiotherapy that includes diagnostic and treatment imaging parameters is envisaged via the 'ukCAT' project at The Christie hospital in Manchester (Price 2016).

Finally, we note that D and m are helpfully well defined, 3 and 1 respectively, as shown by the results for real world images (figures 5 & 7). Whilst we have presented detailed experimental work investigating a wide range of k values, in fact the results we obtained were pleasingly consistent with those from our previously published work on 1D approximate entropy applied to time series data. The k values providing robust operability are 0.6-0.7 in 2D and 0.7-0.8 in 3D, as reported. Maximum ApEn range is of course not essential to achieve for each and every patient image, but these values are expected to provide more than adequate generic operation.

5. Conclusion

We have presented what we believe to be the first, fully 3D approximate entropy algorithm, $ApEn_{3D}$. Its behaviour is consistent with the established heuristic forward-only $ApEn_{1D}$ algorithm when applied to 1D data series in appropriately modified form, $mApEn_{3D}$. This is true for both random and progressively degraded sinusoidal series. In particular the ability to distinguish the degree of regularity in a sequence is preserved. We have then shown that the 3D algorithm for $ApEn_{3D}$ is similarly consistent with the previously described 2D algorithm $ApEn_{2D}$ (Moore 2016) for test patterns progressively ranging from random through to structured.

Both 2D and 3D ApEn algorithms are capable of quantitatively distinguishing fan beam from cone beam CT cross sectional imaging modalities by structural content. In terms of the latter, the need for matching effective slice widths between the modalities in any comparison is clearly demonstrated. Where CT data comes in the form of a volume, the modality distinction is amplified and the structural content thrown into stark relief by $ApEn_{3D}$, which has the potential to quantify differences in detail at both global and local levels. Particularly in respect of addressing CT soft tissue detail, where noise is prominent, $ApEn_{3D}$ visibly outperforms common gradient processing. This is initial evidence that multi-dimensional $ApEn_{3D}$ may indeed be a useful quantitative, self calibrated parameter for both expert and machine interrogation of medical images.

References
Cullen J, Saleem A, Swindell R, Burt P and Moore C 2010 Measurement of cardiac synchrony using approximate entropy applied to nuclear medicine scans *Biomedical Signal Processing and Control* **5(1)** 32-36

Kolmogorov A 1958 A new metric invariant of transient dynamical systems and automorphisms in Lebesque spaces *Dokl. Akad. Nauk SSSR* **119** 861-864

Latora V and Baranger M 1999 Kolmogorov-Sinai entropy rate versus physical entropy *Physical Review Letters* 1999;82:520

Marchant T and Moore C J 2014, Making cone beam CT imaging fit for aggressive targeting & adaptive re-planning of photon and proton radiotherapy *Medical Research Council* grant MR/L023059/1

Moore C 1988 Mathematical analysis and encoding techniques applied to large archives of digital images, *PhD Thesis* University of Manchester

Moore C J, Manickam K, Slevin N 2006 Collective spectral pattern complexity analysis of voicing in normal males and larynx cancer patients following radiotherapy *Biomedical Signal Processing and Control* **1** 113-119

Moore C J 2016 A threshold structure metric for medical image interrogation: the 2D extension of approximate entropy *Proc 20th Intl Conf Information Visualization (IV2016, Lisbon)* **00** 336-341 http://doi.ieeecomputersociety.org/10.1109/IV.2016.30

Ngan PM 1997 Motion detection using approximate entropy *DICTA/I VCNZ97* Massey University New Zealand 379-384

Parker G, Schnabel J and Barker G 1999 Nonlinear smoothing of MR images using approximate entropy - a local measure of signal intensity irregularity *Information Processing in Medical Imaging* 484-489

Pincus S M 1991 Approximate entropy as a measure of system complexity *Proc Natl Acad Sci USA* **88** 2297-2301

Pincus S M and Goldberger A L 1994 Physiological time-series analysis: what does regularity quantify? *AJP - Heart and Circulatory Physiology* **266** H1643-H1656

Price G, Moore C, Johnson C and Finn C F 2016 Computer aided theragnostics in the UK (ukCAT): learning on globally distributed data", *Proc Ann Clinical and Scientific Computing Symp*, Birmingham, UK https://www.ipem.ac.uk/Portals/0/Documents/Conferences/2016/363%20Computing%20Symposium/363%20Computing%20Symposium%20programme%20and%20abstracts.pdf

Richman J S and Moorman J R 2000 Physiological time-series analysis using approximate entropy and sample entropy *AJP - Heart and Circulatory Physiology* **278** H2039-H2049

Singer B H and Pincus S 1998 Irregular arrays and randomization *Proc National Academy of Sciences of the United States of America* **95** 1363-1368

Stern D and Kurz L 1988 Edge detection in correlated noise using Latin square masks *Pattern Recognition* **21(2)** 119-129

Wolf A, Swift J B, Swinney H L and Vastano J A 1985 Determining Lyapunov exponents from a time series', *Physica*, **16D** 285-317