

KLASYFIKACJA MIKROSTRUKTUR MATERIAŁÓW Z ZASTOSOWANIEM STATYSTYCZNYCH WSKAŹNIKÓW WYŻSZYCH RZĘDÓW ORAZ SZTUCZNEJ SIECI NEURONOWEJ HAMMINGA

ALEXANDR MIKHALYOV¹, DENIS MIKHALYOV²

APPLICATION OF HIGHER ORDER STATISTICAL FACTORS AND HAMMING'S ARTIFICIAL NEURAL NETWORK TO CLASSIFICATION OF MATERIAL MICROSTRUCTURE

Abstract

The method of classification of material microstructure accounting for texture is presented in the paper. The texture is defined as orientation of elements in three dimensions. The features of the microstructure are identified using higher order statistical analysis. The analysis is based on evaluation of semi invariants. Finally, the Hamming artificial neural network was applied for classification of microstructures. Example of application of the developed method to classification of five classes of materials is demonstrated in the paper. Application was performed using MatLab environment and 90% accuracy of classification was achieved.

1. WSTĘP

Pojęcie tekstury coraz częściej stosuje się w różnych dziedzinach nauki i techniki, na przykład w materiałoznawstwie, w badaniach procesów górniczych i hutniczych, przy kontroli jakości gotowej produkcji względnie przy obróbce mikrostruktury. Wynika to z faktu, że powierzchnia większości rzeczywistych obiektów w tych dziedzinach charakteryzuje się orien-

tacją (anizotropią). Dlatego rozwiązanie problemu klasyfikacji tekstur jest aktualnie bardzo ważne, w szczególności w geologii, geodezji oraz przy analizie mikro- i makrostruktury materiałów w metalurgii. Natomiast jednorodność tekstury pozwala na jej opis przez małą liczbę parametrów.

Przez teksturę autorzy rozumieją przestrzenną organizację elementów. W wyróżnionym obszarze układu taką organizację można ujawnić poprzez statystyczną analizę rozkładu szarego lub innego koloru. Obszar

¹ Михалев Александр, профессор кафедры прикладной математики и вычислительной техники Национальной металлургической академии Украины, E-mail: mai@dmoti.dp.ua

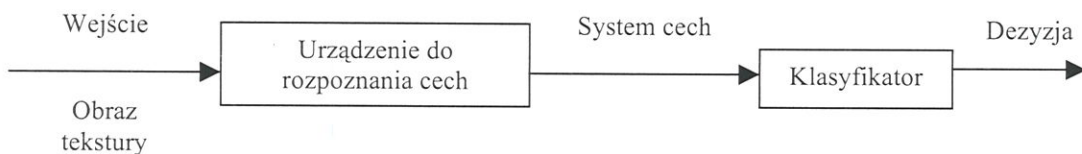
² Михалев Денис, магистр радиофизического факультета Днепропетровского государственного университета

można określić jako steksturowany, jeżeli liczba zauważalnych zmian intensywności kolorów lub kolorów jest dość duża.

Ogólny schemat klasyfikacji tekstur przedstawiono na rysunku 1. Podobnie jak w układach rozpoznawania obrazów, także tutaj można wydzielić dwa główne zadania: rozpoznanie cech oraz klasyfikację tekstur według zaobserwowanych cech.

Pierwsze metody automatycznej analizy tekstury zostały opracowane około dwadzieścia lat temu (Andrzejew i Bazarskij 1984). Podstawowym problemem występującym podczas rozwiązania zadań analizy tekstur w celu ich klasyfikacji jest wyznaczenie systemu cech do opisu przestrzennej tekstury występującej w obrazie. Jako cechę tekstury zazwyczaj przyjmuje się charakterystyczne właściwości, wspólne dla wszystkich tekstur danej klasy.

Podstawą jednej z najszerzej stosowanych metod opisu tekstury jest obliczenie statystycznych wskaźników drugiego rzędu, czyli zestawu macierzy przestrzennych zależności intensywności natężenia światła dwóch elementów obrazu wyznaczonych w zadany sposób. Jednak zmniejszenie ilości informacyjnych oznak powoduje zmniejszenie dokładności klasyfikacji przy zwiększeniu liczby testowych obrazów lub klas (Haralik 1979, Dhawan i Chitre 1996).



