

# Analiza szeregów czasowych obrazów fMRI

Andrzej Chojnowski, Włodzisław Duch

Katedra Metod Komputerowych, Uniwersytet Mikołaja Kopernika,  
Grudziądzka 5, 87-100 Toruń; e-mail: [microbe,duch@phys.uni.torun.pl](mailto:microbe,duch@phys.uni.torun.pl)

20 czerwca 2000 roku

## Streszczenie

Tworzenie „map” aktywności obszarów mózgu zaangażowanego w wykonywanie różnych zadań wymaga analizy serii obrazów. Ze względu na słabą dynamikę interesującego nas sygnału konieczne jest zastosowanie wyrafinowanych metod analizy szeregów czasowych. Na etapie wstępnego przetwarzania zastosować można analizę składowych głównych *PCA* lub nowsze algorytmy analizy składowych niezależnych *ICA*. *PCA*, *ICA* i pozwalają zbadać inne algorytmy rozkładają serię obrazów na niezależne „mapy” neuronowe i przebiegi czasowe aktywności neuronów w tych fragmentach tkanki nerwowej. Tak przygotowane obrazy lepiej nadają się do automatycznej klasyfikacji za pomocą sieci neuronowych.

## 1 Wprowadzenie

Badania medyczne dostarczają bardzo dużej ilości obrazów trudno poddających się interpretacji. Jednym z największych osiągnięć ostatniego dziesięciolecia jest upowszechnienie się metod funkcjonalnego rezonansu jądrowego (fMRI), pozwalającego na obrazowanie działania mózgu w czasie jego normalnej pracy.

W typowym eksperymencie z użyciem techniki fMRI badana osoba wykonuje cykl czynności przeplatanych okresami prostych czynności kontrolnych. Dla przykładu badając zdolności językowe badanemu wyświetla się przez kilkanaście sekund kilka zdań, z których część zawiera błędy syntaktyczne. Badany próbuje ustalić i zapamiętać, które to zdania. W fazie kontrolnej przez kilkanaście sekund badany skupia się na poruszaniu palcami lub wpatruje się w jakiś obiekt wyświetlany na ekranie. W czasie eksperymentu wykonywane jest ciągle skanowanie mózgu badanej osoby. Techniki MRI mają liczne niedoskonałości – długi

czas pojedynczego cyklu badań, słaby sygnał, wymagający wielokrotnego powtórzenia tych samych czynności w celu zwiększenia odstępów sygnał-szum – podczas jednego eksperymentu te same czynności wykonuje się czasami kilkadziesiąt razy. Dlatego eksperymenty trwają długo, nieraz godzinę lub dwie. Łączy się to oczywiście ze sporą niedogodnością dla osoby badanej, zwłaszcza z powodu dużego hałasu aparatury i niemożności poruszania się. Metody komputerowej obróbki obrazów mogą skrócić czas pobytu w urządzeniu MRI zachowując lub nawet poprawiając jakość wyników.

W tym artykule opiszemy tylko metody wstępnego przetwarzania danych, przygotowujące serię obrazów do automatycznej klasyfikacji, skupiając się na metodach przydatnych do analizy obrazów fMRI. Należą do nich statystyczne metody korelacyjne i metody oparte na analizie częstości, oraz metody analizy składowych głównych (PCA, Principal Component Analysis), analizy cech lokalnych (LFA, Local Feature Analysis) i składowych niezależnych (ICA, Independent Component Analysis). Opiszemy też implementację i zastosowanie tych metod do obrazów otrzymanych z eksperymentów fMRI. Pracę kończy krótka dyskusja przydatności tego rodzaju metod do analizy zmiennych w czasie obrazów.

## 2 Badania fMRI

Obrazowanie MRI polega na efekcie precesji jądrowego spinu magnetycznego atomów wodoru zawartego w tkankach miękkich ciała pod postacią wody. Badany umieszczany jest wewnątrz magnesu, gdzie panuje stałe pole magnetyczne o indukcji rzędu kilku tesli. Pole to „porządkuje” spiny atomów w jednym kierunku powodując precesję spinu dookoła wyznaczonej osi. Wirujące spiny powodują powstanie fali elektromagnetycznej odbieranej przez tomograf. Manipulując polem elektromagnetycznym wytwarzanym przez tomograf jesteśmy w stanie pobudzić atomy z różnych miejsc (wokseli – trójwymiarowych pikseli) do precesowania z różną częstotliwością. Wykonując na zmierzonym sygnale dwuwymiarową transformatę Fouriera uzyskujemy obraz tkanek w mierzonym obszarze. Naraz tomograf obrazuje jedną płaszczyznę prostopadłą do osi głównego magnesu. Modyfikując pole magnetyczne dostrajamy do rezonansu atomy z różnych warstw. Pozwala to zeskanować całą, interesującą nas objętość, warstwa po warstwie.

Odmianą badań MRI są badania funkcjonalne (fMRI), gdzie interesującym nas obszarem skanowania jest ludzki mózg podczas normalnego funkcjonowania. Podstawowym efektem umożliwiającym badania funkcjonalne jest zależność zużycia tlenu od aktywności neuronów. Większe zużycie tlenu oznacza większą aktywność neuronów w danej części mózgu. Tomograf fMRI jest przestrojony do rezonansu z atomami żelaza zawartymi w hemoglobinie zamiast z atomami wo-

doru. Mierzmy sygnał generowany przez różnicę stężeń oksy- i deoksyhemoglobiny we krwi, która proporcjonalna jest do różnicy podatności magnetycznej krwi natlenionej i odtlenionej (sygnał ten nazywany jest BOLD, od **B**lood **O**xxygen **L**evel **D**ependent). Nowoczesne techniki tomograficzne takie jak technika echo-planarna pozwalają zdjąć cały skan interesującej nas objętości mózgu w 1.5-2.5 sekundy; daje to do tysiąca trójwymiarowych obrazów z jednego badania. Następnie wyniki są analizowane metodami komputerowej analizy obrazów w celu identyfikacji grup neuronów zaangażowanych w narzucone badaniem zadania.

Oprócz właściwego skanu „funkcjonalnego” wykonywanego na podstawie stałej relaksacji  $T_2^*$  (relaksacja poprzeczna typu spin-spin sąsiednich protonów), zdejmuje się też kilka skanów „anatomicznych” opartych na stałej relaksacji  $T_1$  (relaksacja podłużna spin-sieć) w celu użycia ich jako „tła” przy identyfikacji aktywnych grup neuronów.

Skany wykonywane są w płaszczyźnie strzałkowej, prostopadłej do najdłuższej osi ciała. Po zebraniu wszystkich skanów składających się na trójwymiarowy obraz jesteśmy w stanie wyprodukować dowolny przekrój objętości mózgu. Oprócz płaszczyzny strzałkowej często używamy też płaszczyzny koronalnej i sagittalnej jakkolwiek cięcie objętości może być dokonane wzdłuż dowolnej płaszczyzny – zależy to tylko od oprogramowania, którym dysponujemy.

