

ZASTOSOWANIE METODY *RECURRENCE PLOTS* W ANALIZIE DANYCH POMIAROWYCH

STRESZCZENIE

W publikacji opisano sposoby analizy nieliniowych danych pomiarowych. Przedstawione zostały metody rekonstrukcji przestrzeni fazowej, skuteczne metody poszukiwania wartości zanurzenia wymiaru oraz czasu opóźnienia, a także graficzne metody analizy danych, polegające na konstruowaniu obrazów powrotu (*Recurrence Plot*) oraz wzajemnych obrazów powrotu (*Cross Recurrence Plot*). W artykule zaprezentowano również metodę analizy ilościowej sygnałów powracających (*Recurrence Quantification Analysis*), jako instrument pomocny do określenia parametrów otrzymanego obrazu (a poprzez to także analizowanego układu). Metody graficzne wykorzystano do klasyfikacji pojazdów samochodowych na podstawie sygnałów pomiarowych pochodzących z rzeczywistego systemu pomiarowo-archiwizującego.

Słowa kluczowe: traktor, trajektoria, rekonstrukcja przestrzeni fazowej, zanurzenie wymiaru, czas opóźnienia, wartość graniczna (progowa), sąsiedztwo, norma, obrazy powrotu, obrazy powrotów wzajemnych, analiza ilościowa sygnałów powracających

APPLICATION OF *RECURRENCE PLOTS* IN THE TIME SERIES ANALYSIS

This paper presents special methods of nonlinear time series analysis. Methods for computing the embedding dimension and delay time in phase space reconstruction are shown. A (Cross) Recurrence Plots (CRPs) is proposed as an effective tool for analysis of data series. Applying the CRPs to analysis of car-stream measurement data enables detection, recognition and classification of moving vehicles in traffic.

Keywords: attractor, trajectory, phase space reconstruction, embedding dimension, delay time, threshold, neighborhood, norm, correlation, (Cross) Recurrence Plot (CRP), Recurrence Quantification Analysis (RQA)

1. WSTĘP

Zdolność rozumnej obserwacji zjawisk zachodzących w otaczającej nas rzeczywistości, a w szczególności umiejętność ich analizy, jest szczególną cechą istot rozumnych. Od wieków ludzie dokonywali (bardziej lub mniej świadomie) najróżniejszych obserwacji. Dotyczyły one głównie otaczającego nas świata oraz zachowania naszego organizmu. Zachowane do naszych czasów starożytne opracowania naukowe ukazują wysoki stopień profesjonalizmu poparty interesującym warsztatem naukowo-metodycznym. Już wówczas wypracowano precyzyjne sposoby obserwacji oraz analizy otrzymanych danych. Naukowe metody detekcji i analizy danych pomiarowych ulegały w ciągu wieków licznym przeobrażeniom i kreatywnemu rozwojowi. Wpływ na to miały bez wątpienia zarówno coraz nowsze metody badawcze, poparte rozwijającym się aparatem matematycznym, jak i bez wątpienia szersze spektrum zagadnień podlegających analizie. Przyczyny takiego stanu rzeczy należy się doszukiwać przynajmniej na dwóch płaszczyznach. Po pierwsze jest to zasługa dłuższego okresu, w którym gromadzone były i archiwizowane dane pomiarowe. Implikuje to *ipso facto* lepszą perspektywę postrzegania procesu i zjawisk w nim zachodzących oraz umożliwia dostrzeżenie pojawiających się okresowości, ale i anomalii, zanurzonych w przestrzeni czasu. Po drugie na przestrzeni wieków coraz bardziej nagłać wydawało się znalezienie odpowiedzi na nurtujące ludzi pytania dotyczące zasad funk-

cjonowania ludzkiego organizmu, rozstrzygnięcia problemu deterministyczno-probabilistycznej natury (wszech)świata, jak i dalej – rozwiązanie problemów i zagrożeń związanych z występującymi chorobami i długością ludzkiego życia. Te dwa powody do dzisiaj nie straciły nic ze swej aktualności.

Momentem przełomowym w analizie danych pomiarowych stał się rok 1937, kiedy to angielski matematyk Alan Turing, pracując nad koncepcją obliczalności funkcji matematycznych, opisuje bardzo prostą maszynę logiczną, która dziś nosi nazwę maszyny Turinga [1]. Ten prototyp maszyny cyfrowej zapoczątkował okres wzmoczonej pracy nad systemami wspomagającymi człowieka w zbieraniu i przetwarzaniu danych. Wreszcie – drugie półwiecze minionego wieku to czas ekspansji systemów komputerowych i maszyn liczących, które w przełomowy sposób przyczyniły się do poprawy wydajności, poszerzenia horyzontów oraz rozwoju metodologii analizy danych pomiarowych.

Jedną z dziedzin, w której rozwój techniki komputerowej stał się swoistą cezurą, jest analiza procesów nieliniowych. Ilość i złożoność takich procesów jest zadziwiająca. Szereg z nich spotykamy w sferze nauk przyrodniczych. Zaliczyć do nich możemy m.in. procesy: fizjologiczne (analiza współczynnika zmienności akcji serca [2], analiza fal mózgowych), geologiczne (datowanie profili geologicznych), dotyczące biologii molekularnej (poszukiwanie konkretnych sekwencji kodu DNA) oraz analiza stanów pogody (turbulencje, pływy wiatrów, predykcja opadów) i inne.

* Doktorant na Wydziale Elektrotechniki, Automatyki, Informatyki i Elektroniki AGH w Krakowie

2. KLASYCZNE METODY REKONSTRUKCJI PRZESTRZENI FAZOWEJ

Analiza procesów dynamiki nieliniowej, ze względu na ich złożoność i zmienność w czasie, może przysparzać wiele trudności. Najczęściej celem owej analizy jest znalezienie (przy wykorzystaniu odpowiednio dobranego aparatu obliczeniowego) i zastosowanie w rzeczywistych układach adekwatnego modelu matematycznego. Do tego celu najczęściej wykorzystuje się aparat wykorzystujący analizę nieliniowych równań różniczkowych.

Dobrze znane metody analizy danych bazują m.in. na liniowej estymacji momentów i korelacji, oraz na zasadzie analizy składników. W ostatnich dekadach większość rozważań oparta była na klasycznej analizie liniowej, wzbogaconej o dodatkowe elementy zaczerpnięte z teorii chaosu. Próbowano także estymować nieliniowe pomiary i ich właściwości. Uwaga koncentrowana była najczęściej na informacji płynącej z wymiarowości przestrzeni fazowej, entropii układu czy też z wykładników Lapunowa, jako niosących konkretną wiedzę o właściwościach układu. Jednakże wiele metod nieliniowej analizy danych wymaga długich lub stacjonarnych serii czasowych [3]. Problem w tym, że cechy te trudno odnaleźć w otaczającej nas rzeczywistości. Metody, odpowiednie dla zbudowanych prototypowych modeli, okazują się nie w pełni adekwatne w odniesieniu do obiektów rzeczywistych, z jakimi mamy do czynienia w codzienności. Nie uwzględniają bowiem szeregu (często jednak nieprzewidywalnych) czynników, mających wpływ na wzajemne relacje pomiędzy poszczególnymi cechami opisującymi dany system. W związku z powyższym, wyniki takiej analizy muszą uwzględniać obarczenie błędem, którego często nie udaje się precyzyjnie określić. Za przykład takich zdarzeń może posłużyć szereg zjawisk przyrodniczych (pływy, prądy powietrzne i in.).

Rekonstrukcja przestrzeni fazowej, jako jedna z metod opisu systemu dynamicznego [24], zyskała uznanie i szerokie zastosowanie z powodu swej przejrzystości i możliwości estymowania za jej pomocą charakterystycznych właściwości systemów dynamicznych.

Zależności istniejące pomiędzy elementami systemu dynamicznego powodują, iż o dynamice całego układu decydują wszystkie jego składowe. W związku z tym trajektoria oryginalnej przestrzeni fazowej, zachowująca jej topologiczną strukturę, może być zrekonstruowana tylko z wykorzystaniem jednej obserwacji lub serii czasowej [4, 5].