Rysunek 1. Ogólny schemat urządzenia do rozpoznania i klasyfikacji tekstur.

Celem niniejszej pracy jest wyznaczenie cech obrazów tekstur mikrostruktury technicznych stopów żelaza z węglem przez zastosowanie analizy statystycznej wyższego rzędu i dalsza klasyfikacja tekstur za pomocą sieci neuronowej Hamminga.

2. ANALIZA STATYSTYCZNA WYŻSZEGO RZĘDU

Do wyznaczenia informacyjnych cech tekstury można zastosować analizę statystyczną wyższego rzędu (higher-order statistics), natomiast jako cechy można wykorzystać ocenę półniezmienników (cumulant, semi invariant) 3-ego rzędu (Mendel i Nikias 1993, Kendall i Steward 1973, Stark 1988).

Półniezmiennikiem (kumulantem) jest jedna z liczbowych charakterystyk wielkości losowej, która jest bliska pojęciu momentu wyższego rzędu. Mieszane półniezmienniki rzędu $r = k_1 + k_2 + \dots + k_n$ dla zadane-

go zbioru n rzeczywistych wielkości losowych $\{x_1, x_2, \dots, x_n\}$ są określone następującym wzorem:

$$C_{k_1 \dots k_n} = (-j)^r \frac{\partial^r \ln \Phi(w_1, w_2, \dots, w_n)}{\partial w^{k_1} w^{k_2} \dots w^{k_n}} \Bigg|_{w_1=w_2=\dots=w_n=0} \quad (1)$$

gdzie:

$\Phi(w_1, w_2, \dots, w_n) = E\{\exp j(w_1 x_1 + \dots + w_n x_n)\}$ – wspólna funkcją charakterystyczna półniezmienników. Zauważmy, że mieszane momenty rzędu r dla tegoż zbioru wielkości losowych opisane są wzorem:

$$m_{k_1 \dots k_n} = E\{x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}\} = (-j)^r \frac{\partial^r \Phi(w_1, w_2, \dots, w_n)}{\partial w^{k_1} w^{k_2} \dots w^{k_n}} \Bigg|_{w_1=w_2=\dots=w_n=0} \quad (2)$$

Tak więc, mieszane półniezmienniki można określić poprzez mieszane momenty rozpatrywanych wielkości losowych. Gdy $m_{1 \dots 0} = E\{x_1\} = 0$, wtedy:

$$\begin{aligned} C_{1 \dots 0} &= 0 \\ C_{2 \dots 0} &= m_{2 \dots 0} = E\{x_1^2\} \\ C_{3 \dots 0} &= m_{3 \dots 0} = E\{x_1^3\} \\ C_{4 \dots 0} &= m_{4 \dots 0} - 3C_{2 \dots 0}^2 = E\{x_1^4\} - 3m_{2 \dots 0}^2 \end{aligned} \quad (3)$$

Zakładając, że $\{X(k)\}$, $k = 0, \pm 1, \pm 2, \dots$, jest stacjonarnym procesem losowym z zerową wartością średnią $E\{X(k)\} = 0$, można otrzymać, że szereg momentów tego procesu jest związanych z jego półniezmiennikami w następujący sposób:

1. Półniezmiennik drugiego rzędu $E\{X(k)X(k+\tau_1)\} = m_2(\tau_1) = C_2(\tau_1)$ – autokorelacyjny szereg, występujący jako miara rozrzutu procesu względem jego średniej wartości;
2. Półniezmiennik trzeciego rzędu $E\{X(k)X(k+\tau_1)X(k+\tau_2)\} = m_3(\tau_1, \tau_2) = C_3(\tau_1, \tau_2)$ – szereg momentów lub półniezmienników trzeciego rzędu, który

wyznacza asymetrię rozkładu;