### 3 Wstępna obróbka skanów

Skany funkcjonalne uzyskane ważeniem  $T_2^*$  wymagają wstępnej obróbki przed przystąpieniem do wykrywania aktywnych obszarów. Można wyróżnić kilka źródeł zniekształceń wpływających na jakość obrazu:

1. poruszenia ciała podczas procesu skanowania (zarówno fizjologiczne, np. oddychanie czy bicie serca, jak też przypadkowe)
2. artefakty przepływowe powstające w pobliżu dużych naczyń krwionośnych
3. szum pomiarowy wprowadzany przez tomograf i „systemowy” powodowany przez przypadkowe wyładowania neuronów
4. termiczne niestabilności aparatury pomiarowej („pływanie zera”)
5. artefakty spowodowane obecnością ciał metalicznych, np. plomb w zębach.

Niektóre z wymienionych efektów są nie do usunięcia na drodze obróbki gotowych skanów, np. (2) czy (5); pozostaje jedynie usunięcie uszkodzonych ramek z danych. Inne możemy się przynajmniej starać usunąć.

Małe poruszenia ciała, które powodują „rozmywanie” obrazu można zniwelować poprzez proces **rejestracji** obrazu (Friston i in., 1996). Polega on na obliczeniu parami korelacji przebiegów czasowych dla każdego woksela z danej płaszczyzny. Korekcji ruchu dokonuje się zazwyczaj tylko w płaszczyźnie strzałkowej zakładając, że badany ma znacznie ograniczone możliwości poruszania się wzdłuż osi skanera. Dysponując współczynnikami korelacji wzajemnej można wyznaczyć wektor przesunięcia ramki względem pozostałych i odpowiednio ją skorygować.

Pływanie zera najłatwiej skorygować przeprowadzając dla przebiegu każdego woksela filtrowanie górnoprzepustowe, np. metodą ruchomej średniej (moving average). Od wartości jasności woksela odejmujemy średnią jasność policzoną na N poprzednich wkselach, gdzie N zwykle wynosi 5-20. Proces ten likwiduje wpływ długoterminowych niestabilności aparatury, co jest ważne przy tak długim czasie trwania badania. Filtrowanie takie usuwa także kontury mózgu ze skanu.

Z nadmiernym szumem radzimy sobie stosując filtrowanie dolnoprzepustowe z odpowiednio dobraną częstotliwością odcięcia filtru. Oprócz szumu usuwamy także składowe wysokoczęstotliwościowe tworzące w ogonie widma sygnału szeroki pik przypisywany efektowi zmiany przenikalności magnetycznej ciała przy oddychaniu (efekt ten powoduje zmiana objętości klatki piersiowej).

Metody wstępnej obróbki obrazów fMRI opisano dokładnie w (Krugel i in., 1998).

## **4 Metody detekcji aktywacji funkcjonalnej**

Konwencjonalna analiza statystyczna oferuje kilka metod analizy obrazów fMRI, z których najważniejsze są dwie. Bazują one na analizie korelacyjnej i analizie częstotliwości.

### **4.1 Analiza korelacyjna**

W pierwszej metodzie konstruujemy specjalną funkcję opisującą ogólny przebieg eksperymentu. Najprostsza jej wersja przyjmuje wartość jeden dla momentów czasu kiedy badany wykonuje zadanie kognitywne i zero dla zadania kontrolnego. W praktyce należy uwzględnić opóźnienia w odpowiedzi hemodynamicznej neuronów. Prawdziwy kształt tej odpowiedzi nie jest funkcją prostokątną, przypomina raczej różnicę dwóch funkcji sigmoidalnych, i trudno go jest otrzymać z modeli symulacyjnych. W najprostszym przypadku można przyjąć funkcję prostokątną. Funkcja ta jest niezależnie korelowana z przebiegiem czasowym każdego woksela

za pomocą obliczania współczynnika korelacji Pearsona:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

gdzie  $\bar{x}$  ( $\bar{y}$ ) oznacza średnią ze wszystkich  $x_i$  ( $y_i$ ),  $y_i$  to jasność wokselu w  $i$ -tej ramce, zaś  $n$  to całkowita liczba ramek. Woksele, dla których współczynnik korelacji z funkcją przebiegu eksperymentu przekraczają pewien ustalony próg, są oznaczane jako aktywne. W ten sposób z surowych danych z całego eksperymentu otrzymujemy jedną mapę opisującą region mózgu, którego aktywność skorelowana jest z zadaniami wykonywanymi przez badanego.

Ustalenie właściwej funkcji odpowiedzi hemodynamicznej jest kluczowe dla poprawnego działania analizy korelacyjnej. Generalnie odpowiedź opóźniona jest około 5-6 sekund względem bodźca i rozmyta w czasie na 3-4 sekundy. Niewłaściwe ustalenie opóźnienia powoduje natychmiastowy spadek współczynnika korelacji i „zgubienie” aktywnych wokseli. Dla przykładu zbyt małe opóźnienie powoduje zaznaczenie jako aktywnych wokseli leżących w obrębie tętnic niosących natlenioną krew w aktywne rejony dając jednocześnie błędną identyfikację miejsc aktywności kognitywnej.

## 4.2 Analiza częstotliwości

Opierając się na cyklicznej konstrukcji eksperymentu fMRI możemy wywnioskować, że aktywacja interesujących nas regionów mózgu powinna być również cykliczna z okresem równym okresowi powtarzania bodźca. Wykonując analizę fourierowską (FFT) na przebiegu woksela uzyskujemy omówione wcześniej widmo pozbawione już najniższych częstotliwości („pływanie zera”, kontury mózgu) i najwyższych częstotliwości (szum, pik oddechowy). W tym widmie wyszukujemy piki odpowiadające częstotliwości powtarzania bodźca. Z reguły oprócz fundamentalnej harmonicznnej widoczna jest też druga harmoniczna. trzecia harmoniczna leży zazwyczaj już w zakresie piku oddechowego i zostaje usunięta na etapie obróbki wstępnej.

Na podstawie pomiaru stosunku amplitud pików częstotliwości harmonicznnych z częstotliwością powtarzania bodźca do reszty częstotliwości (tła) możemy orzekać o aktywności bądź braku aktywności danego woksela. Woksele z amplitudą tych pików przekraczającą zadany próg są oznaczane jako aktywne i tworzą mapę aktywności mózgu.

Wadą tej metody jest uzależnianie aktywności woksela tylko od *amplitudy* widma sygnału. Zupełnie ignorowana jest tutaj informacja niesiona przez *fazę* wid-

ma. Przez to wkład do aktywności woksela dają wszystkie możliwe synały pojawiające się z daną częstotliwością, z dowolnym opóźnieniem względem początku cyklu co odpowiada dowolnej fazie. W świetle dyskusji z poprzedniego podrozdziału prowadzi to do wyznaczenia zbyt dużego zbioru aktywnych wokseli. Dla poprzednio przytoczonego przykładu mapa aktywności będzie zawierać zarówno aktywne neurony jak i tętnice doprowadzające do nich krew.