Jednym z elementarnych zagadnień, związanym z analizą systemów dynamicznych, jest rekonstrukcja atraktora, rozumianego jako zbiór w przestrzeni fazowej, do którego w miarę upływu czasu zmierzają trajektorie rozpoczynające się w różnych obszarach przestrzeni fazowej. W wyniku przeprowadzonej analizy wielowymiarowego systemu, otrzymuje się skalarnie serie czasowe, opisujące obserwowaną przestrzeń fazową i trajektorię ruchu. Liczna literatura opisuje stosowane powszechnie techniki rekonstrukcji [m.in.: 4, 6]. Jedną z popularnych metod rekonstrukcji przestrzeni jest metoda opóźnień czasowych τ [5, 7]. Zasadniczą jest zaletą jej prostota i stały poziom zakłóceń na każdym etapie rekonstrukcji. Wadą tej metody jest fakt, iż jej jakość jest funkcją wybranego opóźnienia [8]. Równie złożonym zagadnieniem jest problem wyboru wymiaru zrekonstruowanej przestrzeni m .

Według teorii Takensa [4], poprawną (tzw. zanurzoną – *embedding*) rekonstrukcję, tzn. niezmienną wymiaru i wykładnika Lapunowa, gwarantuje warunek

$$m \geq 2d + 1 \quad (1)$$

gdzie d jest wymiarem topologicznym atraktora.

Najczęściej jednak trudno o informacje dotyczące wymiaru topologicznego, a zatem i o predykcję prawidłowej wartości zrekonstruowanego wymiaru.

2.1. Metody określania parametrów m i τ

Określenie interesujących nas parametrów (m i τ) może odbywać się na podstawie różnych zaproponowanych podejść. Z tym jednakże zastrzeżeniem, że muszą one prowadzić do określenia najmniejszej wartości m i τ pozwalającej poprawnie zrekonstruować trajektorię, a poprzez to odzwierciedlić dynamikę systemu [3, 22].

Przedstawione teraz zostaną cztery metody umożliwiające określenie estymowanych parametrów.

1) Podejście subiektywne.

Wymaga ono zazwyczaj długich serii danych. Jego istota polega na określeniu jakiegoś stałego elementu w zrekonstruowanej przestrzeni, który będzie kryterium poprawności zanurzenia (*embedding*) parametrów m i τ . Gdy stopień zanurzenia jest mały, wówczas faktor zostaje zmieniony, a próba określenia parametrów ponowiona. W przypadku zadowalającego zanurzenia, faktor zostaje zachowany.

2) Podejście predykcyjne.

Pojedyncza wartość rozkładu początkowego zbioru wektorów zrekonstruowanej przestrzeni fazowej ukazuje najmniejszą liczbę niezależnych kierunków przestrzeni fazowej, które mogą być używane jako wymiar zanurzenia (*embedding*).

3) Podejście analityczne.

Niewłaściwie dobrany wymiar zanurzenia przestrzeni fazowej implikuje wzrost liczby fałszywych najbliższych sąsiadów danego punktu przestrzeni. Na drodze badania można przeprowadzić analizę sąsiedztwa punktów przestrzeni fazowej i na tej podstawie określić wymiary m i τ .

4) Podejście zwyczajowe.

Wartość wymiaru m ustala się na drodze inkrementacji wymiaru, lub wykładnika Lapunowa, aż do osiągnięcia zbieżności dla dwóch lub trzech kolejnych wartości m [8, 9].

W rzeczywistym przypadku, rekonstruując przestrzeń fazową z n znormalizowanych parametrów k_i , opisujących określone *a priori* cechy naszego systemu, oraz posiadając określony wymiar przestrzeni zanurzenia (*embedding*) m i opóźnienie τ , musimy postępować zgodnie z następującą $(m \cdot n)$ -krokową regułą (2):

$$x(t) = \begin{pmatrix} k_1(t), & \dots, & k_n(t), \\ k_1(t + \tau), & \dots, & k_n(t + \tau), \\ k_1(t + 2\tau), & \dots, & k_n(t + 2\tau), \\ \vdots & \vdots & \vdots \\ k_1(t + (m-1)\tau), & \dots, & k_n(t + (m-1)\tau) \end{pmatrix} \quad (2)$$

W systemach chaotycznych, w których najczęściej rozważana jest problematyka atraktora, precyzja wyznaczenia parametrów m i τ jest szczególnie ważna (por. rys. 3). Od nich bowiem zależy bezpośrednio dokładność w określeniu pozostałych parametrów opisujących trajektorię procesu w przestrzeni fazowej. Proponowane są więc tutaj dwie drogi, które można by umownie nazwać: „teoretyczna” i „praktyczna”. Różnią się one zarówno sposobem dochodzenia do rozwiązania, jak i narzędziami obliczeniowymi, którymi się posługują.

Podejście teoretyczne

Przyjmuje się tutaj brak korelacji pomiędzy m i τ a wszystkimi innymi parametrami, co implikuje, iż zarówno m , jak i τ mogą być wyznaczone niezależnie. Już Takens udowodnił, że m i τ są niezależne, w przypadku wolnych od zakłóceń/szumów układów chaotycznych, o nieskończonej długości serii czasowych [4]. Wykorzystując powyższą właściwość, zaproponowany został algorytm Grassbergara–Procaccia, pozwalający określić zanurzoną wartość wymiaru m [12]. W celu wyznaczenia wartości τ , podejście to przewiduje trzy grupy możliwości:

- 1) metody korelacyjne (m.in.: autokorelacja, korelacja wyższego rzędu [13], metoda całościowej informacji [14, 19]);
- 2) metody wielokrotnej autokorelacji [15];
- 3) metody rozszerzenia przestrzeni fazowej (m.in.: metoda współczynnika wypełnienia [16, 17], wahającego się produktu [16, 19], dekompozycji pojedynczej wartości [12, 18]).

Ze względu na specyfikę tej publikacji, omówiony teraz zostanie jedynie wybrany przykład drogi teoretycznej. Opisy pozostałych można odnaleźć m.in. w podanych publikacjach.

Próby rekonstrukcji przestrzeni fazowej, a w szczególności określenie τ i m , przysparzają sporo problemów. Tradycyjne metody rekonstrukcyjne bazują na aparacie matematycznym [10], opierającym się na analizie funkcji korelacji własnej sygnału $R_{xx}(\tau)$, charakteryzującą związek pomiędzy wartościami sygnału w danej chwili a wartościami tego sygnału w chwilach kolejnych czasowych (rys. 1).

$$R_{xx}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) \cdot x(t + \tau) dt \quad (3)$$

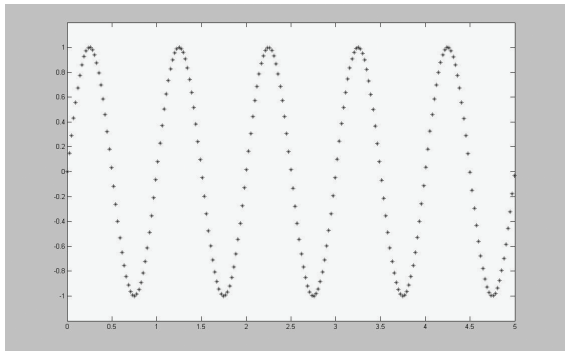
Dlatego też maksymalną wartość funkcja ta przyjmuje dla zerowej wartości opóźnienia.

Istnieje też podejście paralelne, w którym poddawany jest analizie sygnał obiektu $y(t)$, w odniesieniu do wzorcowego sygnału danej klasy $x(t)$.

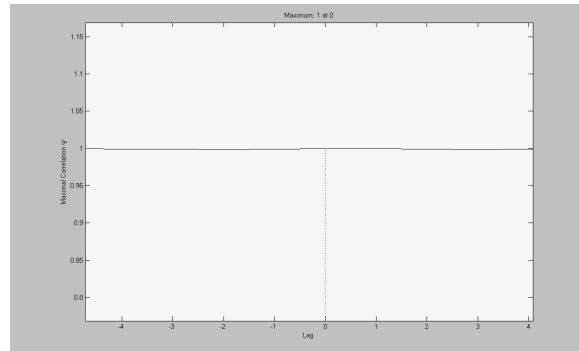
$$R_{xy}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) \cdot y(t + \tau) dt \quad (4)$$

Tak otrzymana funkcja interkorelacji $R_{xy}(\tau)$ określa miarę podobieństwa analizowanego sygnału (obiektu) w stosunku do sygnału (obiektu), którego parametry są dla nas znane (rys. 2).

a)

