

# Nov način razpoznavanja masnih spektrov

## New Approach to Mass Spectra Recognition

Belič I.<sup>1</sup>, VŠNZ, Ljubljana  
L. Gyergyek, FER, Ljubljana

*Masna spektroskopija je le ena mnogih analitskih metod, kjer se veliko dela vlaga v avtomatsko razpoznavanje signalov. Razvito je bilo veliko število numeričnih algoritmov za razpoznavanje spektrov, pri čemer je njihova uporabnost precej različna. Zelo veliko algoritmov odpove, kadar je osnovnemu signalu iz masnega spektrometra primešan šum. Prispevek opisuje uporabo na novo razvitega nevrnskega sistema za razpoznavanje masnih spektrov. Rezultati delovanja so primerjani z nekaterimi klasičnimi metodami razpoznavanja. Nevronski sistem je predstavljen v obliki filtra in v obliki sistema za kvalitativno analizo signala. Predstavljeni nevrnski sistem je možno uporabiti splošno tudi za druga področja spektrometrije.*

*Ključne besede: masna spektrometrija, nevrnski sistemi, razpoznavanje spektrov, simulacija spektrov*

*The mass spectra detection and automatic recognition is one of many analytical methods widely used. Several algorithms were developed for automatic spectra recognition that are more or less effective and accurate. The use of algorithms is often limited by the noise superimposed to the original signal. The paper describes the use of specially designed neural networks for mass spectra recognition. The results of our work are compared with data obtained by procedural Singular Value Decomposition methods. Neural networks are introduced and tested as the signal filtering tool as well as the mass spectra qualitative analysis system. The described system can be used generally and not only for the mass spectra recognition.*

*Key words: mass spectrometry, neural networks, spectra recognition, spectra simulation*

### 1. Uvod

Avtomatsko razpoznavanje komponent masnega spektra je zahteven postopek. Pri klasičnih postopkovnih metodah razpoznavanja predstavlja osnovni problem šum, ki popači izhodni signal masnega spektrometra in povzroči, da so rezultati analize neuporabni.

Namen dela je poiskati principe, kako iz originalnega signala izločiti šum in omogočiti kolikor mogoče natančno analizo. V delu<sup>4,5</sup> je zelo natančno opisan postopek razpoznavanja in detekcije komponent masnega spektra v kvalitativni in kvantitativni obliki. Prav tako so nakazane težave, ki pri tem nastopajo.

Pri signalih, kjer lahko prisotnost motilnih šumov omejimo na zanemarljivo vrednost, lahko z veliko natančnostjo izvedemo razpoznavanje spektrov po klasičnih postopkih (Singular Value Decomposition)<sup>4</sup>.

Zelo pomembno je poznavanje narave motilnih šumov, ki nastopajo pri masni spektroskopiji. Prva vrsta so šumi elektromagnetnega izvora, ki so v vakuumskih sistemih, v katere je masni spekter vključen, zelo pogosti. Obstaja cela vrsta elektronskih filtrov, ki zmanjšajo velikost omenjenih šumov na naj-

manjšo možno mero. Za postopek razpoznavanja komponent masnih spektrov lahko tovrstne šume zanemarimo. Druga vrsta motilnih signalov pa so spektri, ki jih sistem pri razpoznavanju ne upošteva.

Algoritmi za razpoznavanje komponent masnega spektra razpoznavajo le tiste komponente atmosfere, ki so vnaprej pričakovani. Če se pojavi snov, ki ni bila vključena v osnovni nabor, se razpoznavanje poruši. Obstaja več načinov, kako se izogniti omenjenemu problemu. Najlaže je povečati število komponent, ki se v sistemu lahko pojavijo. Tak način zahteva veliko zmogljivost računalnika, ki izvaja razpoznavanje. Razen tega tudi nikoli nismo povsem sigurni, da se vendarle v sistemu ne bi pojavila snov, ki je ni v osnovnem naboru.

