УДК 004.032.026

П.В. Тимощук
Національний університет “Львівська політехніка”,
кафедра систем автоматизованого проектування

ОПТИМІЗАЦІЯ НЕЙРОННОЇ СХЕМИ ІДЕНТИФІКАЦІЇ МАКСИМАЛЬНИХ ДИСКРЕТИЗОВАНИХ СИГНАЛІВ

Пропонується оптимізована структурно-функціональна динамічна нейронна схема типу “K-winners-take-all” (KWTA), призначена для ідентифікації К максимальних серед N невідомих дискретизованих сигналів, де $1 \le K < N$. Для N сигналів така схема містить N жорсткообмежувальних нейронів прямого і один - зворотного поширення, який використовується для визначення динамічного зсуву вхідних сигналів. Від існуючих аналогів схема відрізняється простотою схемотехнічної реалізації, високою швидкістю обробки сигналів, здатністю обробляти сигнали з довільного скінченного діапазону, властивістю збереження впорядкованості сигналів, а також відсутністю потреби скидання і необхідної для цього схеми, що додатково підвищує швидкість обробки сигналів. Запропонована нейронна схема може використовуватись для оптимізації виробництва обчислювальних і керуючих систем.

Optimized structure-functional K-winners-take-all (KWTA) dynamic neural circuit of discrete-time that can identify K maximal from N unknown signals, where $1 \le K < N$ is proposed. For N signals, such circuit contains N feedforward hardlimiting neurons and one feedback neuron that is used to determine a dynamic shift of input signals. In contrast to existing analogs the circuit has simplehardware, high speed of signal processing, ability to process signals of finite range, signal ordering preserving property and absence of necessity of resetting and vorresponding circuit that additionally increases a speed of signal processing. Proposed neural circuit can be used for producing optomozation of computing and control systems.

Постановка проблеми. Як відомо, нейронні мережі типу “K-winners-take-all” (KWTA-мережі) здійснюють вибір K серед N елементів, де $1 \le K < N$, з більшими значеннями активаційних функцій, ніж у решти $N - K$ елементів. Коли K дорівнює одиниці, KWTA-мережа є мережею типу “Winner-takes-all” (WTA-мережею), яка може розрізняти нейрон з максимальною активацією [5, 10, 11]. Вибір K найбільших елементів з множини даних N дійсних чисел є ключовою задачею мереж прийняття рішень, розпізнавання образів, пов’язаних пам’ятей і конкуруючого навчання [12, 15]. Задачі такого типу природно зустрічаються при розв’язанні задач класифікації і застосовуються для розроблення класифікаційних нейронних мереж, для розв’язання задач розпізнавання і класифікації зразків [4]. KWTA-мережі застосовуються в телекомунікації, для оптимізації виробництва систем керування пакетними перемикачами даних [1]. KWTA-механізми мають важливі застосування для підвищення якості виробництва обчислювальних систем, при розв’язанні задач класифікації k найближчих об’єктів, кластеризації k значень та ін. [3, 6].

Аналіз останніх досліджень і публікацій. Для розв’язання задач типу “K-winners-take-all” було запропоновано різні види нейронних мереж. Зокрема, KWTA-механізм, який для надійної збіжності К переможців використовує нейронну мережу Хопфілда неперервного часу, пропонується в [7]. Встановлено, що стани такої мережі збігаються до стійкої рівноваги. Мережа використовує модель Хопфілда (адитивну модель Гросберга) і характеризується затримуючими взаємозв’язками, для яких коефіцієнт підсилення сигмоїдної активаційної функції повинен бути достатньо великим. Стани рівноваги мережі є асимптотично стабільними. У встановленому режимі мережа демонструє K компонентів із значеннями $\alpha > 0$ і N-K компонентів з $\beta < 0$. Частковий випадок взаємно затримуючих мереж, призначених для проектування KWTA-мережі за допомогою інтерактивних активацій, описується в [13]. Доведено, що при відповідно вибраних параметрах така KWTA-мережа є дуальною мережі з [7].

KWTA-нейронні мережі обробки дискретизованих сигналів порівняно з аналоговими мережами є більш надійними і демонструють вищу точність обробки сигналів [2]. Так, проста і швидкодіюча KWTA-нейронна мережа обробки дискретизованих сигналів, яка не використовує концепції обопільного затримання для великої кількості вхідних сигналів, запропонована в [14]. Мережа, що має одношарову структуру, визначає динамічне зсування необхідної кількості переможців для досягнення KWTA-режиму. Мережа характеризується обмеженим діапазоном обробки сигналів, потребує скидання початкового динамічного зсуву до центру діапазону зміни вхідних сигналів і вимагає додаткової контролюючої схеми.