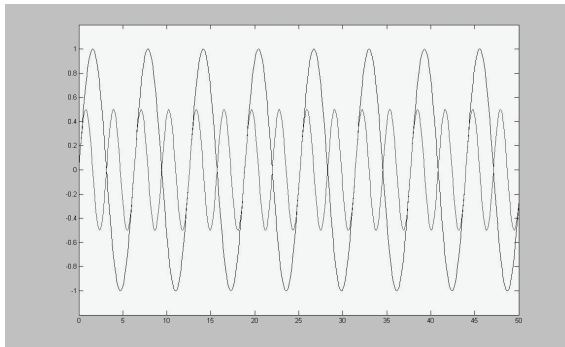


b)

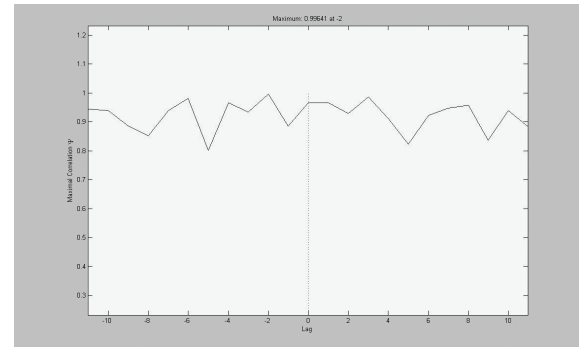


Rys. 1. Wykres sygnału sinusoidalnego $x_1 = 0:01:99$; $x = x_1 * 3 * \pi / 4$; $y = \sin(2 * \pi * x)$ (a) i jego autokorelacja (b)

a)



b)



Rys. 2. Zależność pomiędzy dwoma sygnałami sinusoidalnymi, takimi że: $x_1 = 1:1:99$; $y_1 = \sin(x_1)$; $x_2 = x_1$; $y_2 = \sin(x_2) * \cos(x_2)$ (a) i ich korelacja wzajemna (b)

Na potrzeby analizy i próby oszacowania parametrów rozważanego sygnału, możemy również posłużyć się współczynnikiem korelacji $\rho(x, y)$, opisanym wzorem (5).

$$\rho(x, y) = \frac{\text{Cov}(x, y)}{\sqrt{\sigma^2(x) \cdot \sigma^2(y)}} = \frac{E[(x - E(x)) \cdot (y - E(y))]}{\sqrt{\sigma^2(x) \cdot \sigma^2(y)}} =$$

$$= \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{\left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \left[n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}} \quad (5)$$

gdzie:

$\text{Cov}(x, y)$ – kowariancja sygnału x i y ,

$E[x]$ – estymata sygnału,

$\sigma^2(x)$ – wariancja sygnału x , będąca jego miarą rozrzutu,

$\sqrt{\sigma^2(x)}$ – odchylenie standardowe sygnału x .

Jeżeli współczynnik korelacji, przyjmujący wartości z przedziału $[-1; 1]$, wynosi 1 lub -1 , to zmienne x i y są całkowicie skorelowane (odpowiednio: dodatnio lub ujemnie). Dzieje się tak wówczas, gdy między nimi występuje zależność liniowa. Jeśli współczynnik korelacji jest równy zero, rozważane zmienne sygnały należy uznać za nieskorelowane. Zasadniczo τ może zostać określone z zastosowaniem kryterium obniżenia wartości $R_{xx}(\tau)$ poniżej określonej wartości progowej, bądź też przy detekcji fluktuacji funkcji autokorelacji [11].

Inną stosowaną metodą jest szybka transformata Fouriera (*Fast Fourier Transform* – FFT). Pomimo iż została ona spopularyzowana, oraz faktu, że jej atutem jest prędkość wykonywania obliczeń, zdarza się, że metoda FFT jest nieefektywna [12]. Objawia się to tym, iż kryterium określone za jej pomocą dla jednego systemu dynamicznego, w innym przypadku może okazać się zupełnie niewłaściwe. Może się tak zdarzyć w przypadku źle zdefiniowanej zależności pomiędzy autokorelacją pojedynczych serii czasowych a przestrzenną rekonstrukcją atraktora [8].

Podejście praktyczne

Ukazuje ono sytuację, z którą o wiele częściej mamy do czynienia na co dzień (stąd też przyjęta nazwa metody). W podejściu praktycznym zakładamy, że m i τ są ze sobą ściśle powiązane, jako że w rzeczywistych układach mamy do czynienia z niepozbowionymi szumami seriami czasowymi o skończonej długości. W celu rekonstrukcji przestrzeni fazowej kładziemy założenie (6).

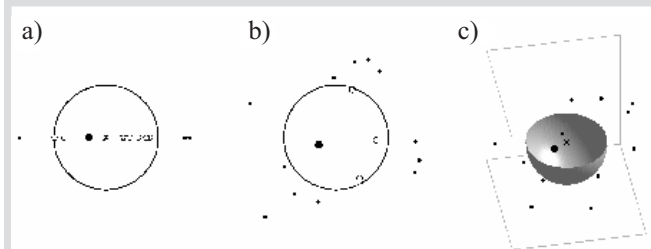
$$t_w = (m - 1) \cdot \tau \quad (6)$$

Istotą doboru zanurzonych parametrów dla danej serii czasowej jest zapewnienie utrzymania wartości t_w na zasadniczo niezmiennym poziomie. Należy jednak zaznaczyć, iż niewłaściwie dobrana relacja pomiędzy wartościami m i τ , znajduje bezpośrednie odzwierciedlenie w poprawności zrekonstruowanej przestrzeni fazowej. Z tego też powodu, przy wyznaczaniu wartości m i τ , coraz częściej znajdują zastosowanie

złożone rozwiązania metodyczne, takie jak np.: zautomatyzowany algorytm wyznaczania wartości progowych (*embedding algorithm*) [15] czy też metoda małych okien [20].

2.2. Sposób określania wartości granicznej ϵ

Osobnym problemem jest określenie wartości granicznej, określanej jako promień $r(i)$, bądź wartość progowa ϵ . Wskazane jest, aby wybrać możliwie najmniejszą wartość ϵ . Należy jednak mieć na uwadze fakt, że szum może zniekształcić strukturę RP (*Recurrence Plots*). Dlatego też zwykle się przyjmować wartość progową $\epsilon > \epsilon_{\min}$ na poziomie kilku procent z maksymalnej średnicy przestrzeni fazowej, jednakże mniej niż 10% ze średniej (ϵ_{sr}) lub maksymalnej (ϵ_{\max}) wartości ϵ .



Rys. 3. Wpływ wymiarowości na prawidłowość określenia lokalizacji punktów w przestrzeni
a) $m = 1$; b) $m = 2$; c) $m = 3$ [22]

Jak już napisano, istnieją także inne metody rekonstrukcji opierające się na tradycyjnej analizie matematycznej [8]. Celem tego artykułu nie jest jednak prezentacja ich wszystkich, a pokazanie nowego jakościowo-metodycznego podejścia, opartego nie tylko na analizie parametrycznej, lecz także zbudowanego za pomocą graficznej analizy sygnału.

3. GRAFICZNE METODY REKONSTRUKCJI PRZESTRZENI FAZOWEJ

Analiza trajektorii przestrzeni fazowej, jak już zostało napisane (v.s.), jest jedną z podstawowych metod analizy danych nieliniowych. Spośród wielu dróg wykorzystujących to podejście, proponowanych przez dzisiejszą naukę, na szczególną uwagę zasługuje metoda obrazów powrotu RP. Została ona zaproponowana w 1987 roku przez J.-P. Eckmanna [21]. Bazuje ona na elementarnej własności rozproszonych systemów dynamicznych, jaką jest właśnie powtarzalność. Pomimo iż nieraz nawet małe zaburzenia są w stanie wprowadzić w systemie wykładniczą rozbieżność, to po pewnym czasie układ wraca do swego pierwotnego stanu (choć nie jest on całkowicie zbieżny z pierwotnym). W praktyce metoda ta koncentruje się na krótkich i niestacjonarnych przebiegach czasowych [22]. Właśnie te cechy wydają się jej zasadniczą zaletą.

Przestrzeń fazowa najczęściej nie posiada 2–3 wymiarów, które pozwoliłyby nam dokonać wizualizacji jej układu, lecz znacznie więcej. Wizualizacja jednak przestrzeni wielowymiarowej nie jest prostym zadaniem. Narzędzie zaproponowane przez Eckmanna – dające nam właśnie taką

możliwość – wydaje się znakomitym rozwiązaniem. Należy jednak na samym początku zaznaczyć, iż realizacja ta dokonuje się w nowej dwuwymiarowej quasi-przestrzeni, do której zostały sprowadzone wszystkie pierwotne wymiary.

