

Multiscale increment entropy: An approach for quantifying the physiological complexity of biomedical time series

Xue Wang, Xiaofeng Liu, Wei Pang, Aimin Jiang
Information Sciences 586 (2022) 279-293



Contents lists available at [ScienceDirect](#)

Information Sciences

journal homepage: www.elsevier.com/locate/ins



Multiscale increment entropy: An approach for quantifying the physiological complexity of biomedical time series



Xue Wang^{a,b}, Xiaofeng Liu^{a,c,d,*}, Wei Pang^e, Aimin Jiang^{a,c,d}

^a College of IoT Engineering, Hohai University, Changzhou 213022, PR China

^b School of Microelectronics and Control Engineering, Changzhou University, Changzhou 213164, PR China

^c Jiangsu Key laboratory of Special Robot Technology, Changzhou 213022, PR China

^d Changzhou Key Laboratory of Special Robot and Intelligent Technology, Changzhou 213022, PR China

^e School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, Scotland EH14 4AS, UK

Contents

Abstract

Introduction

- 理論
- 各種熵分析法回顧

Method

2.1 MSE、RCMSE

2.2 RCMPE

2.3 RCMDE

2.4 MIE

Experiments and results

4.1 Experiments on simulated data

- validation
- stability
- computational efficiency

4.2 Experiments on real physiological data

EKG

RR

OSV

Discussion

Conclusion

Abstract

★ multiscale increment entropy (MIE) 多尺度增量熵

繼承特性

增量熵 incremental entropy (IncrEn) :

- 考慮時間序列的波動方向和振幅
- 合成和真實信號實驗表明，MIE作為一個複雜性指標表現更好

多尺度分析 multiscale analysis :

- 為表徵生理信號特徵的重要方法
- 生理系統紀錄之時間序列資料通常在長時間尺度上呈現固有的生理複雜性變異

優點

- 計算時間短
- 變異性低
- 在各個尺度上都有定義，對於短時間序列也是
- 更可靠的區分能力
- 適合分析未知的生理時間序列

1. Introduction

Introduction - 理論

- 生理系統由跨多個時間尺度的複雜調節反饋環所控制，
- 呈複雜波動，反映了生理狀態和行為的變化，及許多無意識輸出。
- 健康人的生理信號表現出長程（分形）相關。
- **疾病和衰老的複雜性損失理論**：熵矩陣值與生理信號的複雜性隨著疾病和年齡增加而降低。

Introduction – 各種熵分析法回顧

ApEn 近似熵	測量時間序列的不規則性以表徵生理複雜性	<ul style="list-style-type: none"> ● 心臟間期之熵值不符合理論： 心房纖維顫動患者-高；健康個體-低。 證明不規則性與生理複雜性之間沒有直接相關性。 ● 無法正確反映生理時間序列中的動態變化
SampEn 樣本熵		
MSE 多尺度熵	將SampEn擴展到多個時間尺度、長程。 計算每個粗粒化時間序列的SampEn值。	<ul style="list-style-type: none"> ● 僅在部分尺度上符合理論。 部分因為SampEn對信號結構的描述不夠全面。 ● 尺度高、序列短時準確性降低。 ● 計算效率差。
CMSE 綜合多尺度熵	尺度 τ 下粗粒化時間序列的 τ 個SampEn值的平均值（所有滑動窗口重疊），以提高準確性。	<ul style="list-style-type: none"> ● 僅在部分尺度上符合理論。 ● 計算效率更差。 ● 增加了引起未定義熵的可能性。
RCMSE 精細綜合多尺度熵	用 τ 個粗粒化時間序列的匹配模板向量對的總數取代每個粗粒化時間序列的匹配模板向量，來計算負自然對數，以減少短時間序列的未定義問題。	<ul style="list-style-type: none"> ● 僅在部分尺度上符合理論。
MPE 多尺度排列熵	基於排列熵（PE）計算粗粒化時間序列的熵值，PE是一種基於時間序列的自然順序模式的符號動態量度，具有計算效率。	<ul style="list-style-type: none"> ● 未完全符合生理複雜性的概念。 可能是因為PE忽略了時間序列中相鄰元素的大小變化
RCMDE 精細綜合多尺度離散熵	計算效率提高 穩定性提高 識別能力提高	<ul style="list-style-type: none"> ● 在某些尺度上仍然存在與理論不一致的問題。
RCMPE 精細綜合多尺度排列熵		

2. Methods

MSE

2.1 MSE 多尺度熵分析

Multiscale entropy 是對不同時間尺度上時間序列複雜性進行評估的方法

• 1. coarse-grained producers 粗粒化

- 將各非重疊窗口 (長度為 τ) 內的數據點取平均

$$x_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} u_i, 1 \leq j \leq \lfloor \frac{L}{\tau} \rfloor = N,$$

τ = scale factor

N = length of the coarse-grained time series
with respect to τ , equal to the floor of L/τ .

• 2. 計算連續粗粒化時間序列 $x^{(\tau)}$ 的SampEn值

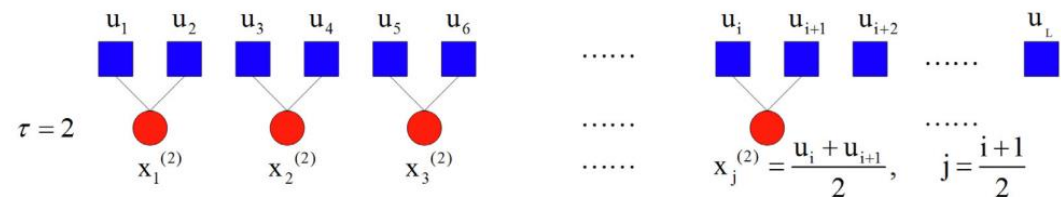
$$MSE(\mathbf{u}, \tau, m, r) = SampEn(\mathbf{x}^{(\tau)}, m, r) = -\ln \frac{n_{m+1}}{n_m}, \quad (\text{模板向量匹配對相除後取負自然對數})$$

r = similarity criterion 相似性標準，對每個粗粒化時間序列， r 保持恆定，等於第一個尺度上的值 (通常為原始信號的 $0.15 \cdot SD$)

n_m, n_{m+1} = number of m -dimensional and $(m+1)$ -dimensional template vector matching pairs derived from the time series $x^{(\tau)}$
with the distance between vector pairs less than r , respectively.