### 4.3 Metody multiwariacyjne

Dotychczas omówione metody analizy opierały się wyłącznie na danych z jednego woksela przy ustalaniu stopnia aktywności tegoż woksela. Były to zatem metody uniwariacyjne. W drodze poszukiwań lepszych metod sięgnięto po metody multiwariacyjne wykorzystujące zarówno zależności przestrzenne jak i czasowe sygnału. Problem jest trudny, gdyż wymaga analizy czterowymiarowych danych (trzy wymiary przestrzenne plus czas). Aby zmniejszyć nakład pracy potrzebny do wykonania takiej analizy można bezpiecznie założyć, że nie istnieją zależności przestrzenne między kolejnymi skanami w płaszczyźnie strzałkowej (wzdłuż osi tomografu). Oczywiście nie jest to do końca prawdą, gdyż takie zależności na pewno istnieją, jednak za takim założeniem stoi sam proces obrazowania MRI. Woksele należące do tej samej płaszczyzny strzałkowej uzyskiwane są dwuwymiarową odwrotną transformatą Fouriera. Aby zmienić głębokość skanu (położenie płaszczyzny strzałkowej, w której wykonujemy skanowanie) potrzeba manipulacji gradientem pola magnetycznego wzdłuż osi strzałkowej, by dostroić atomy z żądanej płaszczyzny do rezonansu jądrowego. Powoduje to pewien, nieunikniony odstęp pomiędzy kolejnymi płaszczyznami skanów (*gap*). Można zatem domniemywać, że woksele między skanami są słabiej związane zależnościami przestrzennymi niż woksele należące do jednego skanu.

Podstawowymi metodami multiwariacyjnymi są analiza składowych głównych (**P**incipal **C**omponent **A**nalysis) i ostatnio wynaleziona analiza składowych niezależnych (**I**ndependent **C**omponent **A**nalysis) omówione w następnych rozdziałach.

## 5 PCA, analiza składowych głównych

Istotne polepszenie dokładności analizy zmienności szeregów czasowych obrazów przynieść mogą metody pozwalające na odrzucenie nieistotnych składowych sygnału. Analiza składowych głównych jest dobrze znaną metodą redukcji wymiarowości danych.

Wektor obserwowanych danych  $v$  (w wypaku fMRI zawiera on pojedynczy skan) jest liniowo transformowany w wektor  $x = Pv$  tak, by sygnały  $x$  były parami

nieskorelowane:

$$\langle xx^T \rangle = I \quad (2)$$

Macierz transformacji  $P$  możemy znaleźć obliczając wartości i wektory własne macierzy autokowariancji (bądź autokorelacji) wektora  $v$ . Jeżeli oznaczymy przez  $D$  macierz zawierającą na głównej przekątnej wartości własne macierzy  $\langle vv^T \rangle$ , zaś przez  $E$  macierz odpowiednich wektorów własnych, macierz transformacji PCA obliczamy według wzoru:

$$P = D^{-\frac{1}{2}} E^T \quad (3)$$

Dowód jest prosty:

$$\begin{aligned} \langle xx^T \rangle &= \langle Pv(Pv)^T \rangle = P \langle vv^T \rangle P = PEDE^T P^T = \\ &= D^{-\frac{1}{2}} E^T EDE^T E (D^{-\frac{1}{2}})^T = D^{-\frac{1}{2}} DD^{-\frac{1}{2}} = I \end{aligned} \quad (4)$$

Wektory zawarte w macierzy  $P$  nazywamy *składowymi głównymi*. Rzutując dane  $v$  na te wektory otrzymujemy parami zdekorelowane dane  $x$ . Taka transformacja zapewnia znalezienie ortogonalnych kierunków największej wariancji danych. Składowa główna, związana z największą wartością własną (wartości własne macierzy autokowariancji są nieujemne) reprezentuje kierunek w przestrzeni danych, w którym dane te wykazują największą wariancję. Sama wariancja dana jest przez wartość własną powiązaną z pierwszą składową główną. Druga składowa główna opisuje następny w kolejności ortogonalny kierunek w przestrzeni o największej wariancji danych itd.

Zwykle tylko kilka pierwszych składowych głównych jest odpowiedzialnych za większość wariancji danych. Dane wyrzutowane na pozostałe składowe główne często nie przekraczają swoją amplitudą szumu pomiarowego i mogą być bezpiecznie usunięte, zmniejszając w ten sposób wymiarowość danych bez istotnego uszczerpkienia informacji w nich zawartej.

Metoda ta może być użyta w charakterze filtru odcinającego szum w danych. Decydujemy się na pierwsze  $N$  składowych głównych odrzucając pozostałe i odtworzamy sygnał tylko na podstawie zachowanych składowych.

Składowe główne są wzajemnie ortogonalne i nie są całkowicie niezależne, co w niektórych przypadkach nie daje dobrej reprezentacji danych, prowadząc do powolnej zbieżności. Interpretacja sensu składowych głównych może czasami być niejasna.

## 6 ICA, analiza składowych niezależnych

### 6.1 Idea

Niezależność statystyczna jest czymś innym niż usunięcie korelacji wzajemnej między zmiennymi losowymi.

Założmy, że mamy  $M$  zmiennych losowych  $v_1, v_2, \dots, v_M$ , będących kombinacjami liniowymi  $N$  niezależnych czynników  $s_1, s_2, \dots, s_N$ . Niezależność oznacza, że wspólny rozkład prawdopodobieństwa wszystkich  $N$  składowych powinien być iloczynem rozkładów prawdopodobieństwa:

$$p(s_1, s_2, \dots, s_N) = \prod_{k=1}^N p_k(s_k) \quad (5)$$

Hyvärinen opierając się na powyższym wzorze twierdzi, że niezależność statystyczna charakteryzuje się zerową kroskorelacją tak jak w PCA, ale nie tylko dla korelacji drugiego rzędu (statystyka drugiego rzędu), ale wszystkich wyższych rzędów (Hyvärinen i Oja, 1996). Prowadzi to do niewygodnego czasami założenia, że składowe niezależne są pozbawione korelacji drugiego rzędu.

Alternatywną definicję proponują Bell i Sejnowski (Bell i Sejnowski, 1995). Ich wersja algorytmu opiera się na tzw. regule *InfoMAX* polegającej na minimalizacji wspólnej informacji w składowych niezależnych. Ilość informacji opisywana jest *entropią* sygnału:

$$H(X) = -\sum_k p_k \log p_k \quad (6)$$

gdzie  $p_k$  jest prawdopodobieństwem  $k$ -tego zdarzenia. Wspólna entropia dwóch zmiennych losowych nie jest addytywna:

$$H(X, Y) = H(X) + H(Y) - I(X, Y) \quad (7)$$

$$I(X, Y) = H(X) - H(X|Y) \quad (8)$$

gdzie  $I(X, Y)$  jest wspólną informacją zmiennych losowych  $X$  i  $Y$ . Algorytm ICA minimalizuje tę wartość dążąc do wyizolowania składowych statystycznie niezależnych.