3.1. Recurrence Plot

Recurrence Plot jest tablicą punktów o wymiarach $N \times N$, gdzie każdy z punktów charakteryzują współrzędne (i, j) , takie że:

- $x(i)$ – i -ty punkt orbity opisujący system w m -wymiarowej przestrzeni ($i, j = 1, \dots, N$),
- $x(j)$ – punkt przestrzeni położony wystarczająco blisko $x(i)$.

W praktyce zmierza się do tego, aby RP był otrzymywany z serii czasowych $\{u_i\}$.

Proces tworzenia RP jest oparty na prostej trzyetapowej regule.

1. Wybieramy zanurzony wymiar m (*embedding dimension*), konstruując m -wymiarową orbitę $x(i)$, na podstawie metody opóźnień czasowych. Jeśli u_i są skalarami, to $x(i) = (u_i, u_{i+1}, u_{i+2}, \dots, u_{i+m-1})$.
2. Wybieramy promień $r(i)$, zaczepiony w punkcie $x(i)$ w \mathbf{R}^m , i wykorzystując go, tworzymy kulę tak, aby wewnątrz niej zawarta była rozsądna liczba punktów $x(j)$ orbity. Jeśli oczekiwana liczba punktów, mających znajdować się wewnątrz kuli, jest niezrealizowana, należy odpowiednio zwiększyć lub zmniejszyć długość promienia.
3. Tworzymy obraz RP taki, że każdy pojedynczy punkt tworzący go odpowiada punktowi (i, j) , gdzie $x(j)$ jest punktem wewnątrz kuli o promieniu $r(i)$ i centrum $x(i)$.

Obrazy RP lubią zachowywać symetrię wzdłuż diagonali $i = j$. Dzieje się tak, ponieważ elementy $x(i)$ oraz $x(j)$ są wzajemnie blisko siebie. Nie jest to jednak najczęściej symetria zupełna, gdyż nie wymagamy, aby $r(i) = r(j)$.

Ponieważ nie ma praktycznej możliwości zrealizowania całkowitej powtarzalności stanu tak, aby $\vec{x}(i) \equiv \vec{x}(j)$, powrót (*recurrence*) definiowany jest jako stan $\vec{x}(i)$ możliwie bliski $\vec{x}(j)$.

Matematyczny opis metody RP

Z matematycznego punktu widzenia powracanie realizowane jest wg wzoru (7)

$$R_{i,j}^m = \Theta(r(i) - \|x(i) - x(j)\|) \quad (7)$$

gdzie:

$$i, j = 1, 2, \dots, N,$$

$r(i)$ – odległość graniczna,

$\|\cdot\|$ – norma, predefiniowana metryka (np. Euklidesowa, maksymalna, minimalna i in.),

$\Theta(\cdot)$ – funkcja Heavyside'a zgodna z przepisem (8)

$$\Theta(x) = \begin{cases} 0 & \text{dla } x < 0 \\ 1 & \text{dla } x \geq 0 \end{cases} \quad (8)$$

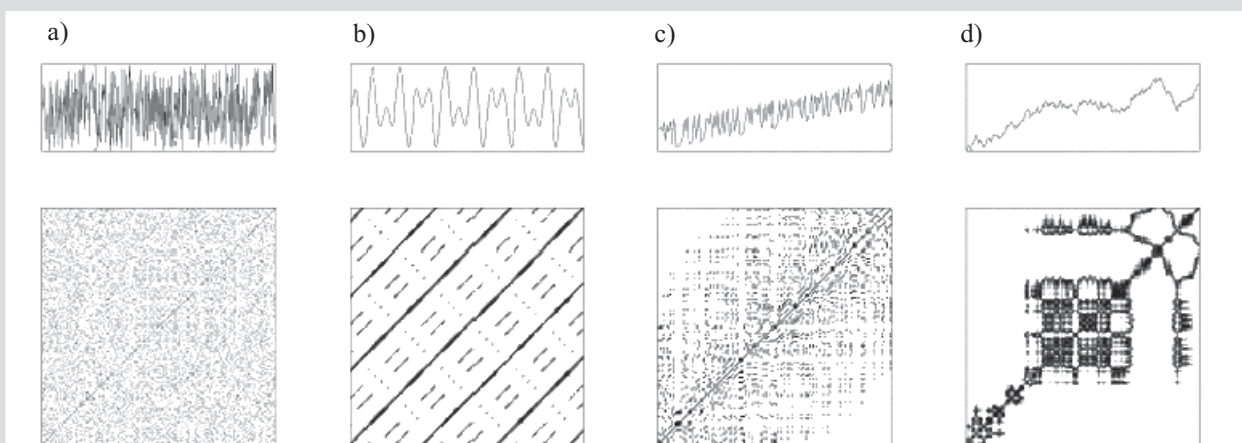
Wartości 0 i 1, realizowane za pomocą funkcji Heavyside'a, można prosto przedstawić graficznie kolorami: białym i czarnym.

Można zatem powiedzieć, że RP jest obrazem ukazującym te wszystkie czasy, dla których stany systemu dynamicznego powtarzają się. Metoda ta pokazuje bowiem wszystkie czasy, kiedy trajektoria przestrzeni fazowej systemu dynamicznego powraca ponownie w sąsiedztwo danego obszaru.

Struktury występujące w metodzie RP

Dla typowych systemów dynamicznych RP przyjmuje charakterystyczne obrazy/wzorce [21]. Posiadają one cechy charakterystyczne, których cztery grupy ukazano na rysunku 4.

Na rysunku 4a możemy zaobserwować jednorodny i nieregularny rozkład pojedynczych punktów, które nie tworzą żadnej koherentnej struktury. Taki plot charakterystyczny jest dla procesów stochastycznych i stacjonarnych, niezależnych systemów o krótkim czasie relaksacji w porównaniu z czasem tworzenia RP. Przykładem takiego systemu mogą być losowe serie czasowe. Rysunek 4b ukazuje układ okresowy. Cykliczne obserwowalne struktury, także w postaci linii diagonalnych oddzielonych od siebie o te same odległości świadczą o powtarzalnej dynamice układu, a także o występujących w układzie oscylacjach.



Rys. 4. Charakterystyczne wzory RP ukazujące typowe cechy różnych systemów dynamicznych [31]

Objaśnienia w tekście

Z tego typu obrazów jesteśmy w stanie wywnioskować o częstotliwościach występujących w układzie. Koncentrowanie się punktów wokół głównej diagonal, przy równoczesnym zanikaniu ich wraz z oddalaniem się od przekątnej (rys. 4c), świadczy o niestacjonarności procesu, objawiającej się dryftem występującym w amplitudzie sygnału. Nieregularne kształty: czarne klustry układające się na przemian z białymi obszarami świadczą o nieregularnym zachowaniu się systemu, które może być następstwem skorelowanego szumu. Kształty, które możemy obserwować na rysunku 4d, mogą też być następstwem nagłych zmian dynamiki układu, lub ukazywać rzadkie stany układu, dalekie od normalności [24, 25].

3.2. Cross Recurrence Plot

Specjalną odmianą metody RP jest metoda obrazu powrotów wzajemnych CRP (*Cross Recurrence Plots*). Podczas gdy w metodzie RP analizie poddawany był jeden sygnał, w metodzie CRP porównywane są dwie różne serie czasowe $x(i)$ i $y(j)$ ($i = 1, \dots, N; j = 1, \dots, M$), zanurzone w tej samej przestrzeni fazowej. W następstwie metody CRP otrzymujemy tablice $N \times M$.

Matematycznie metodę CRP opisuje formuła (9)

$$CR_{i,j}^m = \Theta(r(i) - \|x(i) - x(j)\|) \quad (9)$$

A zatem CRP jest obrazem ukazującym te wszystkie czasy, dla których stan jednego systemu dynamicznego równoległe występuje w drugim systemie dynamicznym. Metoda ta pokazuje wszystkie czasy, kiedy trajektoria przestrzeni fazowej pierwszego systemu dynamicznego znajduje się w sąsiedztwie trajektorii systemu drugiego.

Prezentowana metoda pozwala porównać dwie serie czasowe i znaleźć podobieństwa ich stanów. Jedną z cech charakterystycznych CRP jest brak głównej diagonal, za pomocą której można interpretować częstotliwości i stany fazowe. W klasycznej postaci występuje ona w metodzie RP (jako że mamy w niej do czynienia z plotem tego samego sygnału). W CRP występują dwa różne porównywane sygnały. Dlatego też najczęściej nie mamy do czynienia z główną diagonalną, lecz z pewną krzywą, zawierającą informacje o przeskalowaniu. Ponieważ może być ona użyta w celu synchronizacji obu serii czasowych, nazywamy ją linią synchronizacji LOS (*line of synchronization*).