Переважна більшість існуючих KWTA-мереж мають обмежену роздільну здатність, відзначаються обчислювальною складністю, обмеженим діапазоном обробки сигналів, для повторного використання такі мережі вимагають прецизійного відновлення їх початкових станів. Відповідні енергетичні функції мереж містять багато локальних мінімумів і не мають глобального мінімуму. Тому вихідні сигнали мереж можуть прямувати до різних встановлених режимів, що перешкоджає їх застосуванню для обробки сигналів у реальному часі [5, 8, 13, 15].

Формування мети дослідження. Метою статті є розроблення структурно-функціональної динамічної нейронної схема типу "K-winners-take-all" обробки дискретизованих сигналів, побудованої у вигляді одношарової конкуруючої архітектури, що виконує динамічне зсування вхідних сигналів. Схема, що складається з N нейронів прямого поширення і одного жорстко обмежуючого нейрона зворотного зв'язку, який використовується для визначення необхідного зсуву вхідних сигналів, повинна мати можливість реалізації у цифровому апаратному забезпеченні на основі суматорів, цифрового інтегратора, перемикачів і зовнішніх джерел напруги або струму, які придатні для обробки сигналів у реальному часі, з використанням ВІС-технології [2]. Необхідно порівняти схему з найбільш відомими сучасними аналогами. Схема повинна мати низьку обчислювальну складність і складність схемотехнічної реалізації, високу швидкість обробки сигналів, здатність до обробки сигналів будь-якого скінченного діапазону і володіти властивістю збереження впорядкування сигналів. Схема не має потребувати скидання і відповідної керуючої схеми, що спрощує апаратне забезпечення і підвищує швидкість обробки сигналів.

Постановка проблеми. Нехай задано N дійсних чисел від a_1 до a_N, $N > 1$, тобто $a_1, a_2, ..., a_N$ як миттєвих значень невідомих вхідних сигналів і необхідно вибрати K найбільших з них, де $1 \leq K < N$ – ненегативне ціле. Припустимо, що задані числа розподілені у відомому діапазоні $a \in (A_{min}, A_{max})$. Покладемо, що ці числа не однакові (відрізняються між собою за значеннями) і впорядковані за спадною послідовністю, що задовольняються нерівності

$$a_1 > a_2 > \cdots > a_N, \tag{1}$$

де індекси $1, 2, \cdots, N$ у загальному випадку можуть відрізнятись від оригінальних номерів входів, означаючи, що компоненти вектора $a = [a_1, \cdots, a_N]$ впорядковані. Побудуємо математичну модель нейронної схеми, яка обробляє вхідний вектор дискретизованих сигналів *a* так, що після скінченної кількості ітерацій отримуються вихідні сигнали схеми $b = [b_1, \cdots, b_N]$, які задовольняють нерівності

$$b_i > 0, i \in 1, 2, \cdots, K; b_j < 0, j \in K+1, K+2, \cdots, N. \tag{2}$$

Нерівності (2) виражають KWTA-властивість, тобто що саме вихідні сигнали від b_1 до b_K "виграють" конкуренцію і той факт, що тільки вони є позитивними компонентами вектора b, свідчить про те, що вхідні сигнали від a_1 до a_K є K найбільшими компонентами вектора a.

Викладення основного матеріалу дослідження. Виконаємо попередню обробку заданого вектора a вхідних сигналів, віднявши від усіх його компонентів значення A_{min} і отримаємо додаткові сигнали

$$c_1 > c_2 > \cdots > c_N, \tag{3}$$

де $c_n = a_n - A_{min}$, $n = 1,2,...,N$. Неважко побачити, що сигнали (3) знаходяться у діапазоні $(0, A)$, де $A = A_{max} - A_{min}$>0, тобто $c \in (0, A)$, де $c = [c_1, c_2,..., c_N]$. Оскільки вхідні сигнали (1) не однакові і розподілені у відомому діапазоні, то сигнали (3) також різні і обмежені в діапазоні $(0, A)$. Отже, для будь-яких $1 \le K < N$ існують такі значення $x \in \Re$, які задовольняють нерівності

$$c_i > x, i \in 1,2,\cdots,K; c_j < x, j \in K+1, K+2,\cdots,N. \tag{4}$$

Віднімання x від (4) дає

$$c_i - x > 0, i \in 1,2,\cdots,K; c_j - x < 0, j \in K+1, K+2,\cdots,N. \tag{5}$$