隨著尺度增加，MSE的準確性降低。MSE通常會引起未定義的熵。

一維離散原始時間序列： $\mathbf{u} = \{u(i), 1 \leq i \leq L\}$



連續的粗粒化時間序列： $\mathbf{x}^{(\tau)} = \{x_j^{(\tau)}\}$

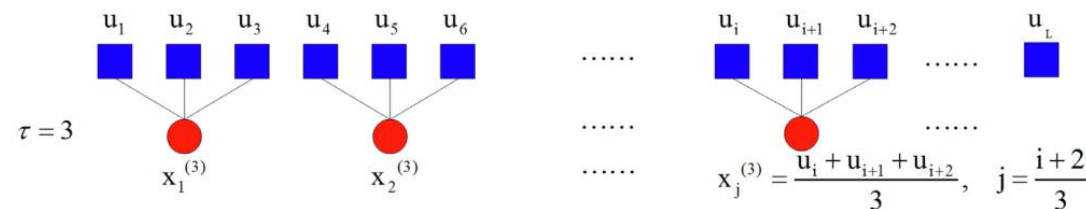


Fig. 1. Schematic illustration of the coarse-grained procedure. Modified from reference [32].

RCMSE, RCMDE, RCMPE

- 2.1 Refined Composite MultiScale Entropy (RCMSE)

MSE之變種：精細綜合多尺度熵

改進粗粒化過程：對尺度 τ 存在 τ 個不同粗粒化時間序列： k^{th} 粗粒化時間序列 $x_k^{(\tau)} = \{x_{k,1}^{(\tau)}, x_{k,2}^{(\tau)}, \dots\}$

$$x_{kj}^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{k+j\tau-1} u_i, 1 \leq j \leq \lfloor \frac{L}{\tau} \rfloor = N, 1 \leq k \leq \tau. \quad RCMSE(\mathbf{u}, \tau, m, r) = -\ln \left(\frac{\bar{n}_{m+1}^{(\tau)}}{\bar{n}_m^{(\tau)}} \right), \quad \bar{n}_{m+1}^{(\tau)} = \frac{1}{\tau} \sum_{k=1}^{\tau} n_{k, m+1}^{(\tau)}, \bar{n}_m^{(\tau)} = \frac{1}{\tau} \sum_{k=1}^{\tau} n_{k, m}^{(\tau)} \cdot n_{k, m}^{(\tau)}$$

- 2.2 Refined Composite Multiscale Dispersion Entropy (RCMDE)

精細綜合多尺度離散熵

DisEn將粗粒化的時間序列 $x(s,k)$ 映射為 $y = [y_1, y_2, \dots, y_N]$ ， N 是粗粒化時間序列的長度。

$$y_j = \frac{1}{\delta \sqrt{2\pi}} \int_{-\infty}^{x_{kj}^{(\tau)}} e^{\frac{-(t-\mu)^2}{2\delta^2}}, \quad DisEn(\mathbf{x}_k^{(\tau)}, m, c) = -\sum_{\pi=1}^{c^m} p_{k,\pi} \ln p_{k,\pi}. \quad RCMDE(\mathbf{u}, m, c, \tau) = -\sum_{\pi=1}^{c^m} \bar{p}_{\pi} \ln \bar{p}_{\pi},$$

- 2.3 Refined Composite Multiscale Permutation Entropy (RCMPE)