3.3. Analiza ilościowa

Omawiając metodę graficznej analizy RP, warto wspomnieć także o synkretycznej metodzie nieliniowej analizy danych, łączącej elementy parametrycznej analizy ilościowej z elementami metody RP. Metodę analizy ilościowej sygnałów powracających RQA (*Recurrence Quantification Analysis*) zaproponowali w latach 90. XX wieku Webber i Zbilut [2, 25]. Koncentruje się ona na analizie ilości i czasu trwania powtórzeń systemu dynamicznego, na podstawie jego trajektorii w przestrzeni fazowej. W tym celu wykorzystuje ona szereg parametrów (m.in. takich jak (10)–(16)):

- współczynnik powracania RR (*recurrence rate*), ukazujący procentowy udział powracających punktów

$$RR = \frac{1}{N^2} \sum_{i,j=1}^N R_{i,j} \quad (10)$$

- determinizm (DET), ukazujący procentowy udział powracających punktów, tworzących diagonalę w strukturze w RP (gdzie $P(l)$ – histogram długości linii diagonalnych)

$$DET = \frac{\sum_{l=l_{\min}}^N l \cdot P(l)}{\sum_{i,j}^N R_{i,j}} \quad (11)$$

- laminarność (LAM), ukazujący procentowy udział powracających punktów, tworzących wertykale w strukturze w RP (gdzie $P(v)$ – histogram długości linii wertykalnych)

$$LAM = \frac{\sum_{v=v_{\min}}^N v \cdot P(v)}{\sum_{v=1}^N v \cdot P(v)} \quad (12)$$

- średnia/maksymalna długość diagonali (L/L_{\max})

$$L = \frac{\sum_{l=l_{\min}}^N l \cdot P(l)}{\sum_{l=l_{\min}}^N P(l)} \quad (13)$$

- dywergencja (DIV) powiązana z sumą pozytywnych wykładników Lapunowa

$$DIV = \frac{1}{L_{\max}} \quad (14)$$

- entropia, w sensie Shannona (ENTR), ukazująca prawdopodobieństwo rozkładu długości linii diagonalnych $p(l)$

$$ENTR = - \sum_{l=l_{\min}}^N p(l) \cdot \ln[p(l)] \quad (15)$$

- trend (TREND) ukazujący koncentrowanie się punktów wokół głównej diagonal, przy równoczesnym zanikaniu ich wraz z oddalaniem się od przekątnej; informuje nas o niestacjonarności procesu

$$TREND = \frac{\sum_{i=1}^N \left(i - \frac{N}{2}\right) (RR_i - \langle RR_i \rangle)}{\sum_{i=1}^N \left(i - \frac{N}{2}\right)^2} \quad (16)$$

– *tramping time* (TT) – średnia długość struktur wertikalnych; informuje nas jak długo system znajduje się w danym stanie

$$TT = \frac{\sum_{v=v_{\min}}^N v \cdot P^{\varepsilon}(v)}{\sum_{v=v_{\min}}^N P^{\varepsilon}(v)} \quad (17)$$

4. PRAKTYCZNE WYKORZYSTANIE METODY RP/CRP

Od momentu zaproponowania przed dwudziestu laty metody RP/CRP do analizy danych nieliniowych, stała się ona niezwykle popularna i jest stosowana w różnych dziedzinach wiedzy: począwszy od analizy matematycznej, poprzez wibroakustykę, analizę obrazów, analizę systemów finansowych, predykcję stanów klimatycznych, morfologię geologiczną, analizę reakcji chemicznych, aż do farmakologii oraz szeroko rozumianych nauk medycznych (genetyka, kardiologia m.in. EKG, neurologia m.in. EEG, psychologia i in.) [30]. Także dziedzina komunikacji i techniki transportu stara się spożytkować korzyści płynące z analizy danych pomiarowych metodą Recurrence Plot. Poniższe rozważania będą dotyczyły próby zastosowania metody RP/CRP do klasyfikacji pojazdów samochodowych.

4.1. Problematyka analizy ruchu drogowego

Od wielu lat obserwuje się lawinowy wzrost liczby pojazdów na drogach całej Europy. Od kilkunastu lat sytuacja ta zaczęła dotyczyć także Polski. Problem staje się niebanalny, jeśli kumuluje się on z nienajlepiej rozwiniętą siecią komunikacyjną oraz stanem polskich dróg. Tym bardziej staje się konieczne ustalenie jasnych i czytelnych kryteriów dotyczących zarówno prędkości, tonażu, jak i jakości pojazdów samochodowych (a dokładnie parametrów opisujących ich stan) poruszających się po polskich drogach [26]. Stąd pomysł, aby poddać analizie i klasyfikacji pojazdy samochodowe. Klasyczne pojęcie metrologiczne polega na ważeniu i badaniu składowych pojazdów na wadze statycznej. Jednak rozwiązanie to, choć efektywne, jest zbyt powolne oraz zbyt inwazyjne w system drogowy. Dlatego też powstała metoda ważenia pojazdów w ruchu WIM (*Weigh-in-Motion*) [27]. Tworzy ją system składający się z szeregu czujników zamieszczonych w samej jezdni lub bezpośrednio nad nią, odczytujących w sposób dynamiczny parametry ruchu pojazdów (długość pojazdu, jego prędkość, liczba osi, odległość między osiami, nacisk osi na podłoże, profil magnetyczny i in.). Są one automatycznie rejestrowane i, w zależności od potrzeb i możliwości danego systemu, analizowane bezpośrednio lub z pewną inercją, założoną przez konstruktorów. Tematem tej dysertacji nie jest omówienie warunków niezbędnych, jakie system musi spełnić, by otrzymane za jego pośrednictwem dane okazały się poprawne i użyteczne (m.in.: parametry systemu, sposób jego kalibracji i przetwarzania danych) [28], lecz wskazanie metody RP/CRP jako narzędzia, za pomocą którego można podjąć próbę klasyfikacji pojazdów samochodowych. W poniższej analizie wykorzystano dane otrzymane z rzeczywistego układu pomia-

rowego, usytuowanego na jednej z polskich dróg, oraz bibliotekę Matlab'a CRP Toolbox 5.3 (by Norbert Marwan).

4.2. Zastosowanie metody RP/CRP do klasyfikacji pojazdów samochodowych

W celu klasyfikacji pojazdów samochodowych dokonano preselekcji sygnałów. Ze względu na specyfikę ruchu samochodowego w Polsce pojazdy podzielono wstępnie na cztery grupy pojazdów:

- 1) osobowe,
- 2) dostawcze,
- 3) autobusy,
- 4) TIR-y.

Ten wstępny podział nie spełnia jednak wszystkich kryteriów. Zdarza się bowiem, że po drogach poruszają się pojazdy nietypowe (rolnicze, wojskowe i in.), pojazdy z przyczepami, rowerzyści itp. Wszystkie te pojazdy, jeśli tylko przejeżdżały przez system mierzący zamontowany w asfalcie, były źródłem generowanego sygnału. Na zasadzie analizy porównawczej usunięto te sygnały ze zbioru danych. Postawionym przed projektem celem było udowodnienie, że klasyfikacja pojazdów samochodowych za pomocą metody RP/CRP jest możliwa i że przynosi zadowalające rezultaty. Przyjęte zatem czteroklasowe kryterium służyło polaryzacji klas i typów pojazdów podlegających analizie. Sygnałem podlegającym analizie był sygnał pochodzący z pętli indukcyjnej, o szerokości 1,5 metra, która została umieszczona w asfalcie i po której przejeżdżał obiekt podlegający analizie.