Як можна побачити з (5), сигнали $c_n - x$, де $n = 1,2,...,N$, володіють KWTA-властивістю. Тому такі сигнали можуть бути використані як вихідні сигнали моделі KWTA-нейронної схеми, тобто можна записати рівності

$$b_i = c_i - x, i \in 1,2,\cdots,K; b_j = c_j - x, j \in K+1, K+2,\cdots,N. \tag{6}$$

Для побудови моделі KWTA-нейронної схеми необхідно розробити процедуру знаходження значення скалярного динамічного зсуву вхідних сигналів x, який задовольняє нерівності (4). Використаємо для цього вимогу, що такий зсув у встановленому режимі повинен знаходитись у діапазоні $(0, A)$. Спроектуємо траєкторію дискретного часу $x^{(k)}$, де $k = 1,2,...,m$ – кількість ітерацій до досягнення встановленого режиму, яка може перетнути весь діапазон $(0, A)$. Нехай така траєкторія буде розв'язком відповідного різницевого рівняння $x^{(k+1)} = \varphi(x^{(k)})$ з початковою умовою $x^{(1)}$, де $\varphi(x^{(k)})$ – певна функція, яка повинна бути визначена. Припустимо, що у деякий момент дискретного часу $t^{(m)}$ змінна $x^{(k)}$ набуває у встановленому режимі значення $x^{(k)} = x^{(m)}$, яке задовольняє нерівність (4). Для зупинки обчислювального процесу у момент $t^{(m)}$ визначимо наступну умову, яка керує кількістю переможців і переможених у кожній дискретній часовій точці протягом обчислювального процесу:

$$R(x^{(k)}) = 2K - N - \sum_{n=1}^{N} \mathrm{sgn}(b_n^{(k)}), \tag{7}$$

де $R(x^{(k)})$ – *k*-те дискретне значення нев'язки, $b_n^{(k)} = c_n - x^{(k)}$ – значення n-го вихідного сигналу моделі на *k*-ітерації

$$\mathrm{sgn}(b_n^{(k)}) = \begin{cases} 1, & \text{if} \quad b_n^{(k)} > 0; \\ 0, & \text{if} \quad b_n^{(k)} = 0; \\ -1, & \text{if} \quad b_n^{(k)} < 0 \end{cases} \tag{8}$$

– сигнум – (жорсткообмежувальна) функція $\sum_{n=1}^{N} \mathrm{sgn}(b_n^{(k)})$ – різниця між реальними кількостями переможців і переможених. Сигнум-функція виконує порівняння між k-дискретним значенням *n*-го вихідного сигналу $b_n^{(k)}$ і нулем. Якщо $b_n^{(k)} > 0$, тоді n-на сигнум-функція забезпечує вихідний сигнал $\mathrm{sgn}(b_n^{(k)}) = 1$, якщо $b_n^{(k)} = 0$, тоді вихідний сигнал n-ї сигнум-функції $\mathrm{sgn}(b_n^{(k)}) = 0$, інакше $\mathrm{sgn}(b_n^{(k)}) = -1$.

Визначатимемо динамічний зсув $x^{(k)}$ за допомогою такого рекурсивного алгоритму:

$$x^{(k+1)} = x^{(k)} - A\Delta x^{(k)}, \tag{9}$$

де $\Delta x^{(k)} = \mathrm{sgn}(R(x^{(k)}))\alpha^k$, α – параметр, який гарантує збіжність алгоритму до KWTA-розв'язку; $0 \le x^{(1)} \le A$ – початкова умова; m – число ітерацій до досягнення збіжності пошуковим процесом встановленого режиму.