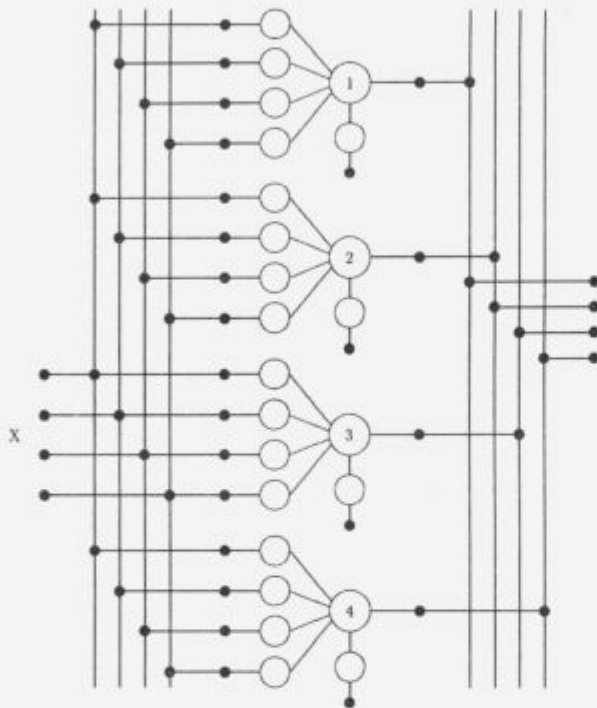
Alternativa je filtriranje vhodnega signala s filtrom, ki omogoča odstranjevanje motilnih komponent. V našem primeru je nevrnski sistem uporabljen v vlogi filtra za odstranjevanje motilnih vplivov. Nevronski sistem izvaja kvalitativno analizo vhodnega signala.

### 2. Uporabljen nevrnski sistem

Za izvajanje kvalitativne analize masnega spektra je bil uporabljen večnivojski nevrnski sistem brez povratnih povezav. **Slika 1** prikazuje en nivo nevrnskega sistema.

<sup>1</sup> mag. Igor BELIČ  
VŠNZ  
Kotnikova 8, 61000 Ljubljana

Nevronski sistem deluje kot posebna izpeljanka asociativnega pomnilnika<sup>1,2</sup>, ki smo ga poimenovali pomnilnik z vektorskim intervalnim naslavljanjem (pomnilnik bo v drugih prispevkih podrobno opisan). Naloga pomnilnika z vektorskim intervalnim naslavljanjem je izvajanje preslikave vhodnega prostora v izhodni pojmovni prostor. Pojmovni prostor vsebuje množico pojmov, kot npr. H<sub>2</sub>O, N<sub>2</sub> itd. Ob nekem poljubnem vhodnem signalu, se v pojmovnem prostoru izberejo tisti pojmi, ki ustrezajo posameznim komponentam atmosfere v vakuumski komori. Proces učenja je postopek<sup>1,2,3</sup>, ki zgradi ustrezno funkcijo preslikave med vhodnim in pojmovnim prostorom. V našem primeru je bil uporabljen nadzorovani postopek učenja.



Slika 1: En nivo uporabljenega nevronskega sistema  
Figure 1: One level of the used neural network system

### 3. Uporaba nevronskega sistema za detekcijo komponent masnih spektrov

Sestavni elementi atmosfere v vakuumskem sistemu imajo poznane oblike spektrov. Označimo spekter *i* tega plina z X<sub>*i*</sub>. X<sub>*i*</sub> je *n* dimenzionalen vektor. Količina posameznega plina je v spektru izražena z multiplikacijsko konstanto, s katero se spekter plina pomnoži. V nekem časovnem trenutku je v vakuumskem sistemu atmosfera sestavljena iz vsote spektrov plinov. Enačba sestave masnega spektra je:

$$S = \sum_{i=1}^n k_i X_i + N \quad (1)$$

kjer so:

- S: spekter atmosfere v vakuumskem sistemu
- n: število plinov, ki sestavljajo atmosfero
- k<sub>*i*</sub>: količina *i*-tega plina v atmosferi
- X<sub>*i*</sub>: spekter *i*-tega plina
- N: naključni šum, ki je dodan spektru

Več je o problemu masne spektroskopije napisano v literaturi<sup>2</sup>.

