FRACTAL DIMENSION ESTIMATION IN DIAGNOSING ALZHEIMER'S DISEASE

VÁCLAV HUBATA-VACEK[a,*], JAROMÍR KUKAL[a], ROBERT RUSINA[b], MARIE BUNCOVÁ[c]

[a] *CTU, Faculty of Nuclear Sciences and Physical Engineering, Department of Software Engineering in Economics, Břehová 7, 115 19 Praha 1, Czech Republic*

[b] *Thomayer's Hospital, Vídeňská 800, 140 59 Praha 4–Krč, Czech Republic*

[c] *Institute for Clinical and Experimental Medicine, Vídeňská 1958/9, 140 21 Praha 4–Krč, Czech Republic*

[*] corresponding author: hubatvac@fjfi.cvut.cz

ABSTRACT. Estimated entropies from a limited data set are always biased. Consequently, it is not a trivial task to calculate the entropy in real tasks. In this paper, we used a generalized definition of entropy to evaluate the Hartley, Shannon, and Collision entropies. Moreover, we applied the Miller and Harris estimations of Shannon entropy, which are well known bias approaches based on Taylor series. Finally, these estimates were improved by Bayesian estimation of individual probabilities. These methods were tested and used for recognizing Alzheimer's disease, using the relationship between entropy and the fractal dimension to obtain fractal dimensions of 3D brain scans.

KEYWORDS: entropy, fractal dimension, Alzheimer's disease, boxcounting, Rényi entopy.

1. INTRODUCTION

Before explaning the relationship between entropy and dimension, we have to introduce the term of dimension. Let $d \in \mathbb{N}$ be a dimension of Euclidean space where a d-dimensional unit hypercube is placed. Let $m \in \mathbb{N}$ be resolution and $a = 1/m$ be edge the length of covering hypercubes of the same dimension d. The number of covering elements is given by

$$N = \mathrm{N}(a) = a^{-D}.$$

Knowledge of N for fixed a enables direct calculation of the hypercube dimension according to

$$\ln \mathrm{N}(a) = -D \ln a$$

$$D = \frac{\ln \mathrm{N}(a)}{\ln \frac{1}{a}}. \tag{1}$$

The very popular *boxcounting method* [1] is based on the generalization of (1) to the form

$$\ln \mathrm{N}(a) = A_0 - D_0 \ln a$$

and its application to the boundary of any set $F \subset \mathbb{R}^d$. As will be shown in the next section, the quantity $\ln \mathrm{N}(a)$ is an estimate of the *Hartley entropy*.

2. RÉNYI ENTROPY

Using a natural logarithm instead of a binary logarithm, we can follow in the definition of *Rényi entropy*. Let $k \in \mathbb{N}$ be number of events, $p_j > 0$ be their probabilities for $j = 1, \ldots, k$ satisfying $\sum_{j=1}^{k} p_j = 1$, and

$q \in \mathbb{R}$. We can define *Rényi entropy* [2] as

$$H_q = \frac{\ln \sum_{j=1}^{k} p_j^q}{1 - q},$$

which is a generalization of *Shannon entropy*. In respect of q, we obtain the specific entropies:

- *Hartley entropy* [3] for $q = 0$ as

$$H_0 = \ln \sum_{p_j > 0} 1 = \ln \sum_{j > 0}^{k} 1 = \ln k = \ln \mathrm{N}(a);$$

- *Shannon entropy* [4] for $q \to 1$ as

$$H_1 = \lim_{q \to 1} H_q = -\sum_{j=1} p_j \ln p_j;$$

- *Collision entropy* [2] for $q = 2$ as

$$H_2 = -\ln \sum_{p_j > 0} p_j^2;$$

The resulting theoretical entropies can be used for defining the *Rényi dimension* [2] as

$$D_q = \lim_{a \to 0+} \frac{H_q}{\ln \frac{1}{a}},$$

which corresponds to the relationship

$$H_q \approx A_q - D_q \ln a \tag{2}$$

for small covering size $a > 0$.