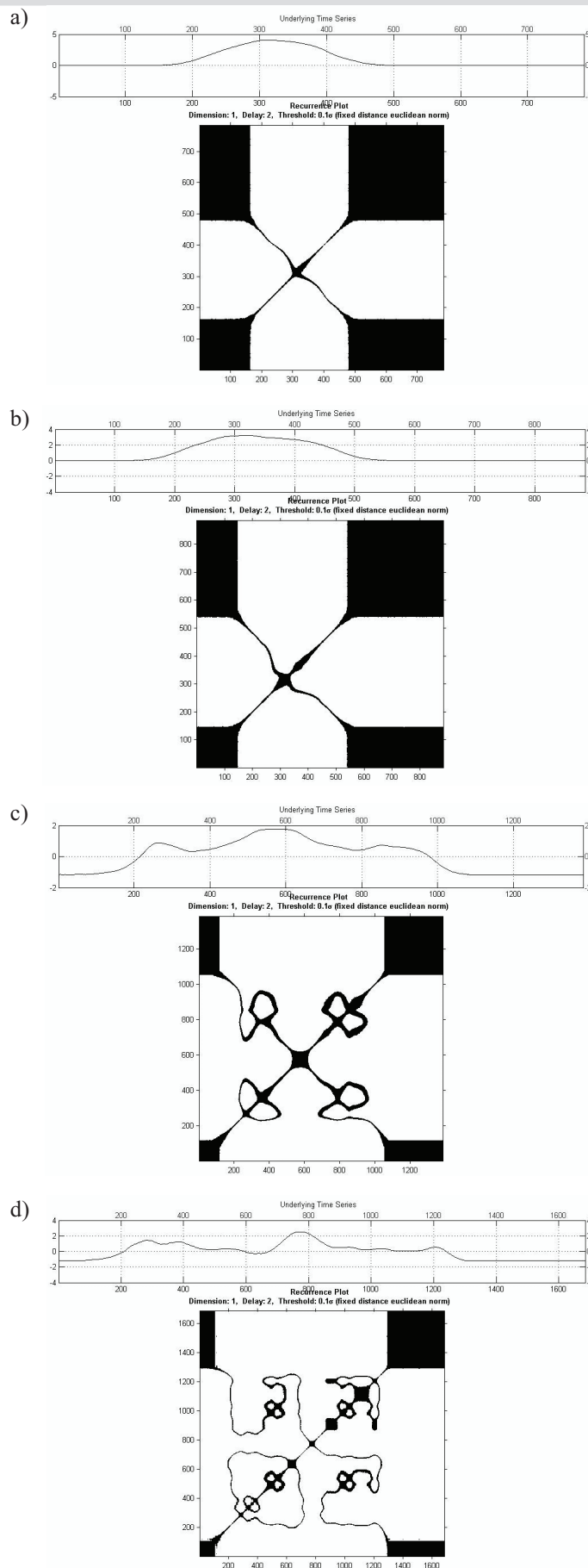
Poniżej przedstawiono efekty zastosowania metody RP dla wzorców ustalonych dla każdej z czterech klas analizowanych obiektów. Należy nadmienić, że wzorzec powstał na drodze uśrednienia wyników pojazdów, zaklasyfikowanych na drodze wizualnej preselekcji do jednej z czterech grup.

4.3. Wnioski wstępne z analizy CRP

Jak można zauważyć, rysunki 5a–d składają się z dwóch części. Na pierwszej (górnej) ukazany jest profil magnetyczny danego pojazdu (klasy), a na dolnym jego CRP. Należy przypomnieć to, co zostało już wcześniej napisane (v.v. podrozdz. 3.1), że wartości 0 i 1, realizowane za pomocą funkcji Heaviside'a, zostały przedstawione graficznie kolorem białym i czarnym.

Analiza profili magnetycznych rysunków skłania do refleksji nad podobieństwem klasy samochodów osobowych i dostawczych (O-D), oraz pewnym podobieństwem występującym pomiędzy profilem magnetycznym pochodzącym od autobusów i samochodów klasy TIR (A-T). Wydaje się to uzasadniona uwaga, jako że samochody osobowe mają często podobną długość oraz zbliżone do samochodów dostawczych rozmieszczenie osi. Pewne podobieństwo zauważalne jest także w przypadku autobusów i TIR-ów. Tu jednak mamy najczęściej do czynienia z inną długością pojazdu oraz zróżnicowaną liczbą osi. W naszej analizie profilu magnetycznego odwołujemy się zasadniczo do długości pojazdu oraz liczby jego osi, bowiem zasadniczo te dwie wartości jesteśmy w stanie odczytać z tego rodzaju przebiegów.

Przyjrzyjmy się teraz plotom CRP odpowiadającym prezentowanym profilom. Wszystkie wykonane zostały przy identycznie zdefiniowanych parametrach rekonstrukcji przestrzeni fazowej, na poziomie: $m = 1$, $\tau = 2$, $\varepsilon = 0,1\sigma$.



Rys. 5. Profil magnetyczny i plot CRP dla klasy: a) samochodów osobowych; b) samochodów dostawczych; c) autobusów; d) TIR-ów

W pierwszej grupie (O-D) dane niezerowe, odpowiadające obiektowi znajdującemu się w zakresie aktywności czujnika, pojawiają się wcześniej i później opadają w przypadku samochodów dostawczych. Odpowiada to ich większej długości. Jednolitość profilu w obu przypadkach wynika z tego, iż najczęściej płyta podłogowa ma dość jednorodną budowę. Co zaś się tyczy CRP – to zauważalna jest pewna zależność, a mianowicie w zakresie danych pomiarowych różnych od zera, wartość CRP jest równa zero, a w pozostałych równa jeden. Oba prezentowane rysunki różnią się jednak kształtem diagonalni. W przypadku samochodów osobowych diagonalne są prostsze, a ich punkt przecięcia się (ukazujący wartość szczytową przebiegu/profilu) bardziej jednorodny. W przypadku samochodów dostawczych linie diagonalne wykazują większą skłonność do zniekształceń, zaś sama wartość szczytowa, odzwierciedlona jest większym zniekształceniem środkowym, które odpowiada szerszemu szczytowi profilu (większej długości samochodu).

W przypadku (A-T) zauważalny jest dłuższy zbiór danych i odpowiadający mu dłuższy przebieg. Załamania w profilu odpowiadają występującym wyraźnie w podwoziu osiom samochodu. Ponieważ TIR-y są z reguły dłuższymi pojazdami od autobusów, oraz mają więcej osi (T: 4÷5, A: 2÷3), ich ploty CRP też różnią się od siebie. Niejednorodność przebiegu, większa w przypadku T aniżeli A, odzwierciedla się

w bardziej złożonej formie CRP. Podobnie jak w pierwszym przypadku zauważalna jest tu główna diagonalna.

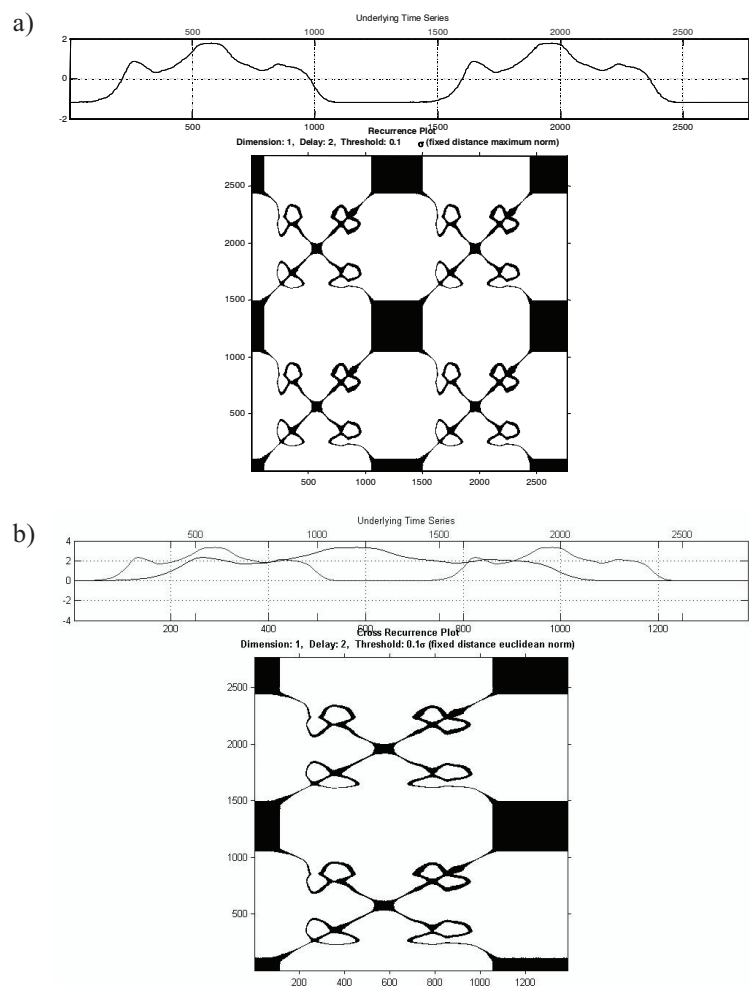
Wszystkie przebiegi są symetryczne w stosunku do prostej $i = i$ (gdzie i – kolejny pomiar), co jest efektem oczekiwanym, jako że analizie poddawany jest za każdym razem jeden i ten sam obraz.

