



УДК 612.821:007

© 2010

Ю. А. Белов, О. В. Тарасюк

Моделювання альфа-хвиль комплекснозначними штучними нейронними мережами

(Представлено академіком НАН України В. Л. Макаровим)

Запропоновано гіпотезу про функціональну роль мозкових ритмів у процесах розпізнавання образів. В основі гіпотези лежить припущення про те, що функціональна роль мозкових ритмів полягає у забезпеченні явища інтерференції нервових імпульсів на сомі нервової клітини, у результаті чого одні й ті ж міжнейронні зв'язки залежно від фазового зсуву мають як збуджуючий, так і гальмуючий вплив. Згідно з цією гіпотезою розроблено та протестовано спеціалізовану архітектуру комплекснозначної штучної нейронної мережі, призначеної для побудови інваріантних ознак розпізнавання образів.

1. Опис експерименту. У 1947 р. Мак-Каллохом та Піттсом було запропоновано гіпотезу сканування [1], згідно з якою альфа-ритм відображує процес сканування зорової кори, тобто считування з неї інформації хвилею збудження. На думку авторів гіпотези, процес сканування забезпечує більш компакту структуру підсистем головного мозку. У 2001 р. І. А. Шевелев опублікував результати роботи [2], у якій експериментальним шляхом отримані безпосередні підтвердження гіпотези Мак-Каллоха та Піттса про скануючу роль альфа-хвиль. І. А. Шевелев вирішив перевірити, чи здатні ритмічні спалахи світла, що створюють у зоровій корі дифузну хвилю збудження, “зупинити” альфа-хвилю у деякий момент її руху й тим самим викликати її зорове сприйняття як нерухомого об'єкта. В експерименті І. А. Шевелева людина сиділа з заплющеними очима перед лампою стробоскопу, частоту спалахів якого комп'ютер змінював довільно. При виникненні зорових ілюзій досліджуванний натискав на кнопку і комп'ютер фіксував частоту, на якій з'являвся образ. Судячи з описів та малюнків сприйнятих зображень, досліджувані бачили на темному фоні світлі кільця, круги, спіралі або решітки з лініями, що косо перетинаються. Індивідуальна, домінуюча у спектрі потужності ЕЕГ частота альфа-ритму та оптимальна частота стимуляції для отримання чіткого ілюзорного ефекту виявилися тісно пов'язаними. На рис. 1 наведені результати цього експерименту.

У даній роботі на базі описаного вище експерименту пропонується розширена гіпотеза сканування, яка пояснює його основні результати. Перше припущення цієї гіпотези полягає у тому, що альфа-хвилі беруть участь не тільки у процесі транспортування інформації, але й виконують функцію її переробки, яка є принципово важливою у всьому процесі розпізна-

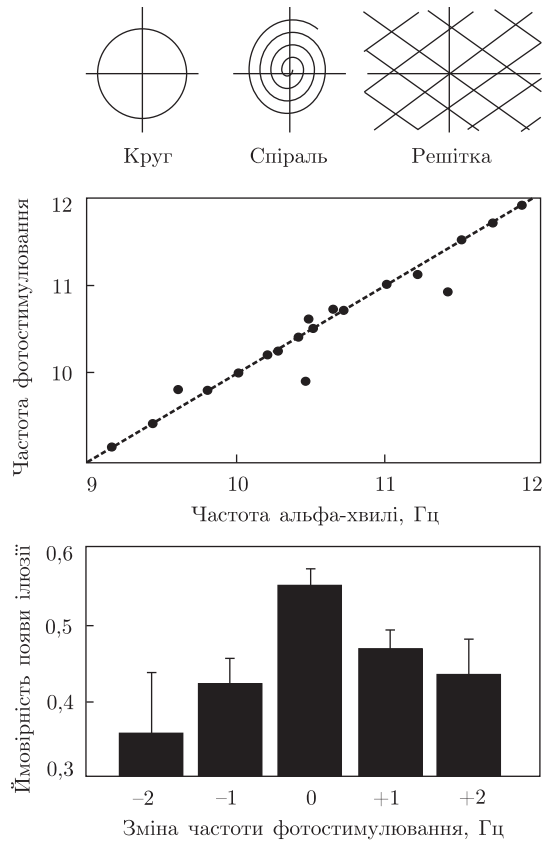


Рис. 1. Залежність ймовірності появи зорової ілюзії від частоти фотостимуляції (взято з [1])

вання зорової інформації. Механізм обробки інформації, який реалізується за допомогою альфа-хвиль, полягає у послідовному здійсненні інтегральних перетворень первинної зорової інформації, які наближено обчислюються на дискретній ґратці, утвореній з нейронів зорової кори головного мозку. Кожне інтегральне перетворення, що здійснюється за допомогою альфа-хвиль, виконується для зняття певної групи симетрії. Можна припустити, що дія спалаху стробоскопу подібна до дії дельта-функції Дірака, за допомогою якої можливо виділити окремі спектральні компоненти ядра інтегрального перетворення. Це дозволяє реконструювати самі інтегральні перетворення.