Jeżeli ułożymy obserwabę  $v_i$  w wektor  $v$ , zaś składowe niezależne  $s_i$  odpowiednio w wektor  $s$ , możemy zapisać zależność między tymi wektorami:

$$v = \sum_{k=1}^N a_k s_k = As \quad (9)$$



gdzie  $A$  jest macierzą, która opisuje proces mieszania sygnałów źródłowych w obserwowane kombinacje. Celem algorytmu ICA jest obliczenie macierzy  $W$ , która spełnia warunek:

$$s \sim Wv \quad (10)$$

Rozkład obserwowanych sygnałów na statystycznie niezależne składowe bez żadnej dodatkowej wiedzy np. o apriorycznym rozkładzie prawdopodobieństw, nazywamy „ślepa separacją” (Jutten i Herault, 1991). Algorytm rozwiązujący to zagadnienie po raz pierwszy zaproponowany został przez Comona w pracy (Comon, 1994).

Zauważmy, że macierz  $W$  nie musi być wcale macierzą odwrotną do  $A$ ; może być to macierz przeskalowana przez jakiś czynnik, a także może mieć poprzestawiane wiersze względem macierzy  $A^{-1}$ . Pierwsza własność  $W$  powoduje niemożność odtworzenia energii i znaków składowych niezależnych, ponieważ dowolna stała, przez którą przemnożymy wektor  $s$ , może być zniwelowana przez dzielenie odpowiedniej kolumny macierzy  $A$  przez tą stałą. Drugie ograniczenie nie jest żadną niedogodnością, ponieważ powoduje tylko permutację składowych niezależnych.

## 6.2 Wybielanie

ICA wymaga jako wstępnego kroku obróbki sygnałów tzw. **wybielania** lub **sferowania** danych. Ten krok jest niczym innym jak analizą składowych głównych. Dla wektora  $v$  obliczane są dane wyrzutowane na składowe główne  $x$ , które są podstawą dalszej analizy. Następnie normalizowana jest wariancja składowych głównych w każdym kierunku oddzielnie. Rzutowanie na wzajemnie ortogonalne kierunki wektorów składowych głównych powoduje rozłożenie danych na hiperelipsie, dalsza normalizacja wariancji przekształca hiperelipsę w hipersferę – stąd nazwa „sferowanie”.

Na tym etapie można także zredukować wymiarowość danych korzystając z właściwości analizy PCA. Pozwala to wstępnie odfiltrować część szumu z danych. Jest to krok ważny dla analizy ICA, gdyż działa ona poprawnie tylko wtedy gdy znamy dokładną liczbę składowych niezależnych w zmieszonym sygnale. Jeśli przystąpimy do wydzielenia mniejszej ilości składowych otrzymamy nie składowe niezależne lecz takie ich kombinacje liniowe, w których wspólna informacja (redundancja) jest zminimalizowana. Okazuje się, że liczba składowych niezależnych może być oceniona na podstawie liczby składowych głównych, dla których wartości własne macierzy kowariancji przyjmują znaczące wartości. Reszta składowych głównych reprezentuje prawdopodobnie szum i może być bez szkody dla sygnału usunięta na tym etapie analizy.

### 6.3 Eksces

Aby rozdzielić składowe niezależne nie możemy używać kryterium niezależności (równanie 5) wprost, ponieważ nieznane są nam źródłowe rozkłady prawdopodobieństwa (problem *ślepej separacji*). Wiele realizacji algorytmu ICA używa *ekscesu* (kumulanta statystycznego czwartego rzędu) jako kryterium poprawnej separacji.

Eksces (ang. *kurtosis*):

$$kurt(v) = E\{v^4\} - 3(E\{v^2\})^2 \quad (11)$$

opisuje „niegaussowskość” zmiennej losowej  $v$ . Dla zmiennych gaussowskich eksces wynosi zero; dla rozkładów posiadających piki eksces jest dodatni (tym większy im większa jest „ostrość” pików). Dla płaskich rozkładów eksces jest ujemny.

Możliwe jest dwojakię podejście do ICA: albo statystykę wyższych rzędów kodujemy w algorytmie wprost używając momentów statystycznych wyższych rzędów (podejście zastosowane w pionierskiej pracy (Comon, 1994)), albo używamy odpowiednich funkcji nieliniowych umieszczając wyższe rzędy statystyczne w algorytmie *niejawnie*. Wiele algorytmów ICA opiera się na maksymalizacji (co do modułu) ekscesu zmiennych. Ma to uzasadnienie, gdyż mieszanina sygnałów ma rozkład bardziej przypominający „gaussowski” niż rozkłady samych sygnałów. Maksymalizacja modułu ekscesu powoduje poszukiwanie składowych niezależnych mających rozkład tak daleki od gaussowskiego jak to możliwe. Wyjaśnia to dlaczego ICA nie potrafi wydzielić sygnałów o rozkładzie normalnym. Użycie ekscesu jako miary odstępstwa od rozkładu normalnego uniemożliwia wydzielenie składowych o gaussowskim rozkładzie prawdopodobieństwa. Na szczęście wiele interesujących sygnałów, w tym szeregi obrazów otrzymywanych w eksperymentach fMRI, mają rozkłady niegaussowskie.

### 6.4 Implementacja neuronowa

Do konstrukcji sieci neuronowej realizującej algorytm ICA powszechnie wykorzystuje się nieliniowe reguły uczenia Hebba oraz reguły antyhebbowskie. Reguły takie mają generalną postać:

$$\Delta w \sim \pm xg(w^T x) + \dots \quad (12)$$

Zmiana wektora wag zależy od wartości sygnałów wejściowych  $x$  oraz od wyjściowego sygnału neuronu nieliniowo zależnego od aktywacji równej ważonej sumie sygnałów wejściowych  $w^T x$ . Znak  $\pm$  oznacza wersje uczenia: hebbowską i antyhebbowską.

Reguły uczenia oparte na ekscesie danych mają postać (za (Hyvärinen i Oja, 1996)):

- dla uczenia hebbowskiego

$$w(t+1) = w(t) + \mu(t)[x(t)g(w(t)^T x(t)) - w(t)] \quad (13)$$

- dla uczenia antyhebbowskiego

$$w(t+1) = w(t) + \mu(t)[-x(t)g(w(t)^T x(t)) - w(t)] \quad (14)$$

gdzie  $w$  jest wektorem wag neuronu,  $\mu$  jest zmienną w czasie stałą uczenia,  $t$  jest krokiem iteracji, zaś  $g$  jest odpowiednią funkcją nieliniową (dla przykładu funkcją logistyczną lub tangensem hiperbolicznym). Reguły te gwarantują obliczenie jednej składowej niezależnej. Pierwszy człon w nawiasach kwadratowych odpowiada uczeniu hebbowskiemu/antyhebbowskiemu, a drugi jest członem normalizującym, zapobiegającym nadmiernemu wzrostowi wag. Wybór odpowiedniej funkcji nieliniowej  $g$  opisano szeroko w (Hyvärinen i Oja, 1998).