Для гарантування того, що модель, яка описується різницевим рівнянням (9) і рівностями (6), буде знаходити K переможців, повинні задовольнятись нерівності $A\alpha^m < c_K - c_{K+1}$. Взявши логарифм від останньої нерівності за основою α для $0.5 < \alpha < 1$, можна отримати наступну нижню границю для кількості ітерацій траєкторій розв'язків рівняння (9) для досягнення збіжності:

$$m > \log_{\alpha} \frac{c_K - c_{K+1}}{A}. \tag{10}$$

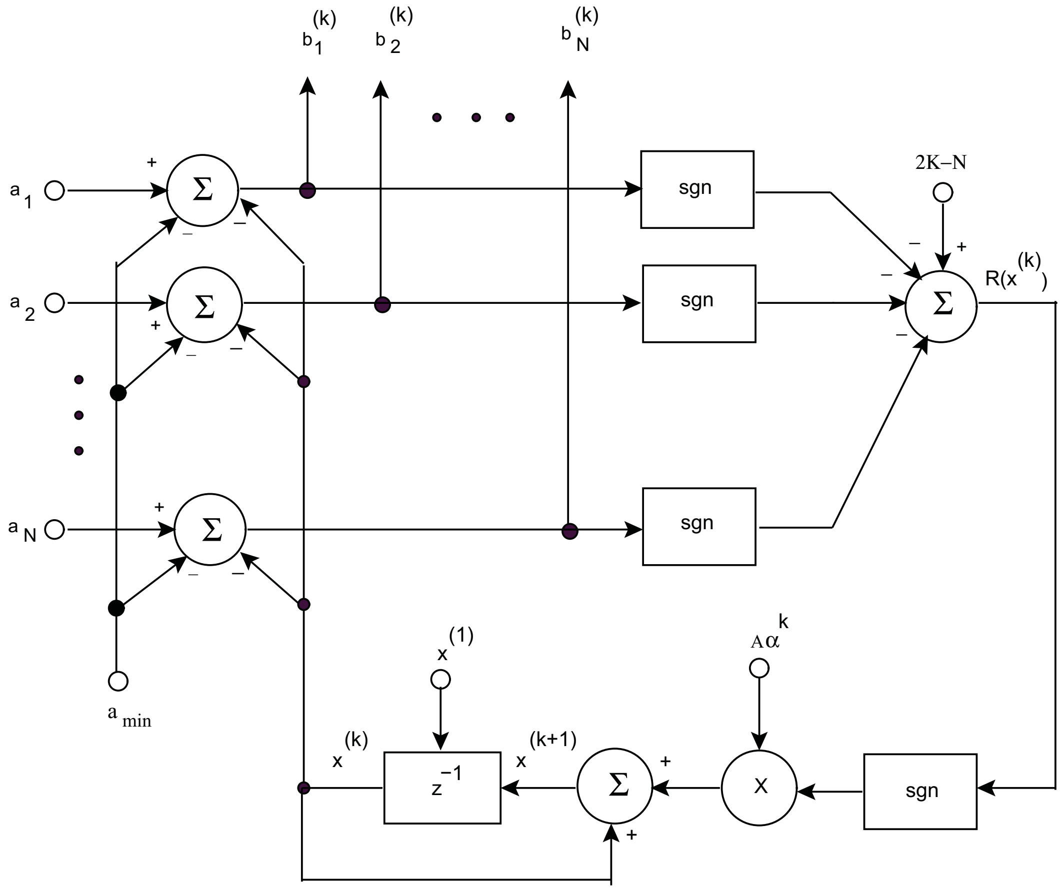
Права частина рівності (10) є скінченною для кожного $c_K - c_{K+1} \ne 0$, тобто для кожного вхідного сигналу, який можна розрізнити. Як випливає з (10), алгоритм, що описується різницевим рівнянням (9) і рівностями (6), має логарифмічну швидкість збіжності.

Структурно-функціональна нейронна схема. Структурно-функціональна динамічна KWTA-нейронна схема обробки дискретизованих сигналів, що описується різницевим рівнянням (9) і рівностями (6), представлена на рисунку, де $x^{(k)} = x((k-1)\tau)$ – динамічний зсув вхідних сигналів у дискретній часовій точці $t^{(k)} = (k-1)\tau$. Схема містить суматори дискретного часу $\sum$, перемножувач, інтегратор z^{-1}, блоки сигнум-функцій sgn (жорсткообмежувачі), зовнішні джерела постійних сигналів $a_{min}, 2K - N, x^{(1)}$ і зовнішнє джерело керованого сигналу $A\alpha^k$. Отже, запропонована динамічна KWTA-нейронна схема обробки дискретизованих сигналів (DTDKWTA-схема) може бути реалізована фізично у сучасному апаратному забезпеченні, використовуючи такі традиційні цифрові компоненти, як суматори, перемножувач, жорсткообмежувальні квантизатори (перемикачі), цифровий інтегратор і зовнішні джерела напруги або струму.