精細綜合多尺度排列熵

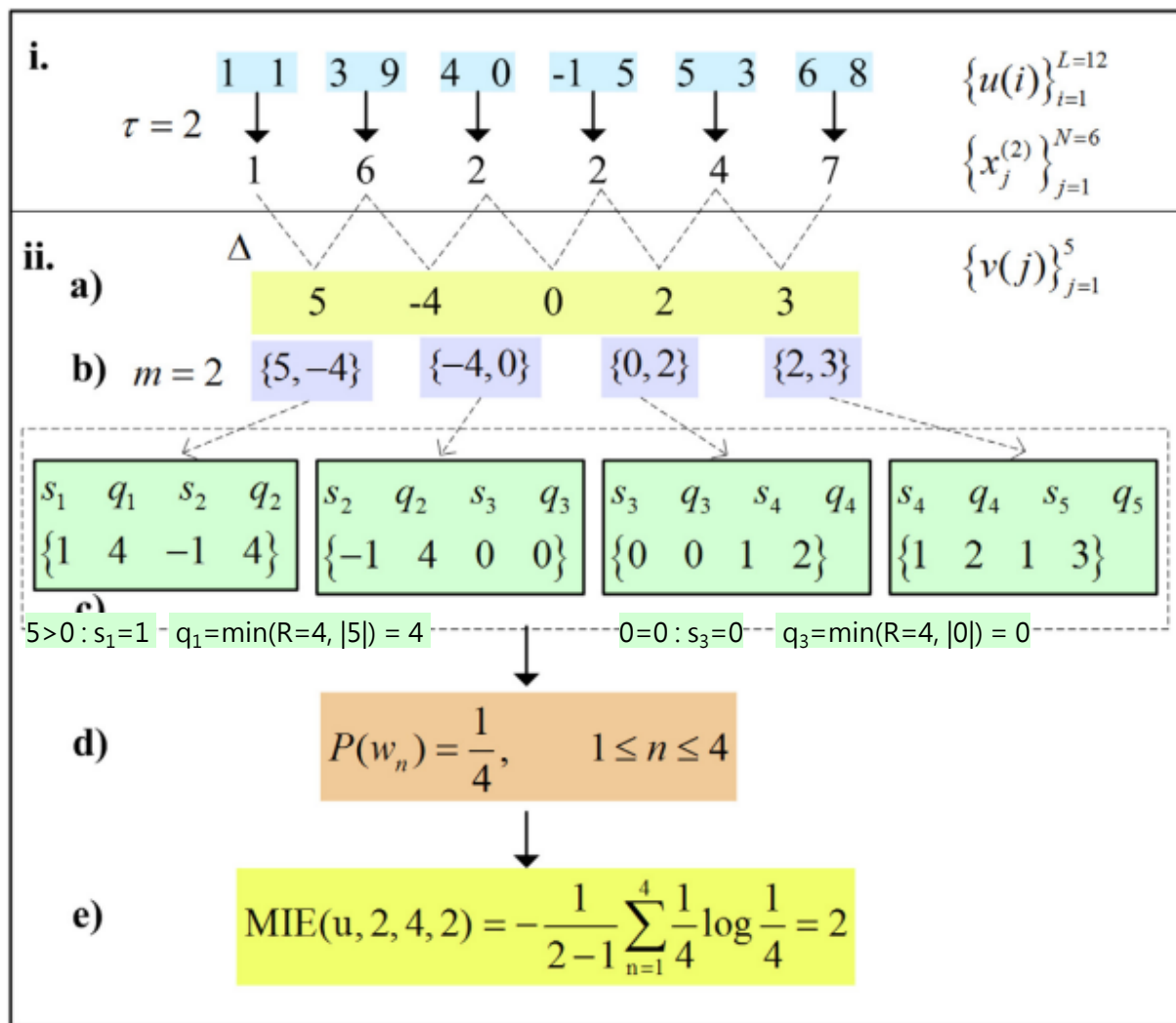
$$RCMPE(\mathbf{u}, m, d, \tau) = -\sum_{\pi=1}^{m!} \bar{p}_{\pi} \ln \bar{p}_{\pi},$$

MIE

2.4 the proposed MIE : i. 粗粒化 (多尺度) ii. 計算IncrEn值 (增量熵)

EXAMPLE

L (時間序列長度 length of the time series) =12, τ (尺度 scale) =2,
 m (維度 dimensionality) =2, R (量化精度 quantifying precision) =4



一維離散原始時間序列： $u = \{u(i), 1 \leq i \leq L\}$

依尺度取平均成連續的粗粒化時間序列： $x^{(\tau)} = \{x_j^{(\tau)}\}$

相減成增量序列 $v = \{v(j), 1 \leq j \leq N-1\}$, where $v(j) = x_{j+1}^{(\tau)} - x_j^{(\tau)}$

分成 N 個 m 維度增量向量 $V(k) = v(k), v(k+1), \dots, v(k+m-1), 1 \leq k \leq N-m$

s_k = 波動方向 of $v(k)$ $s_k = \text{sgn}(v(k))$, $v(k) > 0 \rightarrow 1 = \text{increase}$
 $v(k) = 0 \rightarrow = \text{no change}$
 $v(k) < 0 \rightarrow -1 = \text{decline}$

q_k = 波動大小 of $v(k)$ $q_k = \begin{cases} 0, & \text{step} = 0 \\ \min(R, \lfloor \frac{|v(k)| \times R}{\text{step}} \rfloor), & \text{step} \neq 0 \end{cases}$

頻率 relative frequency of each unique pattern vector $P(w_n) = \frac{Q(w_n)}{N-m}$,

W_n : nth unique pattern vector
 $Q(W_n)$: total number of instances for W_n

$$MIE(u, m, R, \tau) = \text{IncrEn}(x^{(\tau)}, m, R) = \frac{-1}{m-1} \sum_{n=1}^{(2R+1)^m} P(W_n) \log P(W_n)$$

4. Experiments and results

time delay = 1

embedding dimension $m = 2$ in MSE and RCMDE, and $m = 5$ in RCMPE.

parameter r equals 0.15 of the original signal SD in MSE.

number of classes c is set to 6 in the RCMDE.