Różnica między oboma rodzajami uczenia, we wzorze niepozornie sprawdzająca się do zmiany znaku, jest kluczowa w separacji składowych niezależnych. Uceniem hebbowskim jesteśmy w stanie wyseparować tylko składowe o ekscesie ujemnym; uczeniem antyhebbowskim tylko składowe o ekscesie dodatnim. Często nie znamy ilości składowych należących do każdej z tych dwóch klas, a nawet ilości składowych w ogóle. Remedium jest wprowadzenie jednej, uniwersalnej reguły uczenia:

$$w(t+1) = w(t) + \mu(t)[\sigma x(t)g(w(t)^T x(t)) - w(t)] \quad (15)$$

gdzie  $\sigma$  to funkcja opisująca znak ekscesu separowanej składowej. Możemy wystartować z dowolnego znaku i dynamicznie, podczas uczenia sieci estymować ten znak wprowadzając na bieżąco zmiany w regule uczenia różnych neuronów separujących konkretne składowe. W ten sposób mamy pewność, że nie pominiemy żadnej składowej o ile liczba neuronów jest wystarczająca.

Jeden neuron wystarcza do odseparowania jednej składowej niezależnej z wybielonych danych. Liczba składowych niezależnych, jaką możemy separować równa jest liczbie wybielonych zmiennych, którymi dysponujemy. Odbywa się to po prostu przez zestawienie neuronów uczonych powyższą regułą w warstwę sieci neuronowej. Potrzebujemy także jakiegoś rodzaju sprzężenia zwrotnego, aby zapobiec zbieganiu się neuronów tej warstwy do tych samych składowych niezależnych. Najczęściej jest to bezpośrednie sprzężenie zwrotne z wyjść do wejść neuronów, wprowadzające czynnik kary za zbieganie się do tych samych składowych. W podejściu Hyvärinena może to być np. hierarchiczna ortogonalizacja

wyjść neuronów. Można także odpowiednio dobrać wstępne aktywacje, tak by były położone możliwie w basenach atrakcji różnych składowych niezależnych.

Zbieżność algorytmu silnie zależy od wyboru właściwej stałej uczenia. Zbyt duże  $\mu$  niszczy jakąkolwiek zbieżność, a zbyt małe powoduje nieefektywne uczenie, lub wręcz jego brak (ze względu na nieuniknione błędy numeryczne).

Dla algorytmu InfoMAX odpowiednia reguła uczenia wygląda następująco (za (McKeown i in., 1998a)):

$$C = WX \quad (16)$$

$$y = g(C) \quad (17)$$

$$\Delta W = -\mu \left( \frac{\partial H(y)}{\partial W} \right) W^T W = \mu (I + \hat{y} C^T) W \quad (18)$$

$$\hat{y} = \frac{\partial}{\partial C_i} \ln \left( \frac{\partial y_i}{\partial C_i} \right) \quad (19)$$

gdzie  $\mu$  to stała uczenia,  $H(y)$  dane jest wzorem (6), zaś  $X$  to macierz zawierająca wybielone dane. Funkcja nieliniowa zapewnia tu konieczną statystkę wyższych rzędów. Zwykle stosujemy to funkcję logistyczną:

$$g(C) = \frac{1}{1 + e^{-C}} \quad (20)$$

Uczenie przeprowadzamy w trybie batchowym zmieniając wagi dla kilku danych wejściowych jednocześnie zamiast za każdym razem gdy pobieramy daną wejściową. Zaletami tego trybu są większa stabilność procesu uczenia i większa prędkość uczenia gdy danych treningowych jest dużo. Uczenie kończymy gdy macierz wag  $W$  przestanie się zmieniać w sposób znaczący.

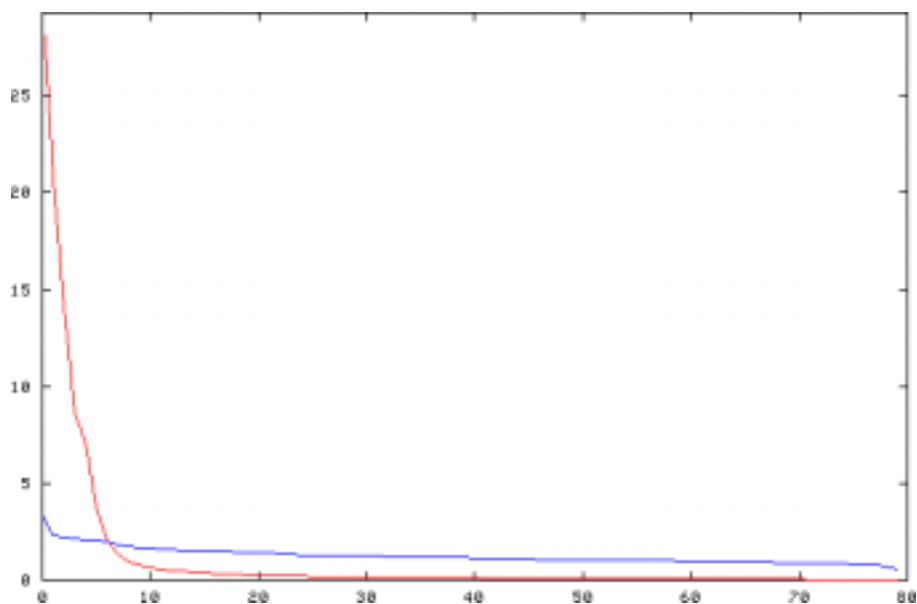
Neuronowe wersje algorytmu ICA są szeroko dyskutowane w (Oja i in., 1995; Cichocki i in., 1998) oraz pracy przeglądowej (Hyvärinen, 1999). Hyvärinen zaproponował także szybką nieneuronową wersję algorytmu dyskutowaną w (Hyvärinen i Oja, 1997).

Gotowe realizacje różnych algorytmów ICA w Matlabie można znaleźć w Internecie na stronach A. Hyvärinena (FastICA) por. (WWW, a) lub T. Bella (InfoMAX ICA) por. (WWW, b).

## 7 Rozkład PCA/ICA dla danych fMRI

Rozkład PCA nie wprowadza wiele nowego do analizy danych fMRI. Nie jest uzyskiwana prawie żadna separacja. Powodem tego jest bardzo mała amplituda sygnału użytecznego w danych. Typowo nie przekracza ona 5-10%. Tak mała amplituda sygnału oznacza jego małą wariancję, co z kolei powoduje, że zostanie

on uwzględniony dopiero w dalszych składowych głównych. Algorytm ICA „wyrównuje” wariacje zawarte w składowych niezależnych co widać na rysunku 1.