3. ENTROPY ESTIMATES

There are several approaches to entropy estimation from experimental data sets. Assuming that the number of experiments $n \in \mathbb{N}$ is finite, we can count the

events and obtain $n_j \in \mathbb{N}_0$ as the event frequencies for $j = 1, \ldots, k$. The first approach to entropy estimation is *naive estimation*. We directly estimate k and p_j as

$$k_\mathrm{N} = \sum_{n_j>0} 1 \leq k,$$

$$p_{j,\mathrm{N}} = \frac{n_j}{n}.$$

These biased estimates also produce biased entropy estimates

$$H_{0,\mathrm{N}} = \ln k_\mathrm{N},$$

$$H_{1,\mathrm{N}} = -\sum_{n_j>0} p_{j,\mathrm{N}} \ln p_{j,\mathrm{N}},$$

$$H_{2,\mathrm{N}} = -\ln \sum_{n_j>0} p_{j,\mathrm{N}}^2.$$

The second approach is based on *Bayesian estimation* of probabilities p_j as

$$p_{j,\mathrm{B}} = \frac{n_j + 1}{n + k_\mathrm{N}}.$$

This technique is called here *semi-Bayesian estimation*. We obtain other, but also biased, entropy estimates

$$H_{1,\mathrm{S}} = -\sum_{n_j>0} p_{j,\mathrm{B}} \ln p_{j,\mathrm{B}},$$

$$H_{2,\mathrm{S}} = -\ln \sum_{n_j>0} p_{j,\mathrm{B}}^2.$$

The estimate $H_{2,\mathrm{S}}$ can be improved as

$$H_{2,\mathrm{S}2} = -\ln \sum_{n_j>0} u_j,$$

where $u_j = \frac{(n_j+2)(n_j+1)}{(n+k_\mathrm{N}+1)(n+k_\mathrm{N})}$ is a Bayesian estimate of p_j^2. A *direct Bayesian estimate* of H_1 was also calculated as

$$H_{1,\mathrm{B}} = -\sum_{i=1}^{k_\mathrm{N}} \frac{n_i + 1}{n + k_\mathrm{N}} \big(\psi(n_i + 2) - \psi(n + k_\mathrm{N} + 1)\big),$$

where ψ is the digamma function.

4. BIAS REDUCTION

Miller [5] modified the naive estimate $H_{1,\mathrm{N}}$ using first order Taylor expansion, which produces

$$H_{1,\mathrm{M}} = H_{1,\mathrm{N}} + \frac{k_\mathrm{N} - 1}{2n}.$$

Lately, Harris [5] improved the formula to

$$H_{1,\mathrm{H}} = H_{1,\mathrm{N}} + \frac{k_\mathrm{N} - 1}{2n} + \frac{1}{12n^2}\left(1 - \sum_{p_j>0} \frac{1}{p_j}\right)$$

From the theoretical point of view, it is prohibited to estimate p_j by its estimates. However we are trying

to investigate biased estimates of H_1 in the forms

$$H_{1,\mathrm{HN}} = H_{1,\mathrm{N}} + \frac{k_\mathrm{N} - 1}{2n} + \frac{1}{12n^2}\left(1 - \sum_{n_j>0} \frac{1}{p_{j,\mathrm{N}}}\right),$$

$$H_{1,\mathrm{HS}} = H_{1,\mathrm{N}} + \frac{k_\mathrm{N} - 1}{2n} + \frac{1}{12n^2}\left(1 - \sum_{n_j>0} \frac{1}{p_{j,\mathrm{B}}}\right),$$

$$H_{1,\mathrm{HB}} = H_{1,\mathrm{N}} + \frac{k_\mathrm{N} - 1}{2n} + \frac{1}{12n^2}\left(1 - \sum_{n_j>0} r_j\right),$$

where $r_j = \frac{n+k_\mathrm{N}-1}{n_j}$ is Bayesian estimate of $\frac{1}{p_j}$.