2. Аналіз даних експерименту. Незважаючи на те, що у реальному процесі переробки інформації мозком бере участь кінцева кількість нервових клітин, на етапі аналізу цією обставиною можна знехтувати, оскільки їх кількість є досить великою для того, щоб було кожен шар зорової кори наближено розглядати як неперервну двовимірну поверхню. У цьому наближенні характеристики окремих нейронів ми можемо розглядати як функції щільності цих характеристик, які неперервно розподілені у кожному шарі зорової кори. Нехай функцією

$$f(z) = f(x + iy) \in \mathbb{R}$$

задається розподіл інтенсивності зображення, що подається на вхідний шар нейронної мережі, де $(x, y) \in \mathbb{R}^2$ — декартові координати точки на вхідному шарі, які для зручності скомбіновані в одне комплексне число $z = x + iy \in \mathbb{C}$.

Розглянемо двовимірне комплекснозначне перетворення Фур'є:

$$g(k) = \frac{1}{2\pi} \iint e^{-i(k_1x+k_2y)} f(z) dx dy. \quad (1)$$

В експерименті спалахи стробоскопу сприймалися піддослідними із заплющеними очима, що давало приблизно рівномірний розподіл інтенсивності зображення на сітківці. Для функції $f(z) = f(x, y)$, симетричної за обома своїми аргументами, цей вираз набуде вигляду

$$g(k) = \frac{1}{2\pi} \iint \cos(k_1x) \cos(k_2y) f(z) dx dy.$$

Припустимо, що спалах стробоскопу обмежує час ефективної дії альфа-хвилі таким чином, що функція $g(k)$ вироджується у

$$g(k) = \delta(k - k').$$

Вираз вигляду

$$\cos(k'_1x) \cos(k'_2y)$$

описує двовимірну решітку з лініями, що косо перетинаються. Мета, з якою мозок може виконувати двовимірне перетворення Фур'є первинного зображення, досить очевидна. Таким чином можливо зробити процес розпізнавання образів інваріантним відносно зсувів зображення у просторі першої проєктивної зони сітківки. В цьому разі будь-який паралельний зсув об'єкта, що розпізнається, перетворюється у відповідний фазовий зсув його Фур'є-образу. Для того щоб цей фазовий зсув не завдавав значного впливу на подальший процес розпізнавання образів, його необхідно якимось чином усунути. У випадку комплекснозначного перетворення це робиться елементарно: досить результат перетворення взяти за модулем. У випадку дійснозначної нейронної мережі подібна арифметика виглядає занадто штучною для реалізації у біологічних нейронних мережах. Інша можливість полягає у тому, що фазовий зсув кодується деяким упередженням або затримкою нервового імпульсу, що сприймається нервовою клітиною, відносно деякого ритму, в якому осцилює клітина-приймач. Схожі явища дійсно спостережувались експериментально [3]. В такому разі одні й ті ж синаптичні зв'язки повинні забезпечувати як збуджуючий, так і гальмуючий вплив на нервову клітину залежно від того, як розподілені у часі всі інші імпульси, які надходять у цю клітину.

Тепер доречно сформулювати основне припущення гіпотези: процеси переробки інформації біологічними нейронними мережами у корі головного мозку організовані таким чином, щоб емулювати квантові ефекти, при цьому імпульси збудження окремих нейронів відіграють роль фотонів, а самі нейрони — роль атомів. Тоді найбільш продуктивним підходом до моделювання біологічних нейронних мереж має бути застосування формалізму, подібного до фейнманівського формалізму інтегрування за траєкторіями. Ймовірність взаємодії двох окремих нейронів дається квадратом комплекснозначної амплітуди ймовірності відповідного нервового імпульсу — “фотона” у точці його взаємодії з нервовою клітиною — “атомом”. У цьому випадку фазовий зсув компенсується автоматично за фактом переходу нейрона у певний визначений стан. Кожний синаптичний зв'язок характеризується деякою провідністю та фазовим зсувом, який додається до амплітуди ймовірності нервового імпульсу, що

проходить через цей зв'язок. У точці перетину альтернативних шляхів поширення нервових імпульсів вони інтерферують між собою, внаслідок чого одні й ті ж самі синаптичні зв'язки мають як збуджуючий, так і гальмуючий вплив на нервову клітину. В даному контексті необхідно відзначити, що картини, подібні до інтерференційної, досить часто з'являються в експериментах з дослідження біологічних нервових мереж: послідовна зміна зон збудження та гальмування в межах одного шару нервових клітин як реакція на елементарний подразник є скоріше правилом, ніж виключенням з нього [4].