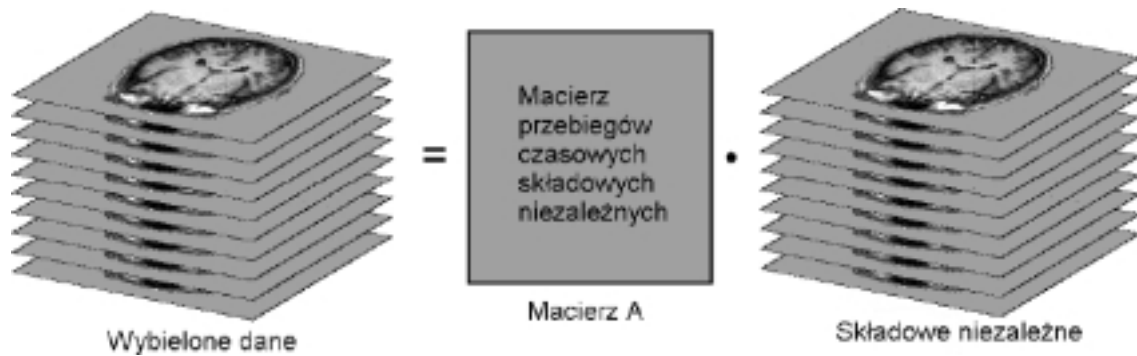


Rysunek 1: Procentowe wkłady do rekonstrukcji danych od składowych głównych (czerwona linia) i niezależnych (niebieska linia)

McKeown (McKeown i in., 1998a) zaproponował użycie algorytmu ICA dla obrazów fMRI. Jako zbiór zmiennych losowych traktujemy tutaj szereg czasowy skanów wykonanych na ustalonej głębokości. Możliwe jest dwojaki sposób do problemu ICA na danych fMRI: możemy poszukiwać składowych niezależnych *przestrzennie* lub składowych niezależnych *czasowo*. Przestrzenna niezależność oznacza, że wygenerowane mapy będą w większości lokalne i nienakrywające się. Proces ten wydaje się dobrze opisywać rzeczy dziejące się w mózgu ludzkim, gdyż korzysta z założenia, że neurony odpowiedzialne za jedną czynność są położone fizycznie blisko siebie. To podejście zastosowano w (McKeown i in., 1998a). Pojedynczy obiekt jest tu dwuwymiarowym obrazem – poszukiwane składowe niezależne będą zatem także obrazami dwuwymiarowymi. Oczekujemy od algorytmu ICA rozłożenia serii obrazów fMRI na zbiór niezależnych obrazów-map neuronowych, oraz przebiegów czasowych aktywacji neuronów podczas eksperymentu (rysunek 2). Na dalszym etapie obróbki możemy użyć konwencjonalnych metod statystycznych lub metod klasyfikacji do separacji interesujących nas przebiegów pokazujących aktywność różnych obszarów mózgu.

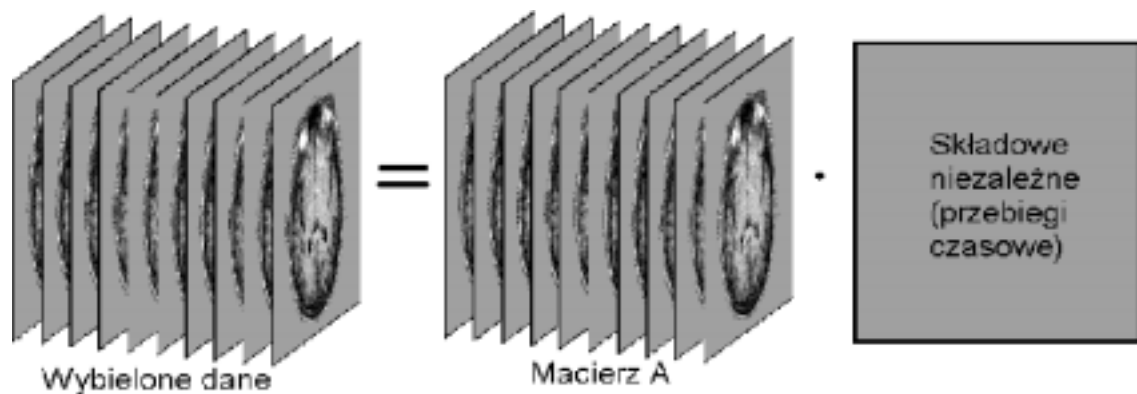
Separacja jest tutaj „ślepa”, gdyż informacja o eksperymencie pod postacią np. funkcji przebiegu eksperymentu opisanej w podrozdziale 4.1 potrzebna jest

dopiero na etapie selekcji interesujących nas map neuronowych.



Rysunek 2: Wersja algorytmu ICA wyodrębniająca *przestrzennie* niezależne składowe

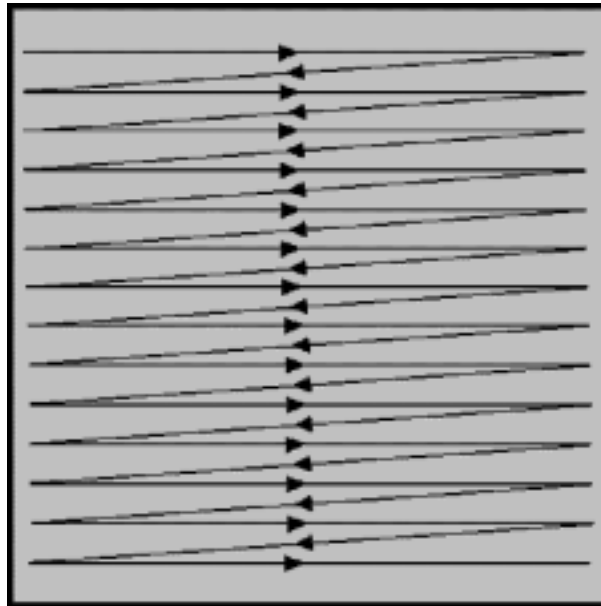
Rozkład na czasowo niezależne składowe powoduje wyodrębnienie map aktywności wokseli wykazujących niezależną od innych grup aktywność w czasie (rysunek 3). Takie grupy nie muszą być przestrzennie niezależne od innych co nie spełnia założenia o lokalizacji procesów zachodzących w mózgu. Poza tym rozkład na czasowo niezależne składniki jest obliczeniowo znacznie bardziej wymagający, głównie ze względu na rozmiary danych (ilość wokseli na płaszczyźnie jest znacznie większa niż ilość skanów danej płaszczyzny w czasie, co powoduje znacznie większe rozmiary macierzy kowariancji używanej w wybielaniu; obliczenie wartości i wektorów własnych większej macierzy zabiera znacznie więcej czasu).



Rysunek 3: Wersja algorytmu ICA wyodrębniająca *czasowo* niezależne składowe

Bazując na pakietach Hyvärinena i Bella/Sejnowskiego powstała w ramach programu badawczego Instytutu Maxa Plancka w Lipsku realizacja ICA (*vica*), służąca dekompozycji serii obrazów tomograficznych z badań fMRI. Opracowany program wchodzi w skład szerszego zbioru programów **BRIAN** por. (WWW, d) opartego na pakiecie Vista por. (WWW, c).

Dla potrzeb algorytmu ICA każdy skan jest przekształcany do postaci wektorowej przez zapis wartości wokseli linia po linii (rysunek 4).