5. ESTIMATION METHODOLOGY

Naive, semi-Bayesian, Bayesian and corrected entropy estimates were subjected of testing on 2D and 3D structures with known Hausdorff dimension. The list of involved estimates is included in Tab. 1. A Sierpinski carpet with $D_q = 1.8928$ for any $q \geq 0$ of size 81×81 is a typical 2D fractal set model. Using the estimates from Tab. 1 and a linear regression model (2), we estimated the Rényi dimensions $\hat{D}_q$ and then evaluated its z_score as a relative measure of bias

$$z_\mathrm{score} = \frac{\hat{D}_q - D_q}{S_{D_q}}.$$

The results are included in Tab. 2. The best estimations with $|z_\mathrm{score}| \leq 1.960$ are $H_{1,M}$ followed by Harris estimations $H_{1,HN}, H_{1,HS}, H_{1,HB}$. A structure of $D_q = 2.3219$ and size $128 \times 128 \times 128$ was then used for 3D testing and the results are also included in Tab. 2. The best estimators are $H_{1,HS}, H_{1,HN}, H_{1,HB}, H_{1,M}, H_{2,S}$.

6. ALZHEIMER'S DISEASE DIAGNOSIS FROM FRACTAL DIMENSION ESTIMATES

Alzheimer's disease (AD) is the most common form of dementia, and is characterised by loss of neurons and their synapses. This loss is caused by an accumulation of amyloid plaques between nerve cells in the brain. Morphologically, the affected areas produce rounded clusters of destroyed brain cells, which are visible on brain scans. On the other hand, Amyotrophic lateral sclerosis (ALS) is a disease of the motor neurons, and it is not visible on brain scans. In this sense, brain scans of ALS patien's look like brain scans of healthy patients.

These entropy estimators were used for diagnosing Alzheimer's disease. We tried to separate two different groups of samples of human brains. In the first group, there were brain scans of patients with Alzheimer's disease (AD) and in the second group brain scans of patients with amyotrophic lateral sclerosis (ALS). We carried out tests on 21 samples (11 for AD and 10 for ALS), represented by $128 \times 128 \times 128$ matrices of thresholded images ($\theta = 40\,\%$). We used a two-sample t-test for null hypotheses, and the alternative

Estimate	Sierpinski carpet $D_q = 1.8928$			Five Box Fractal $D_q = 2.3219$		
	$\hat{D}_q$	SD_q	z_{score}	$\hat{D}_q$	SD_q	z_{score}
$H_{0,\text{N}}$	1.8158	0.0064	-12.0577	2.0897	0.0284	-8.1757
$H_{1,\text{N}}$	1.8472	0.0059	-7.7116	2.1853	0.0320	-4.2690
$H_{2,\text{N}}$	1.8578	0.0076	-4.6212	2.1949	0.0298	-4.2568
$H_{1,\text{S}}$	1.8515	0.0058	-7.0853	2.2367	0.0315	-2.7012
$H_{2,\text{S}}$	1.8657	0.0072	-3.7494	2.2927	0.0298	-0.9798
$H_{2,\text{S2}}$	1.7898	0.0077	-13.4269	2.1189	0.0268	-7.5904
$H_{1,\text{B}}$	1.8170	0.0060	-12.6863	2.1654	0.0297	-5.2638
$H_{1,\text{M}}$	1.8930	0.0059	0.0306	2.3315	0.0349	0.2730
$H_{1,\text{HN}}$	1.8921	0.0059	-0.1203	2.3208	0.0347	-0.0332
$H_{1,\text{HS}}$	1.8921	0.0059	-0.1164	2.3226	0.0347	0.0196
$H_{1,\text{HB}}$	1.8920	0.0059	-0.1328	2.3182	0.0346	-0.1084

TABLE 2. Dimension estimates via various entropy estimates.

Method	H_0	H_1	H_2
Naive	$H_{0,\text{N}}$	$H_{1,\text{N}}$	$H_{2,\text{N}}$
semibayesian (p_j)	*	$H_{1,\text{S}}$	$H_{2,\text{S}}$
semibayesian (p_j^2)	*	*	$H_{2,\text{S2}}$
bayesian	*	$H_{1,\text{B}}$	*
Miller	*	$H_{1,\text{M}}$	*
Harris	*	$H_{1,\text{HN}}$	*
Harris semibayesian (p_j)	*	$H_{1,\text{HS}}$	*
Harris bayesian $(1/p_j)$	*	$H_{1,\text{HB}}$	*

TABLE 1. Entropy estimates.