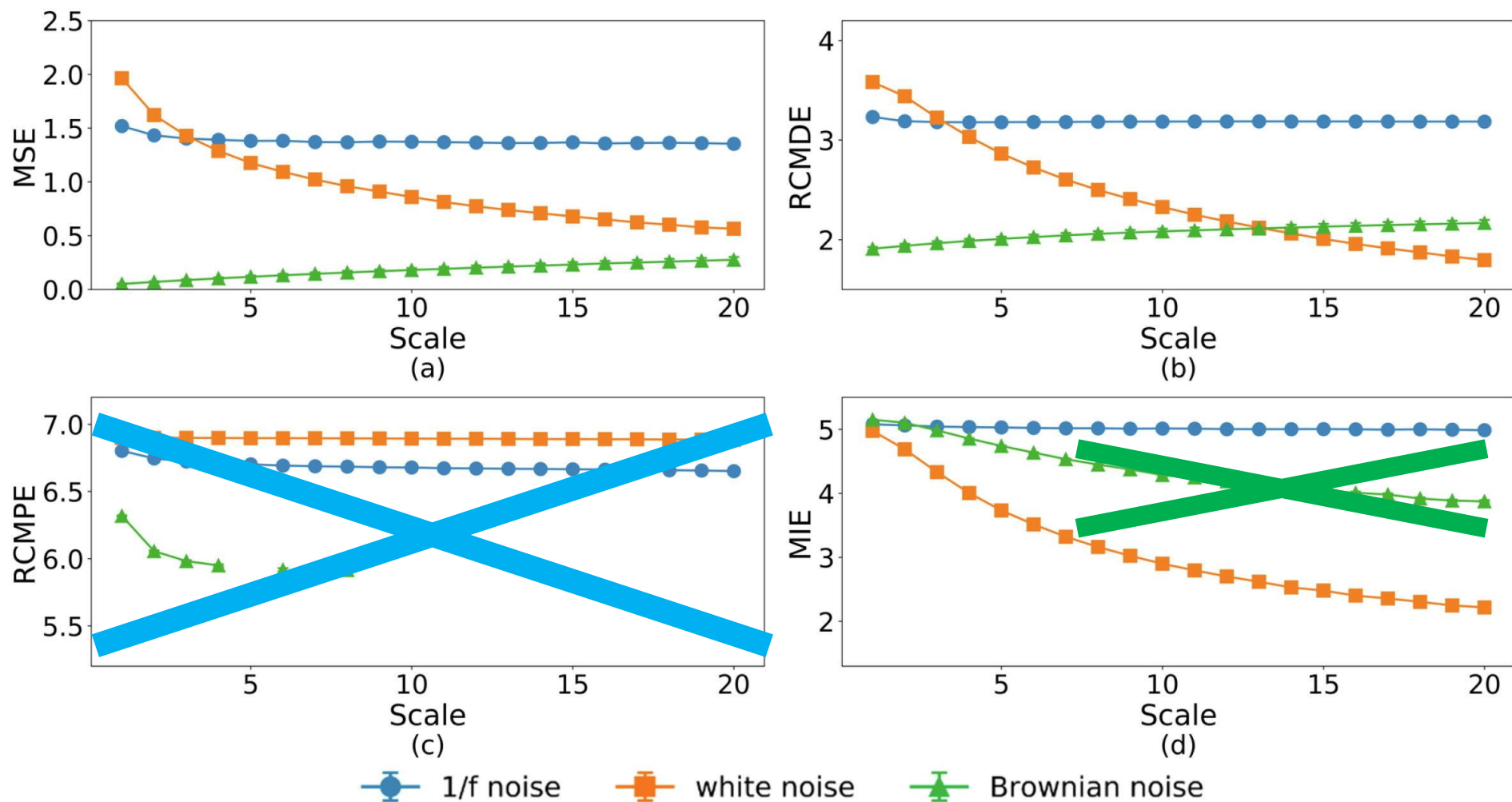
simulation is performed using a PC with an Intel (R) core (TM) i7-5500U CPU working at 2.4GHz and 4GB of RAM with MatLab R2014a.

Simulated time series

3.1 Simulated time series

1/f noise	健康	具有突出的分形特性 存在於許多具有長程相關的複雜動態的生理信號中
white noise	不健康	完全無相關的white noise和Brownian noise 是受損的生理系統最終可能達到的兩種動態終態
Brownian noise		

4.1 Experiments on simulated data - validation



1/ f noise 在多個時間尺度上包含更複雜的結構：熵值最高。

4.1 Experiments on simulated data - **stability**

- 變異係數 (CV) = 標準差 / 均值
- RCMDE和MIE的CV值均小於MSE的CV值。
- 對於MIE來說， $1/f$ noise和Brownian noise 的CV值最小。

Table 2

CV values of the MSE, RCMDE, and MIE algorithms from the analyses for $1/f$ noise, white noise, and Brownian noise (scale factor $\tau = 10$). For each time series type, the smallest CV value is reported in bold.

Time Series	MSE	RCMDE	MIE
$1/f$ noise	0.0325	0.0099	0.0058
white noise	0.0297	0.0128	0.0197
brownian noise	0.5501	0.0715	0.0166

4.1 Experiments on simulated data - computational efficiency

Table 3

The runtime data for the MSE, RCMDE and MIE algorithms used to analyse 30 different $1/f$ noise time series with different data length ($m = 2$). For each data length, the shortest runtime is indicated in bold.

Data Length	MSE	RCMDE	MIE
10,000	518.01 s	66.44 s	23.28 s
5,000	98.95 s	33.98 s	12.96 s
2,000	20.12 s	14.65 s	7.2 s
1,000	7.99 s	8.15 s	5.04 s

MIE的運行時間明顯快於MSE和RCMDE，尤其是對於較長的信號。

計算MSE所需的時間隨著時間序列長度的增加而顯著增加，但對於MIE來說情況不是這樣。

Real-world time series

3.2 Real-world time series

- 1) 來自健康受試者和癲癇患者的腦電圖 (EEG) 時間序列。
- 2) 來自三種不同生理狀態的心臟間距 (RRs) 時間序列。
- 3) 來自不同年齡個體的血氧飽和度變異性 (OSV) 時間序列。