3. Półniezmiennik czwartego rzędu $E\{X(k)X(k+\tau_1)X(k+\tau_2)X(k+\tau_3)\} = m_4(\tau_1, \tau_2, \tau_3) = C_4(\tau_1, \tau_2, \tau_3) + C_2(\tau_1) \cdot C_2(\tau_3 - \tau_2) + C_2(\tau_2) \cdot C_2(\tau_3 - \tau_1) + C_2(\tau_3) \cdot C_2(\tau_3 - \tau_1) -$ szereg momentów czwartego rzędu, który charakteryzuje anizotropię rozkładu. Cechą szczególną półniezmienników, odróżniającą je od innych liczbowych wskaźników charakterystyk funkcji rozkładu, jest niezależność półniezmienników wyższych rzędów od półniezmienników niższych rzędów. To oznacza, że każdy półniezmiennik wyższego rzędu wprowadzi nową informację o właściwościach procesów losowych.

Algorytm wykorzystania teorii półniezmienników dla uzyskania informacji o obrazach tekstury jest następujący (algorytm 1):

1. Obraz tekstury $I(x, y)$ jest przekształcany w masyw jednowymiarowy $y(k)$ wymiaru N poprzez kolejne sczytywanie wierszami, tzn. $N = xy$.
2. Masyw $y(k)$ jest rozbijany na K rozłącznych segmentów o M punktów w każdym, czyli $N = KM$.
3. Dla każdego segmentu odejmowana jest średnia wartość.
4. Dane $\{x^{(i)}(k), k = 0, 1, 2, \dots, M-1\}; i = 1, 2, \dots, K$ w każdym segmencie dają oszacowanie kolejności półniezmienników trzeciego rzędu:

$$C_{3x}^{(i)}(m, n) = \frac{1}{M} \sum_{l=s_1}^{s_2} x^{(i)}(l)x^{(i)}(l+m)x^{(i)}(l+n), \quad (4)$$

gdzie: $i = 1, 2, \dots, K, s_1 = \max(0, -m, -n), s_2 = \min(M-1, M-1-m, M-1-n)$.

5. Obliczana jest średnia wartość $C_{3x}^{(i)}(m, n)$ dla wszystkich segmentów:

$$\hat{C}_{3x}(m, n) = \frac{1}{K} \sum_{i=1}^K C_{3x}^{(i)}(m, n) \quad (5)$$

Wymiar macierzy półniezmienników zależy od zakresu wymiarów n i m i może zmieniać się w zależności od wymiaru pierwotnego obrazu. Dla małych wymiarów obrazu macierz półniezmienników może być większa od samego obrazu, żeby zwiększyć przestrzeń oznak. Natomiast dla dużych wymiarów macierz jest mniejsza, aby przyspieszyć pracę klasyfikatora (zobacz rysunek 1).

3. SIEĆ NEURONOWA HAMMINGA

Spośród różnych konfiguracji sztucznych sieci neuronowych (SSN) wyróżnić można także takie, których na podstawie sposobu uczenia nie można zaliczyć ani do sieci uczonych z nauczycielem ani do sieci uczonych bez nauczyciela (Wasserman 1992). W tych sieciach wagi synaps obliczane są tylko raz przed rozpoczęciem pracy sieci w oparciu o informację o

wejściowych danych i uczenie tych sieci sprowadza się tylko do obliczenia wag. Do sieci o podobnej logice można zaliczyć sieć Hamminga, która zwykle stosowana jest do tworzenia pamięci asocjatywnej, i która została wykorzystana w niniejszej pracy.

Rozpatrzmy zadanie o przynależności obrazu ξ do wybranej klasy X_k , wyznaczonej przez zadane biblioteczne obrazy x_k . Każdy z zadanych obrazów zbioru testowego wyznacza swoją klasę, a więc zadanie sprowadza się do znalezienia „najbliższego” obrazu. W przypadku dwóch binarnych (0-1) obrazów odległość pomiędzy nimi może być wyznaczona według Hamminga jako liczba niezgodnych składników. Dalej, po obliczeniu wszystkich par odległości $\rho_k = \rho(x_k, \xi)$, poszukiwana klasa wyznaczana jest na podstawie najmniejszej wartości spośród nich.