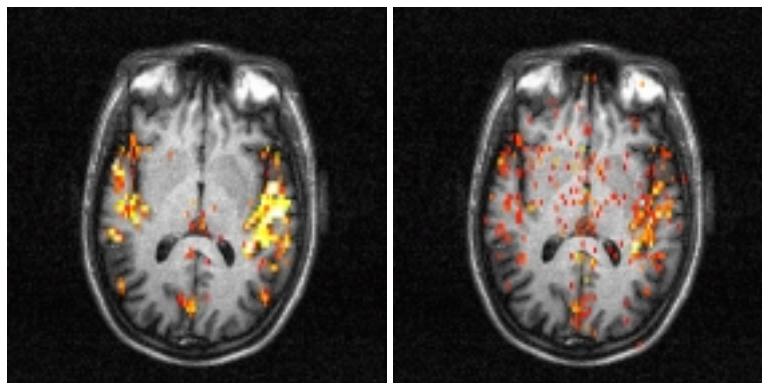


Rysunek 4: Przekształcenie dwuwymiarowego skanu do jednowymiarowego wektora

Tak przygotowane dane zestawiane są w macierz, której kolumny odpowiadają kolejnym wokselsom płaszczyzny skanu, zaś wiersze kolejnym skanom (momentom czasu). Po zastosowaniu ICA otrzymujemy macierz, której wiersze zawierają składowe niezależne, czyli mapy neuronowe. Poddając je teraz procesowi odwrotnemu do skanowania linia po linii otrzymujemy obraz dwuwymiarowy. Taki obraz poddajemy standaryzacji (z-scoring), obcinaniu powyżej pewnego progu aktywacji i rejestracji na odpowiadającym mu obrazie anatomicznym.

## 8 Przykład zastosowania ICA w analizie fMRI

Dla przykładu przeanalizujemy dane z eksperymentu związanego z rozumieniem języka. Zebrano 912 skanów  $T2^*$  na czterech głębokościach w 76 powtórzeniach



Rysunek 5: Obrazy aktywności mózgu otrzymane za pomocą analizy korelacyjnej i korelacji jednej z niezależnych składowych

testu oraz 4 skany anatomiczne  $T_1$ . Jeden cykl eksperymentu trwał 24 sekundy, z czego 6 sekund stanowiła prezentacja zdania, którego poprawność gramatyczną miał za zadanie ocenić badany. Pozostałe 18 sekund stanowiło odpoczynek. Cały eksperyment trwał około 30 minut. Rozdzielczość skanów wynosiła 128x64 woksele dla skanów funkcjonalnych (1.9x3.8x5mm) i 128x128 wokseli dla skanów anatomicznych (1.9x1.9x5mm).

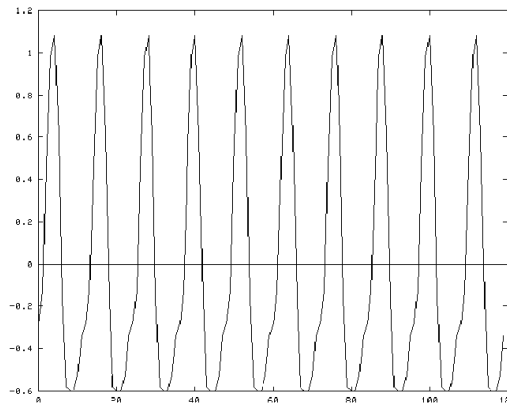
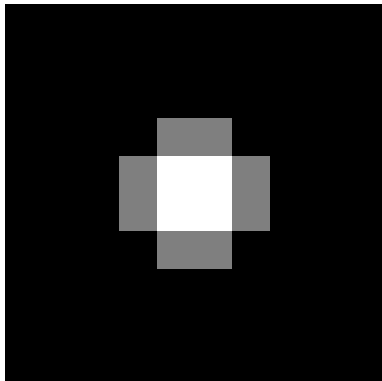
Skuteczność rozkładu pokazują dwa przedstawione tu obrazy. Lewy to mapa aktywności będąca rezultatem klasycznej analizy korelacyjnej. Prawy to jedna z wielu niezależnych składowych, o największym współczynniku korelacji z funkcją przebiegu eksperymentu (0.63).

Generalnie obydwie obrazy ukazują aktywność tych samych regionów mózgu. ICA jako „ślepa separacja” znajduje ponadto w tym przypadku aż 911 innych map aktywności neuronów (zrezygnowano z redukcji wymiarowości danych metodą PCA), z których część może być przypisana poruszeniom badanej osoby w tomografie, oddychaniu czy biciu serca.

Szczegółowe badania algorytmu ICA na wycinkach obrazów MRI ze sztucznie nałożoną aktywacją pokazują jego przydatność w separacji sygnałów. Na obraz nieaktywnego obszaru mózgu rozmiaru 10x10 została nałożona z pewną maską przestrzenną wymodelowana aktywacja neuronów. Następnie na tak spreparowanych danych uruchomiony został algorytm ICA InfoMAX. Otrzymane wyniki pokazuje rysunek 7. Analiza PCA użyta jako wstępny krok analizy (wybielanie) rozłożyła dane na 912 składowych głównych, z których kilkanaście pierwszych zawierało interesujące kształty i kontury np. tzw. artefakty łukowe), zaś reszta składowych niewiele odróżniała się wizualnie od szumu.

Korelacja najlepszej mapy ICA z sygnałem modulującym wynosiła 0.97. Znalezione tylko *jedną* mapę z tak wysokim współczynnikiem korelacji, co dowodzi





Rysunek 6: Maska i sygnał modulujący (aktywacja)

poprawnej dekompozycji obrazu na przestrzennie niezależne składowe.

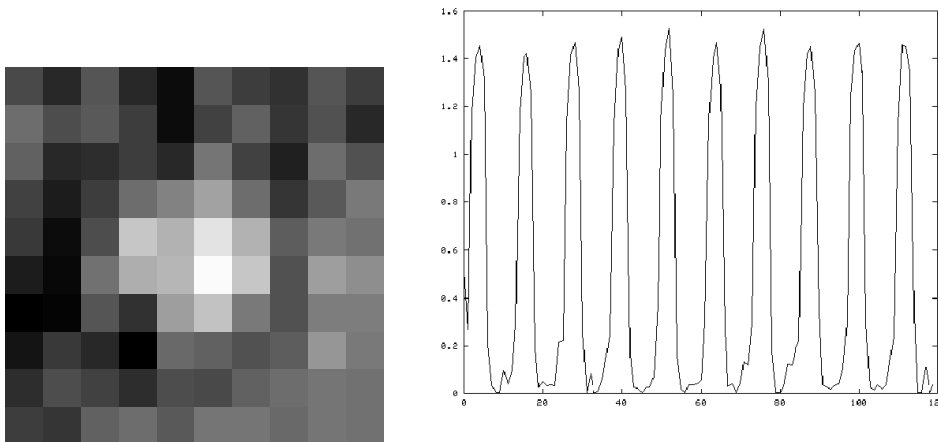
## 9 Podsumowanie

Wraz z postępem fizyki technicznej, oferującej coraz lepsze i szybsze urządzenia dla medycyny, rośnie potrzeba konstruowania efektywnych algorytmów dla celów analizy danych neurologicznych i ich obrazowania. ICA jest ograniczona do separacji *liniowych* kombinacji źródeł. Algorytm ten zgodny jest „ogólnym modelem liniowym” zaproponowanym przez Fristona:

$$X = G\beta + e \quad (21)$$

gdzie  $X$  jest macierzą obserwacji,  $G$  macierzą zawierającą przebiegi czasowe aktywacji źródeł, macierz  $\beta$  zawiera mapy źródłowe, zaś  $e$  jest szumem rezidualnym. Oczywiście takie założenie o liniowości procesów w mózgu jest bardzo restrykcyjne i uniemożliwia modelom czysto liniowym, takim jak PCA czy ICA, poprawną separację zjawisk nieliniowych, co zauważa sam Friston (Friston, 1998).