Analizując omawiane sygnały, można zauważyć, iż zasadniczo zarejestrowane sygnały nie są długie. Liczba próbek dla każdego z nich, ściśle uzależniona od długości analizowanej klasy pojazdu (przy założeniu stałej częstotliwości próbkowania $f_{pr} = 1$ kHz), waha się w granicach O: 784, D: 884, A: 1384, T: 1684. Podjęto więc próbę „przedłużenia sygnału”, poprzez jego skopiowanie. Na rysunku 6 przedstawiono uzyskane efekty dla klasy autobusów.

Jak widać, na rysunku 6, zarówno podwojenie sygnału, jak i CRP dla sygnału pojedynczego i podwojonego nie wnoszą wiele do analizy. Analogiczne wyniki uzyskano dla trzech pozostałych klas.

4.4. Parametryzacja obrazów CRP

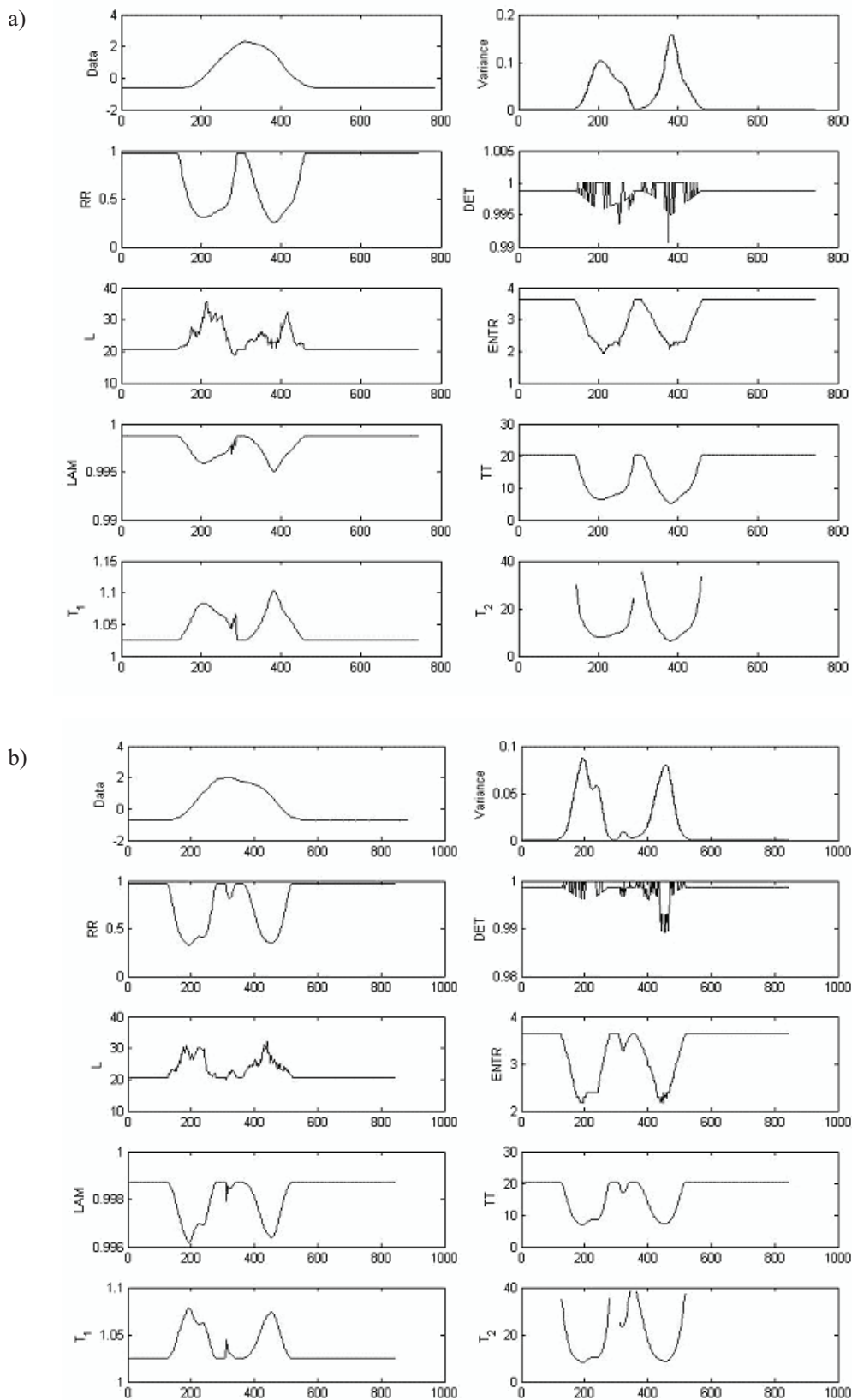
Na koniec, celem ukazania parametrów obrazów CRP, określających charakterystyczne cechy rozpatrywanych obiektów (v.s. podrozdz. 3.3), poddano sygnały analizie ilościowej (RQA). Wyniki obrazują rysunki 7a–d.



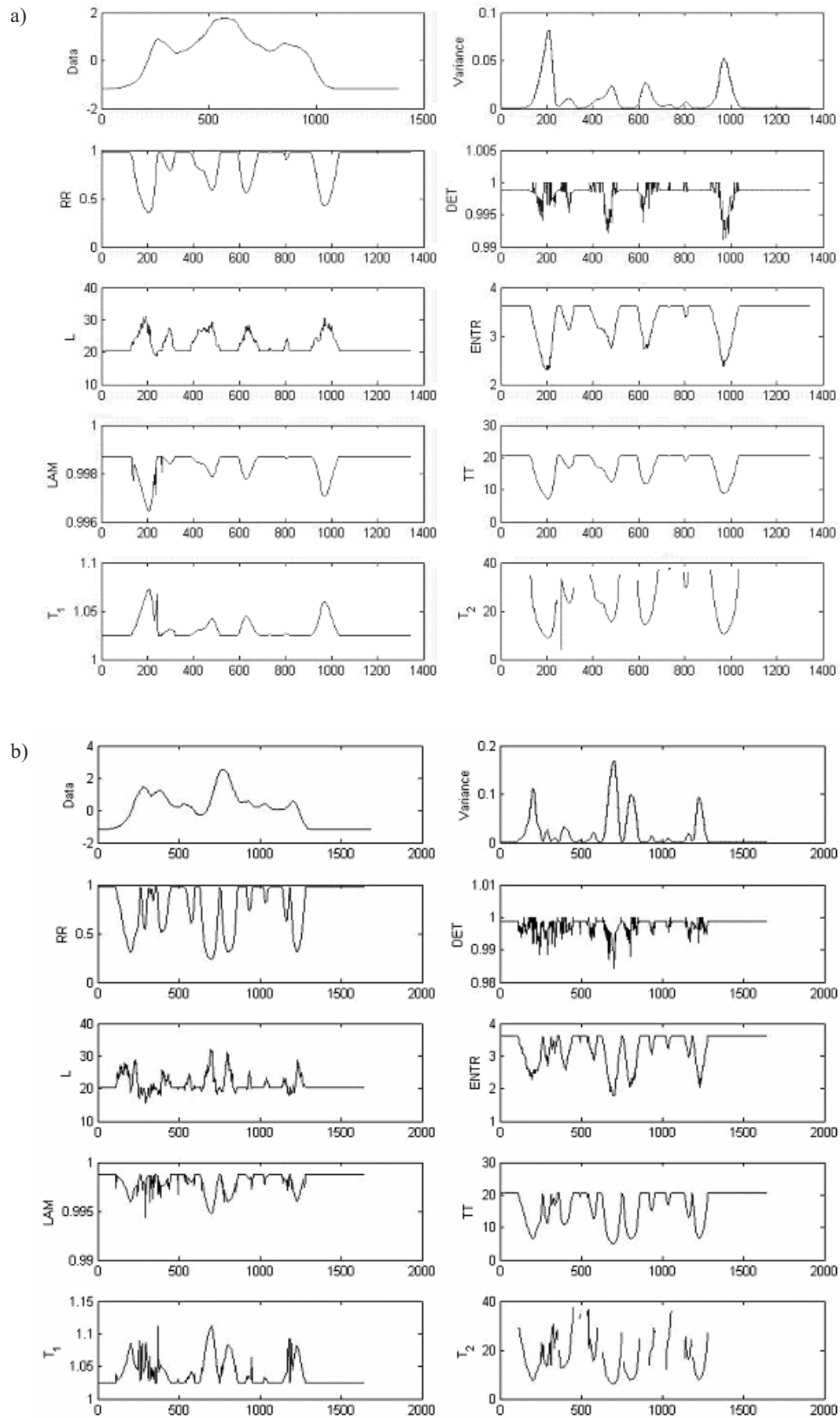
Rys. 6. CRP dla: a) podwojonego sygnału dla klasy autobusów; b) dla pojedynczego i podwojonego sygnału dla klasy autobusów