Rozwiązanie tego zadania przez sieć neuronową może być uzyskane w oparciu o architekturę sieci Hamminga. Sieć zawiera jedną warstwę jednakowych neuronów, liczba których jest równa liczbie klas. Każdy neuron jest połączony z każdym wejściem, a liczba wejść jest równa rozmiarowi obrazów bibliotecznych. Wagi połączeń przyjmowane są jako równe normalizowanym obrazom bibliotecznym:

$$W_n^m = \frac{x_n^m}{\sum_i x_i^m} \quad (6)$$

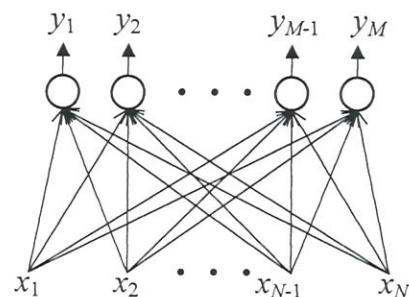
gdzie: W_n^m – wartość wagi połączenia od wejścia n do neuronu m (rysunek 2). Proces wprowadzania informacji o obrazie ξ do sieci neuronowej jest nieiteracyjny, przy czym wejściowy wektor podlega normalizacji:

$$s_n = \frac{\xi_n}{\sum_i \xi_i} \quad (7)$$

oraz neurony przyjmują początkowe poziomy aktywności:

$$y^m(0) = f\left(\frac{1}{2} \sum_n s_n W_n^m - \Theta\right) \quad (8)$$

gdzie: $f(x)$ – funkcja aktywacji neuronu, która jest równa zeru dla $x < 0$, oraz $f(x) = x$ dla $x > 0$. Poziomy Θ przyjmowane są zwykle jako równe zeru.



Rysunek 2. Sieć neuronowa Hamminga.



Sieć powinna wybrać obraz z minimalną odległością Hamminga do nieznanego sygnału wejściowego, w wyniku czego będzie aktywowane tylko jedno wyjście sieci odpowiadające temu obrazowi.

Przy zadaniu wejściowego wektora początkowe wzbudzenie otrzymują wszystkie neurony, dla których iloczyn skalarny wektorów pamięci z wejściowym wektorem przewyższa zadany próg. Dalej spośród tych neuronów należy wybrać jeden, dla którego ten próg jest najwyższy. Osiągane jest to poprzez wdrożenie dodatkowych połączeń (sprzężeń zwrotnych) między neuronami, pracujących na zasadzie „lateralnego hamowania”. Wagi lateralnych połączeń w warstwie neuronowej są normalizowane tak, żeby sygnał sumujący był wzbudzany tylko dla neuronu o największej początkowej aktywności. Pozostałe neurony odczuwają hamowanie:

$$y^m(t+1) = f\left(y^m(t) - \frac{1}{M+1} \sum_{n \neq m} y^n(t)\right) \quad (8)$$

gdzie: M – liczba wejść sieci.

Po wykonaniu kilku iteracji t dla wszystkich neuronów oprócz jednego wartość argumentu funkcji $f(x)$ będzie ujemna, co hamuje aktywność tych neuronów y^m do zera. Jeden pozostały aktywny neuron zostaje zwycięzcą. On wskazuje na klasę, do której należy wprowadzony obraz. Ten sposób postępowania uzyskał nazwę „Zwycięzca Zabiera Wszystko” (Kosko 1992, Terechow 1994).