Зазначимо, що вихідні сигнали N блоків sgn також можна використовувати як вихідні сигнали b схеми. Однак, у цьому випадку буде визначатись тільки число переможців K з N вхідних сигналів. Ніякої інформації стосовно впорядкування вхідних сигналів, яку можна використовувати надалі, наприклад, для розв'язання задач сортування, класифікації, кластеризації та ін. не буде.

Як можна побачити, з погляду складності схемотехнічної реалізації така схема для обробки вхідних сигналів, що знаходяться в діапазоні $a \in (0,1)$, повинна містити один перемножувач, $N+2$ суматори, N+1 жорсткообмежувальний квантизатор, один цифровий інтегратор, 2 зовнішніх джерела постійних сигналів і одне зовнішнє джерело змінного сигналу. У частковому випадку алгоритму (9), коли α^k в $\Delta x^{(k)}$ замінити на α, зовнішнє джерело змінних сигналів заміщується на джерело постійних сигналів, що додатково спрощує схему. Для порівняння, одна з найбільш простих і швидкодіючих DTKWTA-мережа з [14] для обробки таких сигналів потребує двох перемножувачів, $N+4$ суматорів, $N+2$ жорсткообмежувальних квантизаторів, одного цифрового інтегратора, 5 зовнішніх джерел постійних сигналів і одного джерела змінних сигналів. Один з найпростіших найшвидших аналогів неперервного часу, описаний в [3], для обробки вказаних сигналів вимагає наявності N підсилювачів, N+1 суматорів, 4N жорсткообмежувальних квантизаторів, одного інтегратора і 2N+2 зовнішніх джерел постійних сигналів (якщо активаційна функція $g_{\Omega_i}(x_i)$ реалізується на чотирьох жорсткообмежувачах). Отже, вищеописана схема має меншу складність схемотехнічної реалізації, ніж інші аналоги.

З погляду обчислювальної складності DTDKWTA-схема у випадку вхідних сигналів, зосереджених в діапазоні $a \in (0,1)$, на кожній ітерації потребує послідовного виконання одного перемноження, $N+5$ сумувань/віднімань, двох логічних жорсткообмежувальних операцій і однієї інтегруючої операції. Якщо в алгоритмі (9) α^k замінити на α, то така схема вимагатиме замість двох лише одного перемноження. Конкуруюча DTKWTA-схема потребує у цьому випадку на кожній ітерації послідовного виконання двох перемножень, $3N+2$ сумувань/віднімань, двох операцій сигнум-функцій і однієї операції інтегрування. Аналог неперервного часу з [3] у цьому випадку вимагає на кожному такті послідовного виконання одного перемноження, N+3 додавань/віднімань, двох логічних операцій (якщо частина активаційної функції $g_{\Omega_i}(x_i)=x_i, l_i \le x_i \le h_i$ реалізується на двох послідовно з'єднаних жорсткообмежувачах) і однієї інтегруючої операції. Оскільки, як буде продемонстровано нижче, при ітеративній реалізації нейронних мереж неперервного часу шляхом заміни диференційних рівнянь на відповідні різницеві рівняння кількість ітерацій, необхідних для збіжності станів таких мереж до KWTA-режиму, є значно більшою, ніж для аналогів, призначених для обробки дискретизованих сигналів, обчислювальний час DTDKWTA-схеми є меншим, ніж в інших аналогів, зокрема DTKWTA-, CHKWTA- [10] і IAKWTA-мереж [13].



Структурно-функціональна схема ітеративної KWTA-нейронної схеми, що описується різницевим рівнянням (9) і рівностями (6).

Роздільна здатність схеми є теоретично необмеженою, тобто якщо вхідні сигнали нерівні, то схема завжди може ідентифікувати їх згідно з KWTA-властивістю (2). Практична роздільна здатність може обмежуватись точністю схемотехнічної реалізації. Оскільки запропонована схема може обробляти будь-які нерівні сигнали, точність функціонування схеми є такою самою, як у інших аналогів.

Функціонування схеми не залежить від початкової умови $x^{(1)}$, яка може набувати довільного значення в діапазоні [0,A]. Тому схема для повторної обробки сигналів не потребує періодичного скидання $x^{(1)}$, додаткової керуючої схеми скидання і додаткового часу обробки на таку операцію [6, 8, 9, 14]. Це дає змогу при виробництві приладів, які використовують такого типу схеми, спрощувати апаратне забезпечення і підвищувати швидкість обробки сигналів, що важливо для їх функціонування у реальному часі.