Круги та спіралі, що спостерігалися в експерименті І. А. Шевелева, можуть мати аналогічне пояснення, як специфічної форми інтерференційної картини, породжувані звуженням спектра інтегрального оператора до однієї гармоніки. Розглянемо інтегральне перетворення

$$\omega(\lambda, \gamma) = \frac{1}{2\pi} \iint e^{-i\lambda\varphi - i\gamma \ln \rho} d\varphi d \ln \rho, \quad (2)$$

де $x + iy = \rho e^{i\varphi}$. Вираз $e^{-i\lambda\varphi - i\gamma \ln \rho}$, що фігурує під знаком інтеграла, може пояснювати форму оптичної ілюзії, стійко спостережуваної у досліді І. А. Шевелева. Дане інтегральне перетворення може бути застосовано для побудови простору ознак розпізнавання образів, інваріантних відносно поворотів та масштабувань первинного простору зображень. Дійсно, при будь-якому перетворенні вигляду

$$(\rho, \varphi) \rightarrow (\alpha\rho, \varphi + \psi)$$

маємо:

$$\omega(\lambda, \gamma) \rightarrow e^{i\lambda\psi + i\gamma \ln \alpha} \omega(\lambda, \gamma).$$

Як і у випадку двовимірного перетворення Фур'є, для того щоб виключити фазовий множник, який виникає у правій частині останнього виразу, досить взяти його за модулем, що відповідає обчисленню квадратного кореня ймовірності події на основі квантовомеханічної амплітуди ймовірності. Помічаючи, що повороти та масштабування первинного зображення призводять до відповідних поворотів та масштабувань його Фур'є-образу, ми можемо побудувати комбіноване двоетапне двовимірне перетворення, яке полягає у послідовному застосуванні перетворень (1) та (2), за допомогою якого будь-який двовимірний об'єкт перетворюється у сукупність ознак, інваріантних відносно довільних поворотів та масштабувань цього об'єкта. Це двоетапне перетворення має такий вигляд:

$$G(k) = \left| \frac{1}{2\pi} \iint e^{-ik_1x - ik_2y} \sqrt{f(x, y)} dx dy \right|^2; \quad k = k_1 + ik_2;$$

$$\Omega(\lambda, \gamma) = \left| \frac{1}{2\pi} \iint e^{-i\lambda\varphi - i\gamma \ln \rho} \sqrt{G(\rho e^{i\varphi})} d\varphi d \ln \rho \right|^2; \quad k = \rho e^{i\varphi}.$$

Поява квадратних коренів у підінтегральних виразах зумовлена переходом від ймовірностей до амплітуд ймовірностей. Відповідно, функції f , G , та Ω описують щільності ймовірності випромінювання нервового імпульсу.

3. Конструювання. Тепер від неперервного подання перейдемо до дискретного. Комплекснозначна штучна нейронна мережа, яка наближено здійснює потрібне двоетапне перетворення, складається з трьох шарів, апроксимуючих функції f , G , та Ω відповідно. Перший шар є вхідним, він апроксимує функцію f та являє собою двовимірну матрицю:

$$f_{js} = f(x_j, y_s);$$

$$x_j = \Delta x \left(j + \frac{1}{2} \right); \quad j \in \{0, \dots, N_x - 1\};$$

$$y_s = \Delta y \left(s + \frac{1}{2} \right); \quad s \in \{0, \dots, N_y - 1\};$$

Аналогічно, другий шар нейронної мережі повинен бути такою матрицею:

$$G_{pq} = G(k_{1p}, k_{2q});$$

$$k_{1p} = \Delta k_1 \left(p + \frac{1}{2} - \frac{N_{k_1}}{2} \right); \quad p \in \{0, \dots, N_{k_1} - 1\};$$

$$k_{2q} = \Delta k_2 \left(q + \frac{1}{2} - \frac{N_{k_2}}{2} \right); \quad q \in \{0, \dots, N_{k_2} - 1\}.$$

Перетворення від першого шару до другого має такий вигляд:

$$G_{pq} = \left| \sum_{j,s} \left(e^{-ik_{1p}x_j - ik_{2q}y_s} \sqrt{f_{js}} \right) \right|^2.$$