Table 1

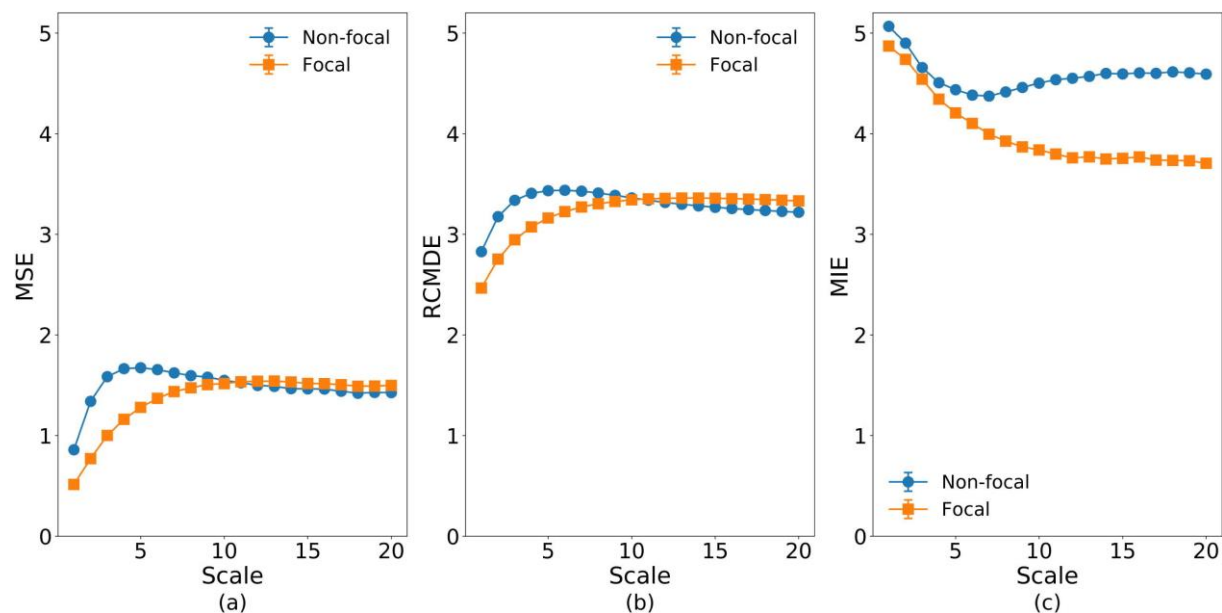
The basic information of the real-world datasets used in this study.

Datasets	Feature		Note
	The number of the time series	Data length	
EEG	5 healthy and 5 epileptic	4,700	health vs. epilepsy during seizure-free intervals
RR	18 healthy, 84 atrial fibrillation, and 15 congestive heart failure	20,000	health vs. the two different pathological states
OSV	20 young and 20 elderly	3,500	healthy young vs. healthy elderly

判斷：

- 應該正確反映時間序列中的動態變化，
在不同尺度上皆能明確顯示出差異，並在不同資料長度上保持一致性。
- 應該符合衰老和疾病的複雜性損失理論：健康或年輕者的熵值高於患者或老年者的熵值

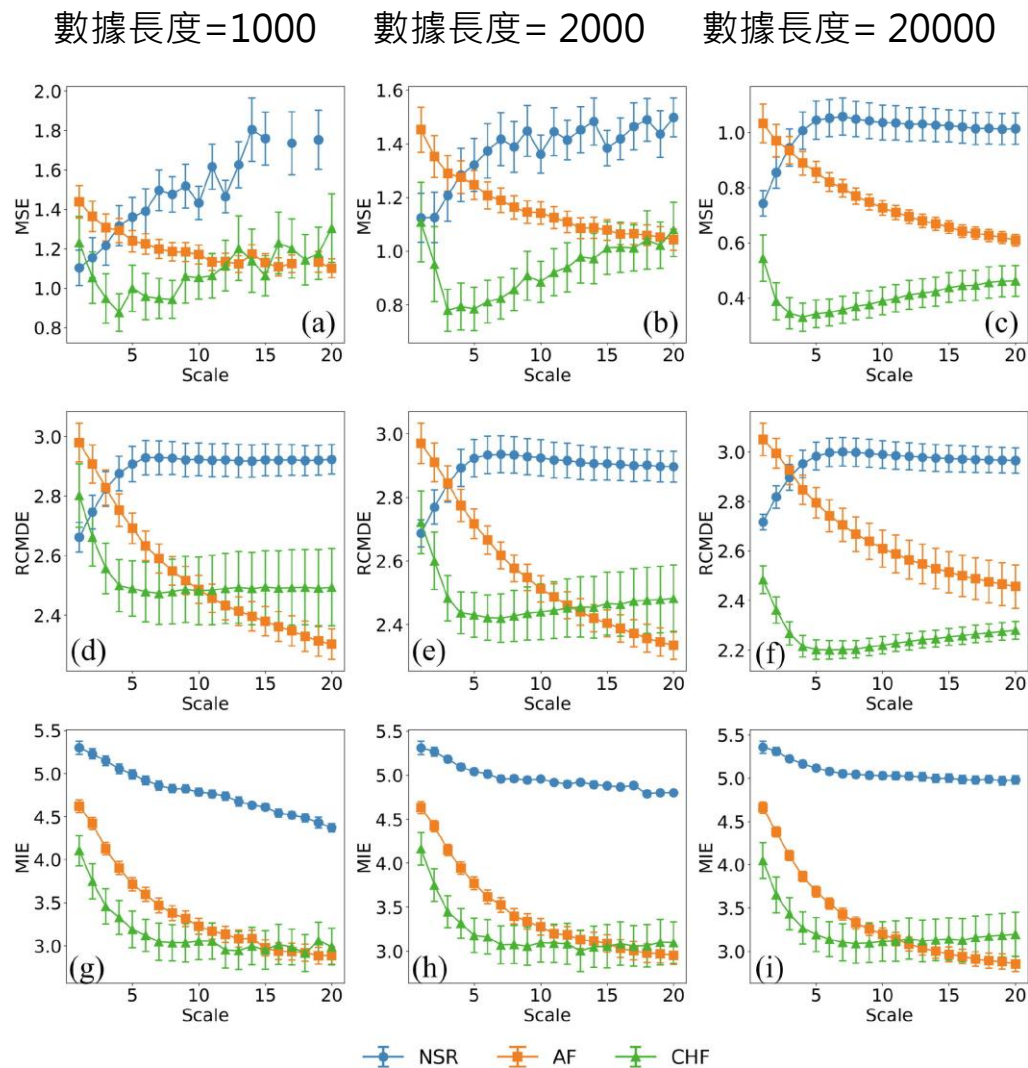
4.2 Experiments on real physiological data



腦電圖 (EEG)

- MSE和RCMDE在不同尺度上並不一致，隨尺度的增加而發生曲線交叉。
- MIE在每個尺度上皆
非癲癇受試者之值 > 癲癇受試者之值。

4.2 Experiments on real physiological data



心臟間距 (RR)