Важливою перевагою запропонованої KWTA-нейронної схеми є те, що на противагу іншим аналогам [3, 6, 14] вона володіє властивістю збереження впорядкованості вхідних сигналів.

Висновки. Запропоновано оптимізовану структурно-функціональну нейронну схему типу "K-winners-take-all", призначену для обробки дискретизованих сигналів, яка функціонує на основі динамічного зсування вхідних сигналів. Результати порівняння свідчать про те, що схема досягає швидкості обробки сигналів, не нижчої, ніж інші аналоги. Складність схемотехнічної реалізації схеми є дещо меншою, ніж в інших KWTA-мереж. Схема може обробляти будь-які нерівні сигнали із скінченними значеннями і володіє властивістю збереження впорядкування сигналів. Періодичного скидання і необхідної для цього додаткової схеми при обробці повторюваних сигналів не потрібно. Запропоновану схему можна рекомендувати до використання, коли необхідна проста KWTA-схема з високими роздільною здатністю, швидкістю обробки сигналів широкого діапазону, незалежністю від початкових умов і властивістю збереження впорядкованості сигналів. Запропонована схема придатна для різноманітних застосувань, зокрема, для оптимізації виробництва як аналогових, так і цифрових обчислювальних та керуючих систем.

1. Bihn L. N. and Chong , H. C. A neural-network contention controller for packet switching networks, IEEE Trans. on Neural Networks 6 (1995) 1402–1410. 2. Cichocki A. and Unbehauen R. Neural Networks for Optimization and Signal Processing (New York: John Wiley and Sons, 1993). 3. Hu X. and Wang J. An improved dual neural network for solving a class of quadratic programming problems and its k-winners-take-all application, IEEE Trans. on Neural Networks, 19 (2008) 2022-2031. 4. Kwon T.M. and Zervakis M. A parallel sorting network without comparators: A neural-network approach, in: Proc. Int. Joint Conf. on Neural Networks, Vol. 1 (1992) 701–706. 5. Lippmann R. P., Gold B. and Malpass M.L. A comparison of Hamming and Hopfield neural nets for pattern classification, MIT Lincoln Laboratory Technical report TR-769 (1987) 1–37. 6. Liu S. and Wang J. A simplified dual neural network for quadratic programming with its KWTA application, IEEE Trans. on Neural Networks, 17 (2006) 1500–1510. 7. Majani E., Erlanson R. and Abu-Mostafa Y. On the K*-winners-take-all network, in: Advances in Neural Information Process. Syst. D.S. Touretzky, Vol. 1 (Kaufmann, San Mateo, 1989) 634–642. 8. Marinov C.A. and Calvert B. D. Performance analysis for a* K*-winners-take-all analog neural network: basic theory, IEEE Trans. on Neural Networks 14 (2003) 766-780. 9. Marinov C.A. and Hopfield J.J. Stable computational dynamics for a class of circuits with 0(N) interconnections capable of KWTA and rank extractions, IEEE Trans. on Cir. and Syst. I: Fundamental Theory and Applications 52 (2005) 949–959. 10. Tymoshchuk P. and Kaszkurewicz E. A Winner-take-all circuit based on second order Hopfield neural networks as building blocks, in: Proc. Int. Joint Conf. on Neural Networks, Vol. II (2003) 891–896. 11. Tymoshchuk P. and Kaszkurewicz E. A winner-take-all circuit using neural networks as building blocks, Neurocomputing 64 (2005) 375-396. 12. Urahama K. and Nagao T. K-Winner-take-all circuit with 0(n) complexity, IEEE Trans. on Neural Networks 6 (1995) 776–778. 13. Wolfe W.J., Mathis D., Anderson C., Rothman J., Gotler M., Bragy G., Walker R., Duane G. and Alaghband G. K-Winner networks, IEEE Trans. on Neural Networks 2 (1991) 310–315. 14. Yang J. F. and Chen C. M. A Dynamic K-Winners-Take-All Neural Network, IEEE Trans. on Syst., Man and Cyb. 27 (1997) 523~526. 15. Yen J.C., Guo J.I. and Chen H.-C. A new* k*-Winners-take all neural network and its array architecture, IEEE Trans. on Neural Networks 9 (1998) 901–912.*