4. SFORMUŁOWANIE ZAGADNIENIA

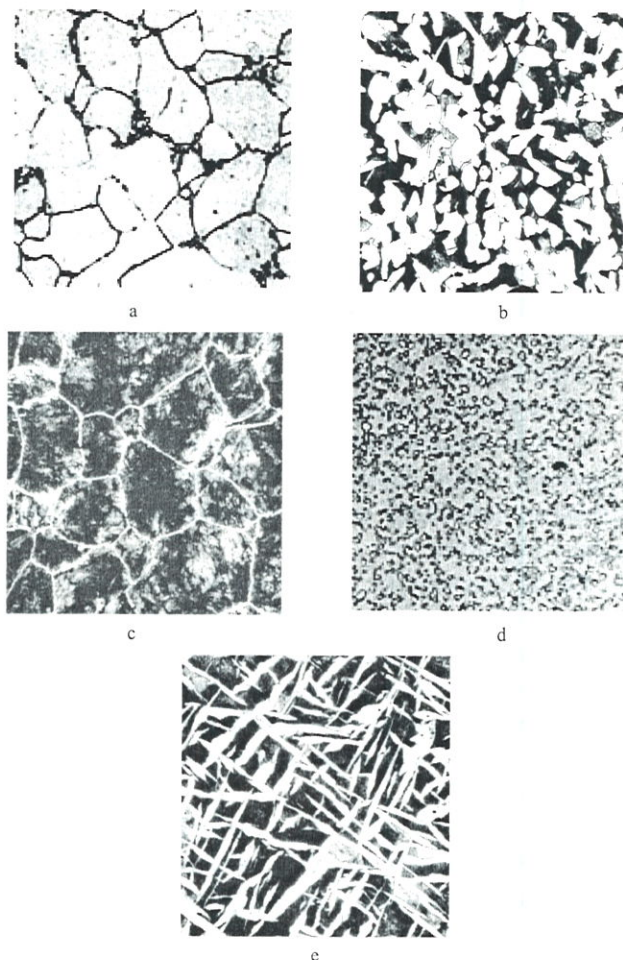
Wybrano 5 testowych obrazów metalograficznych mikrostruktur materiałów. Każdy obraz reprezentuje oddzielną klasę. Rozmiar każdego obrazu wybrano 128×128 .

Klasy mikrostruktur są następujące:

1. Klasa A – stal 08 – ferryt z trzeciorzędowym cementytem na granicach ziaren.
2. Klasa B – stal 30 – ferryt z perlitem.
3. Klasa C – stal nadeutektoidalna \approx N10.
4. Klasa D – stal nadeutektoidalna po sferoidyzacji.
5. Klasa E – staliwo podeutektoidalne.

Obrazy mikrostruktur dla różnych klas pokazane są na rysunku 3. Wybrane klasy pasują do opisanego powyżej sposobu wyznaczenia tekstury, dlatego można do nich zastosować przedstawioną wcześniej metodę nazwaną algorytm 1.

Głównym problemem w klasyfikacji tekstury jest nieoznaczoność orientacji obrazu. Oznacza to, że do wejścia systemu rozpoznania może być podana każda z tekstur przedstawionych na rysunku 3 obrócona o kąt 90° , 180° lub 270° , co ze względu na anizotropię



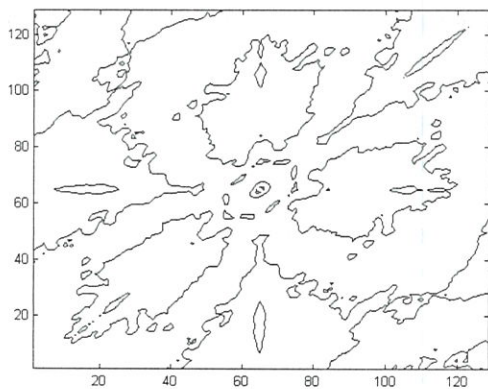
Rysunek 3. Mikrostruktury różnych klas: a – klasa A, b – klasa B, c – klasa C, d – klasa D, e – klasa E.

komplikuje takie rozpoznanie. Wówczas uzyskujemy 20 obrazów oraz 5 klas. Wtedy należy prawidłowo zinterpretować obraz wejściowy i zaliczyć go do jednej z klas.