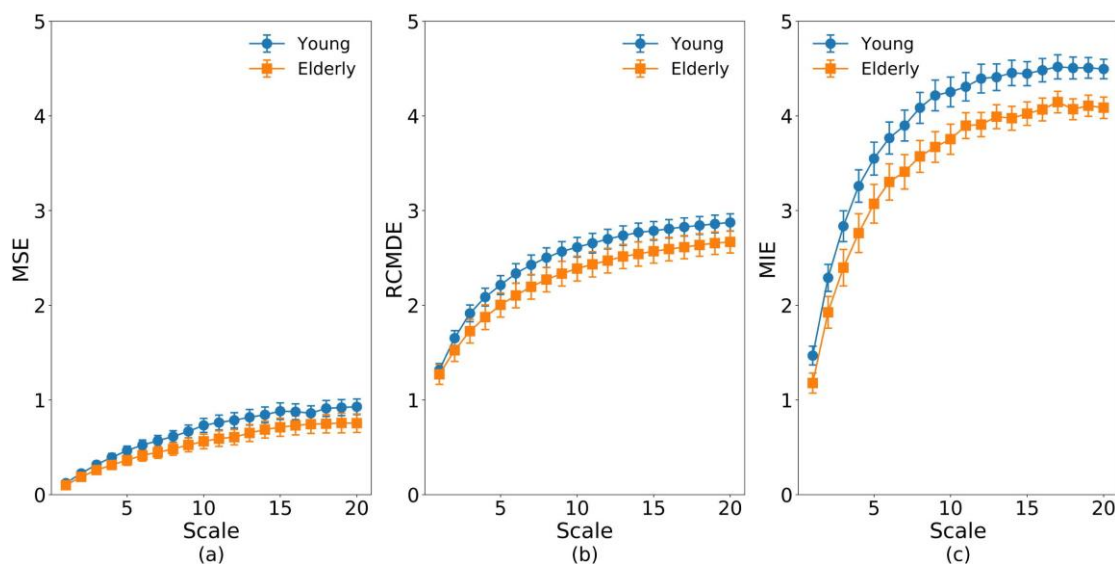
健康：間歇性心跳 (NSR)

不健康：心力衰竭 (CHF)、房顫 (AF)

- MIE能夠在所有尺度上清晰且一致地區分。
- MSE和RCMDE的結果差異在不同數據長度下不一致。
- MSE在數據長度為1000時在大尺度下出現未定義的熵，且在數據長度 ≤ 2000 時隨尺度增加變不穩定並波動。

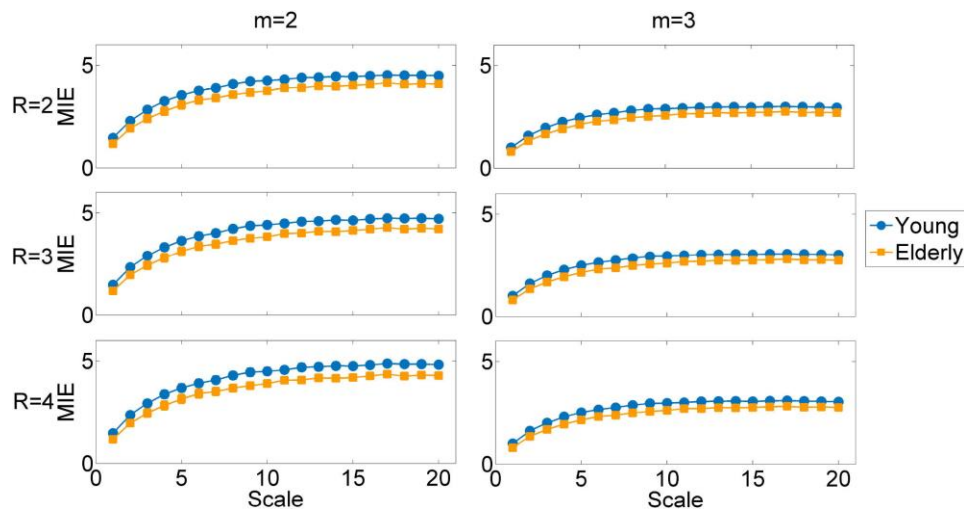
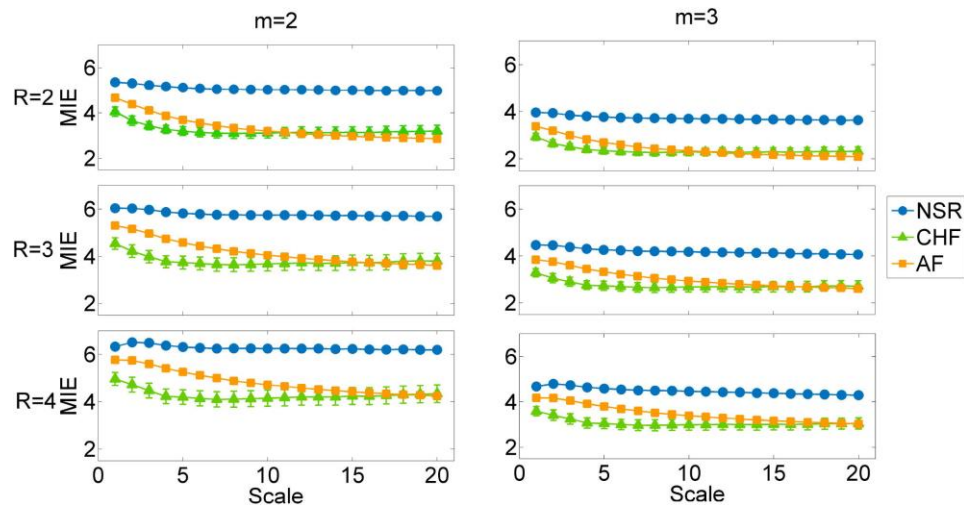
4.2 Experiments on real physiological data