### 4. Testiranje nevronskega sistema

Testno nevronske plast sestavljajo naslednji elementi:

- Štirje nivoji nevronov s po 50 nevronov v nivoju.
- Nevroni z 51 vhodi in sigmoidalno aktivacijsko funkcijo<sup>2</sup>, ki omejuje izhodne vrednosti celic v intervalu (-1, 1).
- Naključni generator povezovalnih uteži med nevroni. Vrednost naključnega izbora uteži je omejena med vrednosti -0.8 in 0.8.

Prvi test ugotavlja sposobnost nevronske plasti, da pravilno razporedi naučene spektre. Spektri so bili sistemu prikazani v relativni obliki.

Delovanje sistema je bilo preizkušeno s sledenjem naslednjih korakov:

- naključni izbor uteži;
- učenje sistema;
- preverjanje odziva plasti.

Test zajema 100 ponovitev omenjenih korakov. V tabeli 1 so zbrani rezultati testov. Tabela prikazuje odstotke pravih razvrstitev vzorcev. V prvem stolpcu so zbrane vrednosti pri naslavljanju intervalnega pomnilnika z eno celico, v drugem pa z dvema.

Tabela 1: Rezultat testa razpoznavanja komponent masnega spektra. Masni spektri so bili sistemu predstavljeni v normirani obliki  
Table 1: Results of the mass spectra recognition process. Mass spectra were introduced to the system in normalized form

plini	pravilnost razvrstitve v %	pravilnost razvrstitve v %
	naslavljanje: 1 celica	naslavljanje: 2 celici
H <sub>2</sub>	100	100
He	99	100
CH <sub>4</sub>	97	100
NH <sub>3</sub>	100	100
H <sub>2</sub> O	100	100
Ne	99	100
C <sub>2</sub> H <sub>2</sub>	100	100
C <sub>2</sub> H <sub>4</sub>	98	100
N <sub>2</sub>	100	100
CO	100	100
C <sub>2</sub> H <sub>6</sub>	100	100
NO	96	100
C <sub>3</sub> H <sub>6</sub>	100	100
C <sub>3</sub> H <sub>8</sub>	100	100
CO <sub>2</sub>	100	100
CH <sub>2</sub> O	100	100
O <sub>2</sub>	97	100
H <sub>2</sub> S	100	100
A	96	100
N <sub>2</sub> O	100	100
C <sub>2</sub> H <sub>2</sub> O	99	100
C <sub>2</sub> H <sub>4</sub> O	98	100
NO <sub>2</sub>	100	100
CH <sub>2</sub> O <sub>2</sub>	100	100
C <sub>4</sub> H <sub>10</sub>	100	100
C <sub>3</sub> H <sub>8</sub> O	99	100
C <sub>2</sub> HC <sub>13</sub>	100	100

#### Ocena rezultatov testa

Ker je bila plast nevronov naučena razpoznavati točno določene vzorce, ki so obenem tudi vsi možni vzorci, ki jih

sistem lahko v svojem delovanju dobi na vhodu, so intervali oz. razredi zelo ozko definirani. Napake, ki se pri razvrstitvi pojavijo v primeru naslavljanja intervalnega pomnilnika z eno celico so posledica naključne nelinearnosti preslikave, ki jo naredi nevronski sistem. Na pojav napake vpliva tudi širina izbranega intervala okoli vzorcev.

**Tabela 2:** Rezultat testa razpoznavanja komponent masnega spektra.

Masni spektri so bili sistemu predstavljeni v nenormirani obliki  
**Table 2:** Results of the mass spectra recognition process. Mass spectra were introduced to the system in unnormalized form