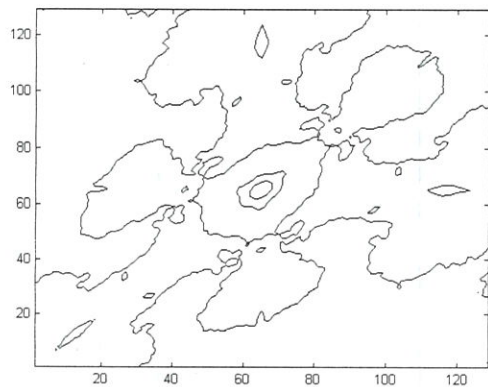
5. WYNIKI I ANALIZA

W fazie początkowej dla uzyskania informacyjnych oznak tekstury zastosowano algorytm 1. Rozmiar segmentów, na które rozbijany jest wejściowy maszyn, wybrano równym $m = 128$. Rozmiar macierzy półniezmienników 3-ego rzędu jest równy $2n + 1 = 129$, to znaczy, że $n = 64$. Przykład takiej macierzy dla klasy mikrostruktur C w reprezentacji graficznej przedstawiono na rysunku 4.

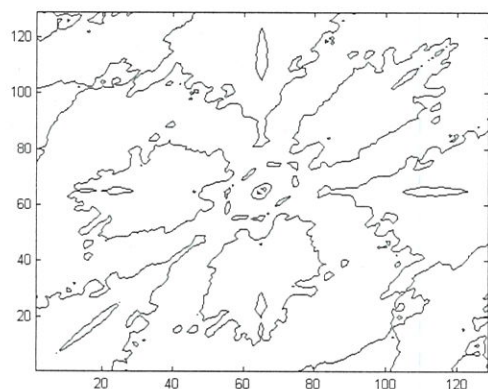
Jak zaznaczono powyżej, wejściowe obrazy dla sieci Hamminga są binarne. Dlatego przed klasyfikacją należy przekształcić obrazy mikrostruktur do postaci binarnej. Poziom progu przekątnika wybrano równy 0. Zatem, jeżeli $C(i, j) \geq 0$, to wtedy $B(i, j) = 1$, natomiast, jeżeli $C(i, j) < 0$, to wtedy $B(i, j) = 0$, gdzie $i, j = 1, \dots, 2n + 1$, \mathbf{B} – macierz binarna cech informacyjnych tekstury. Przykład macierzy \mathbf{B} w reprezentacji graficznej dla klasy C pokazano na rysunku 5.



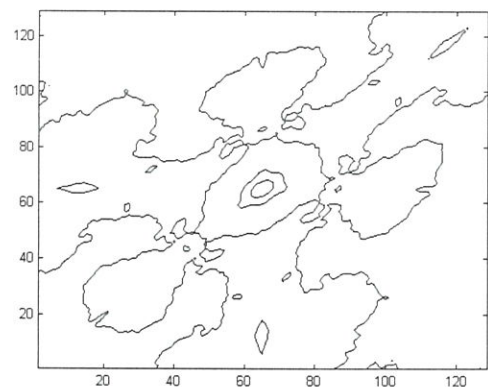
a



b



c

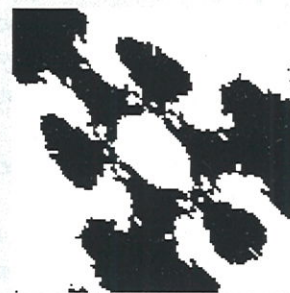


d

Rysunek 4. Pólniezienniki 3-ego rzędu dla tekstury przedstawionej na rysunku 3c w postaci linii poziomych dla trójwymiarowej powierzchni: a – orientacja 0°; b – orientacja 90°; c – orientacja 180°; d – orientacja 270°.



a



b



c



d

Rysunek 5. Pólniezienniki 3-ego rzędu po binaryzacji: a – orientacja 0°; b – orientacja 90°; c – orientacja 180°; d – orientacja 270°.

Na rysunku 5 widać, że macierze cech dla różnych orientacji różnią się znacząco. Związano to najpierw z faktem, że dla uzyskania ocen pólniezienników wejściowy obraz przechodził w jednowymiarowy



masyw, a ponieważ przy obrocie obrazu zmienia się układ odniesienia, to odpowiednio zmienia się również macierz oznak informacyjnych.