血氧飽和度變異性 (OSV)



- MIE顯示出明顯差異
- MIE在不同尺度上保持一致的識別結果。
- 年輕人之值 > 老年人之值
- 曲線皆隨著尺度的增加逐漸增加，與先前的研究一致。

4.2 Experiments on real physiological data

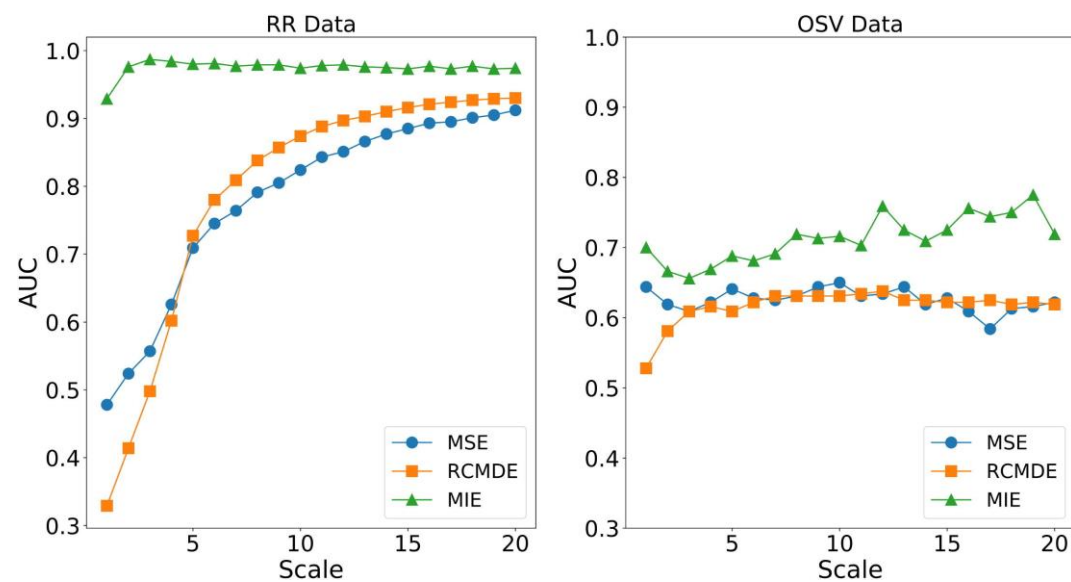


維度參數 m 和解析度參數 R

對MIE在RR和OSV數據集上的影響

- 熵值隨著 m 和 R 的增加而減少
這與參數 m 和 R 對IncrEn的影響相一致。
- 對較大的 R 或較小的 m ，兩組之間差異變明顯，因為MIE能夠識別出更多的時間序列模式。
- 在 $m=2$ ， $m=3$ 且 $2 \leq R \leq 4$ 的情況下，MIE呈現較好的相對一致性。

4.2 Experiments on real physiological data



ROC曲線下面積 (AUC) 值

MIE在任何尺度上產生更大的AUC值。

4.2 Experiments on real physiological data

25	Scale	p-values for Healthy vs. Patient			p-values for Young vs. Elderly		
26		MSE	RCMDE	MIE	MSE	RCMDE	MIE
27	1	0.004	0	0	0.287	0.696	0.054
28	2	0	0	0	0.34	0.351	0.122
29	3	0	0	0	0.286	0.213	0.115
30	4	0	0	0	0.205	0.182	0.091
31	5	0	0	0	0.166	0.198	0.066
32	6	0	0	0	0.186	0.167	0.092
33	7	0	0	0	0.156	0.17	0.08
34	8	0	0	0	0.152	0.169	0.043
35	9	0	0	0	0.159	0.164	0.062
36	10	0	0	0	0.125	0.18	0.041
37	11	0	0	0	0.145	0.188	0.078
38	12	0	0	0	0.131	0.174	0.023
39	13	0	0	0	0.179	0.18	0.046
40	14	0	0	0	0.213	0.16	0.055
41	15	0	0	0	0.202	0.178	0.034
42	16	0	0	0	0.271	0.175	0.012
43	17	0	0	0	0.356	0.174	0.043
44	18	0	0	0	0.212	0.177	0.013
45	19	0	0	0	0.226	0.18	0.007
46	20	0	0	0	0.179	0.169	0.02

單因素變異數分析

- 通過最小顯著差異檢驗來分析MSE、RCMDE和MIE的統計顯著性。
(≥ 0.05 : 綠 / < 0.05 : 紅)
- MIE是對健康的年輕和老年OSV組的三個指標中唯一具有顯著性的指標