plini	pravilnost	pravilnost	pravilnost
	razvrstitve v % naslavljanje: 1 celica	razvrstitve v % naslavljanje: 2 celici	razvrstitve v % naslavljanje: 3 celice
H <sub>2</sub>	62	94	100
He	64	88	100
CH <sub>4</sub>	66	84	100
NH <sub>3</sub>	64	92	100
H <sub>2</sub> O	60	96	99
Ne	68	97	100
C <sub>2</sub> H <sub>2</sub>	63	100	100
C <sub>2</sub> H <sub>4</sub>	100	82	100
N <sub>2</sub>	94	97	100
CO	85	99	100
C <sub>2</sub> H <sub>6</sub>	70	84	100
NO	69	90	100
C <sub>3</sub> H <sub>8</sub>	90	100	100
C <sub>3</sub> H <sub>6</sub>	95	100	100
CO <sub>2</sub>	78	94	100
CH <sub>3</sub> O	69	87	99
O <sub>2</sub>	67	93	100
H <sub>2</sub> S	92	100	100
A	65	84	97
N <sub>2</sub> O	84	99	100
C <sub>2</sub> H <sub>2</sub> O	87	95	100
C <sub>2</sub> H <sub>4</sub> O	83	93	100
NO <sub>2</sub>	65	92	100
CH <sub>3</sub> O <sub>2</sub>	77	90	100
C <sub>4</sub> H <sub>10</sub>	93	93	100
C <sub>3</sub> H <sub>8</sub> O	71	100	98
C <sub>2</sub> HC <sub>13</sub>	65	99	100

#### Ocena rezultatov testa

Tudi v tem primeru se je sistem zelo dobro naučil razvrščanja vzorcev. Širina najmanjšega intervala okoli točke pošameznega vzorca vpliva na število intervalov, ki jih sistem ustvari v procesu učenja.

V prisotnosti šuma, ki ga primešamo osnovnim spektrom se rezultat učenja precej poslabša. Zelo pomembna je naslednja praktična izkušnja.

Če želimo, da bo sistem uspešno izvajal razpoznavanje signalov v prisotnosti dodanega šuma, mora tudi učenje potekati v prisotnosti šuma.

Prisotnost šuma v procesu učenja sistema razširi intervale naslavljanja v posamezne razrede. Učenje sistema ob prisotnosti naključnega šuma lahko postane dolgotrajen postopek, nikakor pa postopek ni zapleten.

Na rezultat učenja razvrščanja močno vpliva omejitev področja iz katerega naključni generator izbira vrednosti. Če so izbrane uteži prevelike, potem je prevajalna funkcija sistema v posameznih področjih zelo strma, na drugih pa se vrednost nahaja v nasičenju. Rezultat je stapljanje razredov (prekrivanje), ki onemogoči učenje.

V drugem testu je bila plast nevronov enaka kot v prvem, le cilj je bil zastavljen nekoliko zahtevnejše. Nevronski sistem se je učil razvrščati vhodne spektralne signale, pomnožene s poljubnimi konstantami (konstante  $k_i$  v enačbi (1)).

Vhodni vzorci so bili zato pripravljani v absolutni in ne več v relativni obliki. Učenje se je pričelo z malimi vrednostmi konstante  $k_i$ . Tabela 2 prikazuje, kako se je sistem naučil razvrščanja v stotih primerih naključno izbranih uteži in pri različnem številu uporabljenih celic za intervalno naslavljanje.

#### 5. Sklep

Razpoznavanje komponent, ki sestavljajo masni spekter je bilo narejeno s simulacijo masnega spektrometra in do sedaj v delo nismo vključili dejanske naprave. Zato je vrednost rezultatov manjša. Nevronski sistem, ki izvaja kvantitativno analizo masnega spektra je uporabljen kot filter za izločanje motilnih signalov. Filtriranju signala sledi kvalitativna analiza, izvedena s klasičnimi postopkovnimi metodami.

#### Literatura

- <sup>1</sup> T. Kohonen, Self Organisation and Associative Memory, Berlin, Springer Verlag, 1985
- <sup>2</sup> G. Matsumoto, Neurocomputing - Neurons as Microcomputers, Future Generations Computer Systems, Elsevier Sci.Publ., 4, 1988, 39-51
- <sup>3</sup> B. Kosko, Neural Networks and Fuzzy Systems, Englewood Cliffs, Prentice Hall International Inc., 1992
- <sup>4</sup> I. Belič, Avtomatsko razpoznavanje komponent masnega spektra, Univerza v Ljubljani, FER, 1985
- <sup>5</sup> M. J. Drinkwine, D. Lichtman, Partial pressure analyzers and analysis, American Vacuum Society Monograph Series, 1980