Перетворення від другого шару до третього більш складне, оскільки потребує проводити розбиття не за змінними (k_1, k_2) , а за (ρ, φ) . Перший підхід до розв'язання цієї задачі полягає в тому, щоб з самого початку подати другий шар нейронної мережі у вигляді набору концентричних кілець та безпосередньо у нього відобразити перший шар. Проте, з точки зору програмної реалізації, цей підхід виявляється не дуже ефективним, оскільки не дозволяє звести двовимірне перетворення Фур'є до суперпозиції одновимірних. Інший підхід припускає виконання заміни розбиття для вже обчисленої функції G_{pq} :

$$G'_{st} = G(\rho_s, \varphi_{st}) = \sum_{p,q} h_{st}^{pq} G_{pq}.$$

Коефіцієнти h_{st}^{pq} , які фігурують під знаком суми, залежать від обраного способу інтерполяції. Слід також зазначити, що для регулювання щільності розбиття функції G розбиття за φ вже буде функцією від ρ , тобто буде залежати від того, в якому з концентричних кіл розбиття за ρ виконується розбиття за φ . Фінальне перетворення має вигляд:

$$\Omega_{lm} = \Omega(\lambda_l, \gamma_m) = \left| \sum_{s>0,t} e^{-i\lambda_l \varphi_{st} - i\gamma_m \ln \rho_s} \sqrt{G'_{st}} \right|^2.$$

4. Тестування. Для того щоб перевірити принципову придатність отриманих таким чином інваріантних ознак для розв'язання практичних задач розпізнавання образів, до описаної вище нейронної мережі було додано ще один одновимірний шар, який реалізує примітивний класифікатор. Навчання цієї мережі виконувалось шляхом розв'язання системи лінійних рівнянь. Вхідний шар класифікатора збігався з шаром Ω_{lm} , а кількість нейронів у вихідному шарі відповідала кількості класів об'єктів. Квадрат модуля результату підсумовування надходячих імпульсів на кожному нейроні вихідного шару інтерпретувалось як ймовірність віднесення тестового примірника до відповідного класу образів. Усього в тесті використовувалось п'ять класів зображень (рис. 1). Кожному класу відповідав, як

Представник класу					
Додатковий зразок					

Рис. 2. Перелік класів зображень

Текстовий зразок						
Клас						
P	0,992891	0,465047	0,481701	1	0,876046	0,983246
Next P	0,00483629	0,331749	0,241304	1,76028e-09	0,0564782	0,0104506

Текстовий зразок						
Клас						
P	0,972887	0,89134	0,838641	0,979266	0,89681	0,538206
Next P	0,0165332	0,0739882	0,100974	0,012475	0,082991	0,208172

Текстовий зразок						
Клас						
P	0,599839	0,741328	0,631781	0,531349	0,876077	1
Next P	0,288412	0,103204	0,241558	0,328385	0,0680591	9,01094e-09

Рис. 3. Результат класифікації

правило, тільки один навчальний примірник. Результат класифікації тестових примірників наведений на рис. 2. Значення P дорівнює ймовірності правильної класифікації тестового примірника, значення $next P$ — максимальній ймовірності віднесення тестового примірника до будь-якого з решти класів. Результат класифікації виявився малосприйнятливим не тільки до масштабу та розміщення класифікованих об'єктів, але й до кількості, в якій вони наведені на вхідному зображенні. Це досить приємний факт внаслідок того, що використані для побудови ознак перетворення є істотно нелінійними. Слід також очікувати, що

більш досконала архітектура класифікатора забезпечила б більш якісний результат класифікації.

1. *Шевелев И. А.* Волновые процессы в зрительной коре мозга // Природа. – 2001. – № 12. – С. 28–40.
2. *Pitts W., McCulloch W. S.* How we know universals: the perception of auditory and visual forms // Bull. Math. Biophys. – 1947. – 9. – P. 127–147.
3. *Skaggs W. E., McNaughton B. L., Wilson M. A., Barnes C. A.* Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences // Hippocampus. – 1996. – 6. – P. 149–172.
4. *Hubel D. H.* Eye brain and vision. – New York: Freeman, 1988. – 246 p.
5. *Ландау Л. Д., Лифшиц Е. М.* Квантовая механика (нерелятивистская теория). Изд. 6-е, исправл. – Москва: Физматлит, 2004. – 800 с. – (Теоретическая физика, том III).
6. *Akira Hirose.* Complex-valued neural networks, theories and applications. – Berlin: Springer, 2006. – 176 p.

Київський національний університет
ім. Тараса Шевченка

Надійшло до редакції 01.04.2010

Yu. A. Belov, O. V. Tarasiuk

Modeling the brain alpha-waves with artificial complex-valued neural networks

The hypothesis about a role of brain waves in the processes of image recognition has proposed. The main point of the hypothesis is the assumption that the functional role of brain rhythms is to provide the phenomenon of interference of nerve impulses on the soma of a nerve cell. As a result, the same synaptic junctions have either stimulant or inhibitory effect depending on the phase shift. On the grounds of the hypothesis, a special complex-valued neural network has developed and tested for solving the task of building the invariant features of images for the recognition.