Estimate	$\text{E}\hat{D}_q(\text{AD})$	$\text{E}\hat{D}_q(\text{ALS})$	p_{value}
$H_{0,\text{N}}$	1.9745	2.0315	0.017486
$H_{1,\text{N}}$	2.0649	2.1096	0.025128
$H_{2,\text{N}}$	2.0687	2.1034	0.067814
$H_{1,\text{S}}$	2.0968	2.1471	0.018828
$H_{2,\text{S}}$	2.1458	2.1903	0.031375
$H_{2,\text{S2}}$	1.9274	1.9666	0.036419
$H_{1,\text{B}}$	2.0011	2.0506	0.018873
$H_{1,\text{M}}$	2.2607	2.3115	0.041142
$H_{1,\text{HN}}$	2.2428	2.2931	0.037608
$H_{1,\text{HS}}$	2.2452	2.2957	0.037729
$H_{1,\text{HB}}$	2.2366	2.2868	0.035800

TABLE 3. Diagnostic power.

hypotheses were

$$\text{H}_0 : \text{E}\hat{D}_q(\text{AD}) = \text{E}\hat{D}_q(\text{ALS}),$$

$$\text{H}_\text{A} : \text{E}\hat{D}_q(\text{AD}) \neq \text{E}\hat{D}_q(\text{ALS}).$$

The results are included in Tab. 3. The most significant differences between AD and ALS were observed for $H_{0,\text{N}}, H_{1,\text{S}}, H_{1,\text{B}}$.

7. Conclusion

In this paper we tested estimates for Hartley, Shannon and Collision entropy. These estimates were improved by Bayesian estimation and tested on fractals with known fractal dimensions. Finally, these estimates were used on two groups of samples of brain scans, in order to obtain the best separator. The best separators, with regard to the experiment, are $H_{0,\text{N}}, H_{1,\text{S}}, H_{1,\text{B}}$, and they have a 2 % level of significance. The rest of the estimates also have results under a 5 % level of significance, except for $H_{2,\text{N}}$,

which was worst. On hte basis of these results, entropy can be used for diagnosing Alzheimer's disease in the future, considering that methods can be still improved, especially by estimating k_N or by image filtering.

References

[1] Theiler, J., *Estimating fractal dimension*. Journal of the Optical Society of America, Vol. 7, No. 6 1990, pp. 1055–1073.

[2] Renyi, A., *On measures of entropy and information*.

Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, 1961, p. 547.

[3] Hartley, R.V.L., *Transmission of information.* Bell System Technical Journal, Vol. 7, 1928, 535.

[4] Shannon, C.E., *A mathematical theory of communication.* Bell System Technical Journal, 1948.

[5] Harris, B., *The statistical estimation of entropy in the non-parametric case.* MRC Technical Summary Report, 1975.

[6] Gomez, C., Mediavilla, A., Hornero, R., Abasolo, D., Fernandez, A., *Use of the Higuchi's fractal dimension for the analysis of MEG recordings from Alzheimer's disease patients.* Medical Engineering & Physics, Volume 31, Issue 3, April 2009, pp. 306–313.

[7] Jouny, C.C., Bergey, G.K., *Characterization of early partial seizure onset: Frequency, complexity and entropy.* Clinical Neurophysiology, Volume 123, Issue 4, April 2012, pp. 658–669.

[8] Lopes,R., Betrouni, N., *Fractal and multifractal analysis: A review.* Medical Image Analysis, Volume 13, Issue 4, August 2009, pp. 634–649.

[9] Polychronaki, G.E., Ktonas, P. Y., Gatzonis, S., Siatouni, A., Asvestas, P. A., H Tsekou, H., Sakas, D. and Nikita, K.S., *Comparison of fractal dimension estimation algorithms for epileptic seizure onset detection.* Journal of Neural Engineering 7 (2010).