Kolejne ograniczenie ICA, dotyczące niemożności separacji sygnałów o rozkładzie normalnym nie jest istotne dla większości sygnałów. Rzeczywiste sygnały mają albo rozkład sub-gaussowski (np. obrazy), albo super-gaussowski (np. dźwięki). Mapy fMRI charakteryzują się zwykle rozkładem super-gaussowskim z dużą wartością ekscesu dzięki silnej lokalizacji klastrów neuronowych.



Rysunek 7: Najlepsza mapa ICA i jej aktywacja

ICA jest jeszcze bardzo młodą techniką obróbki danych. Pierwsze eksperymenty potwierdziły jej użyteczność w separacji różnych rodzajów danych, nie tylko obrazów. Z sukcesem użyto tej metody np. do separacji głosów różnych osób mówiących jednocześnie (tzw. *problem cocktail party*) (Lee i in., 1997). Martin McKeown użył z powodzeniem algorytmu ICA do analizy danych EEG (McKeown i in., 1998b).

Praktyczne zastosowanie tych technik do analizy szeregu obrazów napotyka jednak pewne trudności. Interpretacja poszczególnych składowych nie zawsze jest jasna, a interesujące (tj. skorelowane z przebiegiem czasowym eksperymentu) składowe mogą pojawić się wśród kilkuset niezależnych składowych. Różne wersje ICA dają różne rezultaty i nie ma kryterium, które pozwoliło by zdefiniować, które z tak otrzymanych obrazów są najbardziej interesujące (do pewnego stopnia taką rolę spełniać może współczynnik korelacji z funkcją opisującą przebieg eksperymentu). Nie przeprowadzono dotychczas badań porównawczych pomiędzy różnymi metodami analizy czynników niezależnych ani nie opracowano zestawu problemów, na których można by ocenić zalety i wady poszczególnych rozwiązań.

Lokalna Analiza Cech (Local Feature Analysis, LFA) (Penev i Atick, 1996) jest ciekawą propozycją wstępnego przetwarzania obrazów, bardzo przydatną w rozpoznawaniu obiektów. Jej założenia i realizacja są podobne do ICA. Zakłada się wstępny etap wybielania danych i dalszą obróbkę siecią neuronową, której reguły uczenia (odpowiednie czynniki kary) powodują otrzymanie nienakrywających się wzajemnie map nazywanych tu detektorami cech. W zastosowaniu do rozpoznawania twarzy znajduje lokalne cechy, takie jak nosy, oczy, usta czy policzki, pozwalając następnie algorytmom klasyfikującym skupić się na klasyfika-

cji istotnych cech obrazu. LFA jest obecnie bezkonkurencyjnym algorytmem w rozpoznawaniu obiektów, nie była jednak stosowana do analizy obrazów medycznych. W tej ważnej i trudnej dziedzinie pozostaje jeszcze dużo do zrobienia.

## 10 Podziękowania

A.Ch. dziękuje dr. Frithjofowi Kruggelowi (kruggel@cns.mpg.de), dr. Markusowi Svensenowi (svensen@cns.mpg.de) i dr. Martinowi J. McKeown'owi (martin.mckeown@duke.edu) za ich pomoc. Jego pobyt w Lipsku finansowany był z grantu Tempus nr SJEP 11415/96 kierowanego przez dr. Andrzeja Kowalczyka z Wydziału Fizyki i Astronomii, UMK. Za wsparcie finansowe jesteśmy też wdzięczni Komitetowi Badań Naukowych, grant nr. 8 T11F 014 14.

## Bibliografia

- Bell, A. i Sejnowski, T. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, t. 7, 1129–1159.
- Cichocki, A., Douglas, S., i Amari, S. (1998). Robust techniques for independent component analysis with noisy data. *Neurocomputing*, t. 22, 113–129.
- Comon, P. (1994). Independent component analysis – a new concept? *Signal Processing*, t. 36, 287–314.
- Friston, K. (1998). Modes or models: a critique on independent component analysis for fMRI. *Trends in Cognitive Sciences*, t. 2, nr 10, 373–375.
- Friston, K., Williams, S., Howard, R., Frackowiak, R., i Turner, R. (1996). Movement-related effects in fMRI time series. *Magn. Reson. Med.*, t. 35, 346–355.
- Hyvärinen, A. (1999). Survey on independent component analysis. *Neural Computing Surveys*, t. 2, 94–128.
- Hyvärinen, A. i Oja, E. (1996). Simple neuron models for independent component analysis. *Int. Journal of Neural Systems*, t. 7, 671–687.
- Hyvärinen, A. i Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. *Neural Computation*, t. 9, 1483–1492.
- Hyvärinen, A. i Oja, E. (1998). Independent component analysis by general non-linear hebbian-like learning rules. *Signal Processing*, t. 64, 301–313.

- Jutten, C. i Herault, J. (1991). Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, t. 24, 1–10.
- Kruggel, F., Descombes, X., i von Cramon, D. (1998). Preprocessing of fMRI datasets. W *Workshop on Biomedical Image Analysis (Santa Barbara)*, ss. 211–220, Los Alamitos. IEEE Computer Press.
- Lee, T.-W., Bell, A., i Orglmeister, R. (1997). Blind source separation of real world signals. W *Proceedings of IEEE International Conference Neural Networks*, ss. 2129–2135, Houston.
- McKeown, M., Makeig, S., Brown, G., Jung, T.-P., Kindermann, S., Bell, A., Iragui, V., i Sejnowski, T. (1998a). Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping*, t. 8, 368–372.
- McKeown, M., Makeig, S., Brown, G., Jung, T.-P., Kindermann, S., Lee, T.-W., i Sejnowski, T. (1998b). Spatially independent activity patterns in functional magnetic resonance imaging data during the Stroop color-naming task. *Proc. Natl. Acad. Sci. USA*, t. 95, 803–810.
- Oja, E., Karhunen, J., Wang, L., i Vigario, R. (1995). Principal and independent components in neural networks - recent developments. W *Proc. VII Italian Workshop on Neural Nets WIRN VIETRI95*, ss. 16–35.
- Penev, P. i Atick, J. (1996). Local feature analysis: a general theory for object representation. *Network*, t. 7, 477–500.
- WWW. Strony FastICA. <http://www.cis.hut.fi/projects/ica/fastica/>.
- WWW. Strony ICA InfoMAX. <http://www.cnl.salk.edu/~tony/ica.html>.
- WWW. Strony pakietu VISTA. <http://www.cs.ubc.ca/nest/lci/vista/vista.html>.
- WWW. Strony projektu BRIAN. [http://www.cns.mpg.de/Research/Groups/Group\\_sip/brian/main.html](http://www.cns.mpg.de/Research/Groups/Group_sip/brian/main.html).