Analiza obrazu na rysunku 5 wykazuje kilka regularności:

- 1) Macierz cech informacyjnych jest symetryczna względem głównej przekątnej;
- 2) Dla tekstur, dla których kąt nachylenia jest różny od 180° , macierz cech informacyjnych jest symetryczna.

Na podstawie tych obserwacji udowodniono empirycznie, że najlepszymi wektorami-obrazami dla sieci Hamminga są uśrednione dla wszystkich czterech kierunków późniemienniki 3-ego rzędu.

Jako wynik przedstawionej analizy powstał algorytm pracy systemu wyznaczania obrazów tekstury (algorytm 2) w następującej postaci:

1. Dla każdego z 40 obrazów obliczane są późniemienniki 3-ego rzędu.
2. Dla każdej z 5 klas obliczony jest uśredniony dla czterech obrazów danej klasy późniemiennik 3-ego rzędu, który przedstawiany jest binarnie z progiem 0.
3. Wszystkie 5 binarnych macierzy wprowadzone jest do sieci Hamminga jako biblioteczne wektory-obrazy.
4. Na wejście systemu wyznaczania podawany jest niewiadomy testowy obraz (jeden z dwudziestu wymienionych powyżej), który system powinien zaliczyć do jednej z klas z minimalnym błędem.

5. PODSUMOWANIE

Opisany system klasyfikacji obrazów tekstury programowo został zrealizowany w środowisku MatLab v. 5.2. Dokładność klasyfikacji równa się 90%. Czas obliczeń na PC Pentium 366 MH/ RAM 64MB to 10 minut na obliczenie macierzy cech oraz 2–3 sekundy na klasyfikację obrazu. Duża prędkość klasyfikacji

związana jest z prostotą pracy sieci neuronowej Hamminga. Dla przyspieszenia pracy urządzenia wyznaczenia cech (rysunek 1) próbowano zmniejszyć macierz cech do wymiaru 64×64 , jednak wraz ze zmniejszeniem czasu obliczeń późniemienników dokładność spadła do około 70%.

Do zalet opisanej metody należy zaliczyć wysoką dokładność klasyfikacji oraz krótki czas obliczeń klasyfikatora dla danej klasy PC. Jako wadę można uznać długi czas obliczeń przy wyznaczeniu cech oraz ograniczoną liczbę opracowanych klas tekstury. W przypadku wykorzystania dodatkowej 6-ej klasy, system zmuszony jest zaliczyć ją do jednej z pięciu istniejących. Ten problem związany jest z zasadą działania sieci Hamminga i może być rozwiązany tylko poprzez zasadnicze zmiany ideologii pracy klasyfikatora. Jako przykład, zamiast sieci Hamminga można zastosować paradygmat ART (adaptive resonance theory), z którym będą związane dalsze badania.

LITERATURA

- Andrzejew, G. A., Bazarskij, O. W., 1984, Analiz i sintez sluczajnych prostranstwiennych tekstur, *Zarubieżnaja radioelektronika*, No. 2, 3-34.
- Dhawan, A., Chitre, Y., 1996, Analysis of mammographic microcalcifications using gray-level image structure features, *Proc IEEE Trans. of Medical Imaging*, 15, No. 3.
- Haralik, R. M., 1979, Statistical and Structural Approaches to Textures, *Proc IEEE*, 67, 786-804.
- Kendall M. J., Steward A., 1973, *Statisticheskie wywody i swiazi*, Nauka, Moskwa.
- Kosko, B., 1992, *Neural networks for signal processing*, Prentice Hall, Englewood Cliffs.
- Mendel, J.M., Nikias, C.L., 1993, Signal Processing with Higher-Order Spectra, *IEEE Signal Processing Magazine*, 10, No. 3.
- Stark G., 1988, *Primienienie metodow Fourier-optiki*, Radio i swiaz, Moskwa.
- Terechow S.A., 1994, *Lekcji po teorii i priloženiam iskusstwennych nejronnych setej*, WNIIT, Snežynsk.
- Wasserman F., 1992, *Nejrokomputiernaja tehnika*, Mir, Moskwa.