Jak widać na rysunkach 7 i 8, pomocnicze narzędzie, jakim w naszym przypadku jest RQA, pozwala znaleźć w obrazach CRP dodatkowe elementy wspólne, ale i różnice na „pierwszy rzut oka” niezauważalne. Zgodnie z tym, co zostało już wcześniej podkre-

ślone, obrazy dla klasy autobusów i TIR-ów są zdecydowanie bardziej zróżnicowane, co jest zauważalne także w przypadku analizy RQA. Takie parametry jak wariancja, entropia czy długość głównej diagonalnej, zdają się o tym wyraźnie świadczyć.



Rys. 7. Analiza ilościowa dla klasy samochodów: a) osobowych; b) dostawczych



5. UWAGI KOŃCOWE I PODSUMOWANIE

Dokonawszy określenia klasy pojazdów za pomocą metody *Cross Recurrence Plot*, wysiłek związany z dalszą realizacją pracy doktorskiej autor ukierunkował na następujące zagadnienia:

- Klasyfikacja metodą RP/CRP poszczególnych losowych sygnałów pochodzących od pojazdów samochodowych, na podstawie predefiniowanych klas.
- Szczegółowa analiza wpływu parametrów zanurzenia (m i τ), sąsiedztwa ε oraz przyjętej normy na wynik klasyfikacji.
- Bardziej szczegółowa analiza poszczególnych obrazów CRP.

Jak widać, graficzne metody analizy danych pomiarowych przynoszą ciekawe, a nieraz i zaskakujące efekty. Bardziej niż w metodach parametrycznych odzwierciedlone są tu podobieństwa, ale także i różnice pomiędzy poszczególnymi sygnałami. Zastosowanie do szeregu dyscyplin naukowych metod, ukazujących obrazy powrotów, wskazuje na jej uniwersalność, oraz zdaje się otwierać wciąż nowe perspektywy do jej dalszego rozwoju. Cała trudność tych metod ogniskuje się na wiernym odtworzeniu przestrzeni fazowej, z wykorzystaniem takich parametrów jak wymiarowość przestrzeni i opóźnienie czasowe, oraz na „nauczeniu się” odczytywania obrazów, będących jej efektem końcowym. Jak zostało zasygnalizowane, z końcowych plotów można uzyskać także szereg danych parametrycznych, pomocnych do dalszej analizy zachowania układu. Wydaje się, że dynamiczny rozwój metod RP/CRP przyczyni się do lepszego poznania właściwości wielu układów dynamicznych, dla których analizy są one dedykowane.

Literatura

- [1] Turing A.: *On Computable Numbers with an Application to the Entscheidungsproblem*. 1937
- [2] Webber C.L. jr., Zbilut J.P.: *Embeddings and delays as derived from quantification of recurrence plots*. *Physics Letters A*, 171 (3–4), 1992, 199–203
- [3] Kantz H., Schreiber T.: *Nonlinear time series analysis*. University Press, Cambridge 1997
- [4] Takens F.: *Detecting strange attractors in turbulence*. *Lecture Notes in Mathematics*, 898, 1981
- [5] Packard N.H., Crutchfield J.P., Farmer J.D., Shaw R.S.: *Geometry from a time series*. *Physical Review Letters* 45, 1980
- [6] Farmer J.D., Ott E., Yorke J.A.: *The dimension of chaotic attractors*. *Physica D*, 1983
- [7] Ott E., Sauer T., Yorke J.A.: *Coping with chaos*. Wiley Interscience, 1994
- [8] Rosenstein M.T., Collins J.J., De Luca C.J.: *Reconstruction expansion as a geometry-based framework for choosing proper delay times*. *Physica D* 73, 1994
- [9] Aleksic Z.: *Estimating the embedding dimension*. *Physica D* 52, 1992
- [10] Gajda J.: *Statystyczna analiza danych pomiarowych*. Skrypt AGH, Kraków 2002
- [11] King G.P., Jonem R., Broomhead D.S.: *Phase portraits from a time series: a singular system approach*. *Nuclear Physics B* 2, 1987
- [12] Albano M., Muench J., Schwartz C.: *Singular value decomposition and the Grassberger-Procaccia*. *Physical Review AD* 58, 1992
- [13] Albano A.M., Passamante A., Farrell M.E.: *Using higher-order correlations to define an embedding window*. *Physica D*, 54, 1991, no. 1–2
- [14] Fraser A.M.: *Information and entropy in strange attractors*. *IEEE Transactions on Information Theory*, 35, no. 2, 1989
- [15] Lin J., Wang Y., Huang Z., Shen Z.: *Selection of proper time-delay in phase space reconstruction of speech signals*. *Signal Processing* 15(2), 1999
- [16] Pfister G., Buzug T.: *Comparison of algorithms calculating optimal embedding parameters for delay time coordinates*. *Physica D*, 58(1–4), 1992
- [17] Pfister G., Buzug T.: *Optimal delay time and embedding dimension for delay-time coordinates by analysis of the global static and local dynamical behavior of strange attractors*. *Physical Review A*, 45, 1992
- [18] Kembe G., Fowler A.C.: *A correlation function for choosing time delays in phase portrait reconstructions*. *Physics Letters, A* 179(2), 1993
- [19] Liebert W., Pawelzik K., Schuster H.G.: *Optimal embeddings of chaotic attractors from topological considerations*. *Europhysics Letters*, 14, 521, 1991
- [20] Kugiumtzis D., Christophersen N.: *State space reconstruction: method of delays vs singular spectrum approach*. Research report 236, University of Oslo, 1997
- [21] Eckmann J.-P., Oliffson Kamphorst S., Ruelle D.: *Recurrence Plots of dynamical systems*. *Europhysics Letters*, 4 (9), 1987
- [22] Marwan N.: *Encounters with Neighbours*. University of Potsdam, 2003
- [23] Cao L.: *Practical method for determining the minimum embedding dimension of a calar times series*. *Physica D*, 110 (1–2), 1997
- [24] Eckman J.-P., Ruelle D.: *Ergodic theory of chaos and strange attractor*. *Review of Modern Physics*, 75 (3), 1985
- [25] Webber C.L. jr., Zbilut J.P.: *Dynamical assessment of physiological systems and states using recurrence plot strategies*. *Journal of Applied Physiology*, 76, 1994
- [26] Grabowski T.: *Analiza kalibracji systemów ważenia pojazdów w ruchu*. W „V Ogólnopolskie Warsztaty Doktoranckie OWD 2003”, t. 2, cz. 17, 2003
- [27] *Weigh-in-Motion of Road Vehicle*. Final Report of COST 323 action (ver. 3.0), 1999
- [28] Gajda J., Sroka R., Stencel M., Żegleń T.: *Systemy pomiarowe parametrów ruchu drogowego*. W: „Komputerowe systemy wspomaganie nauki, przemysłu i transportu »Transcomp«”, Zakopane 2001
- [29] Gajda J., Sroka R., Stencel M., Żegleń T.: *Measurement of road traffic parameters usin an inductive single-loop detector*. W: „Electrical Instruments in Industry”, Glasgow 1997
- [30] www.recurrence-plot.tk
- [31] www.wikipedia.org

Wpłynęło: 21.11.2006

Tomasz GRABOWSKI



Urodzony w Krakowie 21.11.1978 roku. Studia magisterskie na Wydziale Elektrotechniki, Automatyki, Informatyki i Elektroniki Akademii Górniczo-Hutniczej w Krakowie ukończył w 2002 roku. Pracę magisterską obronił z zakresu rozpoznawania i klasyfikacji obiektów w Katedrze Metrologii. Obecnie jest doktorantem na Wydziale Elektrotechniki, Automatyki, Informatyki i Elektroniki AGH i zajmuje się analizą sygnałów nieliniowych.

e-mail: tomasz.grabowski@poczta.fm