使用Fisher’s discriminant function判別函數評估這些指標區分每組記錄的性能。
計算了各指標的整體準確度、敏感度和特異度。

所有統計檢驗均使用SPSS版本22.0.0.0進行。
<65：紅／65~85：黃／>85：綠

Indices	Scale	Healthy vs. Patient RR time s			Young vs. Elderly OSV time			Indices	Scale	Healthy vs. Patient RR time se			Young vs. Elderly OSV time se			Indices	Scale	Healthy vs. Patient RR time se			Young vs. Elderly OSV time se		
		Acc(%)	Sen(%)	Spe(%)	Acc(%)	Sen(%)	Spe(%)			Acc(%)	Sen(%)	Spe(%)	Acc(%)	Sen(%)	Spe(%)			Acc(%)	Sen(%)	Spe(%)	Acc(%)	Sen(%)	Spe(%)
MSE	1	81.2	50	86.87	61.11	55	68.75	RCMDE	1	77.78	66.67	79.8	55.56	55	56.25	MIE	1	64.1	88.89	59.6	63.89	65	62.5
	2	78.63	44.44	84.85	58.33	50	68.75		2	76.07	44.44	81.82	58.33	60	56.25		2	91.45	94.44	90.91	66.67	65	68.75
	3	58.12	50	59.6	61.11	50	75		3	68.38	27.78	75.76	61.11	65	56.25		3	94.87	100	93.94	66.67	70	62.5
	4	58.12	55.56	58.59	58.33	50	68.75		4	55.56	61.11	54.55	55.56	60	50		4	94.87	100	93.94	66.67	70	62.5
	5	58.97	55.56	59.6	61.11	50	75		5	58.12	66.67	56.57	55.56	60	50		5	93.16	100	91.92	66.67	70	62.5
	6	61.54	61.11	61.62	58.33	50	68.75		6	63.25	66.67	62.63	61.11	70	50		6	93.16	100	91.92	63.89	70	56.25
	7	62.39	61.11	62.63	63.89	55	75		7	66.67	66.67	66.67	61.11	70	50		7	93.16	100	91.92	66.67	70	62.5
	8	64.1	66.67	63.64	61.11	55	68.75		8	76.92	66.67	78.79	55.56	60	50		8	91.45	100	89.9	63.89	70	56.25
	9	66.67	66.67	66.67	61.11	50	75		9	82.05	72.22	83.84	55.56	60	50		9	88.89	100	86.87	61.11	70	50
	10	70.94	66.67	71.72	63.89	65	62.5		10	82.91	72.22	84.85	55.56	60	50		10	88.89	100	86.87	63.89	70	56.25
	11	76.07	72.22	76.77	61.11	55	68.75		11	85.47	77.78	86.87	55.56	60	50		11	88.89	100	86.87	61.11	75	43.75
	12	75.21	66.67	76.77	58.33	55	62.5		12	86.32	83.33	86.87	55.56	60	50		12	88.89	100	86.87	63.89	70	56.25
	13	82.91	72.22	84.85	61.11	60	62.5		13	85.47	83.33	85.86	52.78	60	43.75		13	88.89	100	86.87	63.89	70	56.25
	14	86.32	77.78	87.88	58.33	60	56.25		14	85.47	83.33	85.86	55.56	65	43.75		14	88.89	100	86.87	61.11	70	50
	15	89.74	77.78	91.92	58.33	55	62.5		15	86.32	83.33	86.87	52.78	60	43.75		15	89.74	100	87.88	66.67	70	62.5
	16	88.89	77.78	90.91	55.56	55	56.25		16	86.32	83.33	86.87	55.56	65	43.75		16	89.74	100	87.88	69.44	75	62.5
	17	89.74	77.78	91.92	58.33	60	56.25		17	86.32	83.33	86.87	55.56	65	43.75		17	88.89	100	86.87	63.89	75	50
	18	88.89	77.78	90.91	58.33	55	62.5		18	84.62	83.33	84.85	58.33	70	43.75		18	89.74	100	87.88	72.22	70	75
	19	88.89	77.78	90.91	63.89	65	62.5		19	83.76	83.33	83.84	55.56	65	43.75		19	89.74	100	87.88	75	75	75
	20	91.45	77.78	93.94	58.33	60	56.25		20	83.76	83.33	83.84	55.56	65	43.75		20	89.74	100	87.88	66.67	70	62.5

- MSE和RCMDE在大尺度上具有更好的識別能力。
- MSE和RCMDE在健康和患者RR時間序列上的準確度、敏感度和特異度均低於MIE的大多數尺度。
- MIE的性能優於從單一熵 (IncrEn , $\tau=1$) 衍生的指標。

5. Discussions

Simulated time series

MIE 在 Brownian noise 及 $1/f$ noise 的結果
在所有尺度上存在不一致性。

但對類似於 Brownian noise 的心力衰竭時間序列

MIE在所有尺度上都符合理論。

Real-world time series

MSE、RCMDE、MIE 中，

只有MIE在所有尺度上都符合理論。

	MSE	RCMDE	MIE
計算速度 實驗結果 與推測一致	最慢 兩個嵌套循環	中 細化在尺度 τ 上計算 τ 個粗粒化序列	最快
時間複雜性	$\sum_{\tau=1}^s O\left(\frac{N}{\tau}\right)^2$	$\sum_{\tau=1}^s \tau O\left(\frac{N}{\tau}\right)$	$\sum_{\tau=1}^s O\left(\frac{N}{\tau}\right)$
變異性	較大 對數據長度 ≤ 200 時 SampEn 對參數非常敏感	較大	小 對短時間序列（ $N=1,000$ ）在各尺度上都有明確定義粗粒化序列的數據長度小於100，而 IncrEn 可以用於分析至少包含100個數據點的信號。
穩定性	最差	相當 細化在尺度 τ 上計算 τ 個粗粒化序列	相當，MIE繼承了IncrEn的優點，在維度 $m = 2$ 和 $m = 3$ 以及分辨率 $2 \leq R \leq 4$ 下表現出一定的穩健性。

SampEn

僅通過評估兩 m 維度之對應點間最大距離判斷

IncrEn

將差分操作用於構建增量序列，對時間序列中的波動變化敏感。
通過量化相鄰元素間的方向和大小，自然地刻劃了由 m 個點構成的向量模式。

6. Conclusions

Merits of the proposed novel entropy metric MIE

- 不需要了解信號本身，不依賴尺度的選擇，適合未知生理信號的複雜性分析。
- 區分能力優於MSE及RCMDE
- 對參數的穩健優於MSE及RCMDE
- 計算時間快於MSE及RCMDE

Future research

將MIE擴展到

- 多變量場景multivariate scenarios (多通道傳感器multi-channel sensors收集時間序列)
- 二維圖像複雜性分析 two-dimensional image complexity analysis 。
- 其他時間變動的波動時間序列 time-varying volatility time series的分析應用 。

THANK YOU