

DOI: 10.4467/29567610PIB.24.029.20791

prof. zw. dr hab. inż. Tadeusz Niedziela

Wyższa Szkoła Kształcenia Zawodowego we Wrocławiu

ORCID 0009-0000-2625-6520

prof. zw. dr hab. dr h. c. mult. Brunon Hołyst

Menedżerska Akademia Nauk Stosowanych

Uniwersytet w Siedlcach

Nasarawa State University Keffi (Nigeria)

ORCID 0000-0001-8673-1485

WYKORZYSTANIE FILTRÓW GABORA DO IDENTYFIKACJI OBRAZÓW PISMA ODRĘCZNEGO

USING GABOR'S FILTERS TO IDENTIFY HANDWRITTEN IMAGES

Streszczenie

Celem artykułu jest ocena efektywności zastosowania filtrów Gabora w identyfikacji oraz klasyfikacji obrazów przedstawiających pismo odręczne, w szczególności w kontekście rozpoznawania liter, cyfr lub całych słów. W artykule przedstawiono optoelektroniczną metodę identyfikacji cech charakterystycznych obrazów pisma odręcznego bazującą na transformacie falkowej Gabora. Metoda ta nie była nigdzie dotychczas publikowana a otwiera nowe perspektywy badania pisma ręcznego. Problem badawczy sformułowano: Czy zastosowanie filtrów Gabora pozwala na efektywną ekstrakcję cech z obrazów pisma odręcznego, umożliwiającą dokładną identyfikację i klasyfikację znaków w porównaniu do innych metod ekstrakcji cech? Odpowiednio do postawionego problemu badawczego sformułowano hipotezę badawczą, która zakłada, iż filtry Gabora, dzięki swojej zdolności do wykrywania wzorców lokalnych w obrazach (takich jak krawędzie, tekstury i struktury kierunkowe), są skuteczną metodą ekstrakcji cech w zadaniach związanych z identyfikacją obrazów pisma odręcznego, przewyższającą inne powszechnie stosowane metody pod względem dokładności klasyfikacji. Autorzy w artykule zaproponowali prosty algorytm rozpoznawania obrazów w przestrzeni dyfrakcyjnej dla celów kryminalistyki. Pozostaje pytanie, czy metoda ta może znaleźć zastosowanie w praktyce? W celu uzyskania odpowiedzi na problem badawczy i weryfikacji hipotezy badawczej zastosowane metody badawcze takie

jak analiza literatury zarówno krajowej jak zagranicznej, analiza porównawcza, jak również modelowanie matematyczne.

Słowa kluczowe: filtry Gabora, kryminalistyka, pismo odręczne, system optoelektroniczny, filtr multifalkowy

Summary

The aim of the article is to assess the effectiveness of the use of Gabor filters in the identification and classification of images depicting handwriting, in particular in the context of recognition of letters, numbers or whole words. The article presents an optoelectronic method of identifying the characteristics of handwriting images based on the Gabor wavelet transform. This method has not been published anywhere before, but it opens up new perspectives for the study of handwriting. The research problem was formulated: Does the use of Gabor filters allow for effective feature extraction from handwriting images, enabling accurate identification and classification of characters compared to other feature extraction methods? Accordingly to the research problem posed, a research hypothesis was formulated, which assumes that Gabor filters, thanks to their ability to detect local patterns in images (such as edges, textures and directional structures), are an effective method of extracting features in tasks related to the identification of handwriting images, surpassing other commonly used methods in terms of classification accuracy. In the article, the authors proposed a simple algorithm for recognizing images in the diffraction space for forensic purposes. The question remains, can this method be used in practice? In order to obtain an answer to the research problem and verify the research hypothesis, research methods such as analysis of both domestic and foreign literature, comparative analysis, as well as mathematical modeling were used.

Keywords: Gabor filters, forensics, handwriting, optoelectronic system, multi-wavelet filter

Wstęp

Współczesne rozpoznawanie obrazów bazuje na ekstrakcji ich cech charakterystycznych. Ekstrakcja cech charakterystycznych obrazów może być łatwo zrealizowana z tekstury obrazów. Tekstura jako jedna z podstawowych cech powierzchni, odgrywa ważną rolę w rozpoznawaniu obrazów. Zazwyczaj różnice w skali szarości lub w kolorze pomiędzy sąsiednimi pikselami są niewystarczające. Tekstura, zawiera przestrzenny rozkład oraz informacje statystyczne o skali szarości^{1,2,3}. Tekstura (struktura) to uprzywilejowana

- 1 U. Marmol, G. Lenda: *Filtry teksturalne w procesie automatycznej klasyfikacji obiektów*, Archiwum Fotogrametrii, Kartografii i Teledetekcji, 21, 2010, s. 235-243.
- 2 T. Andrysiak, M. Chora: *Image retrieval based on hierarchical Gabor filters*, Int. J. Appl. Math. Comput. Sci., 15, 4, 2005 pp. 471-480.
- 3 S. Arivazhagan, L. Ganesan, *Texture segmentation using wavelet transform*, Pattern Recognition Letters 24, 2003, pp. 3197-3203.

orientacja przestrzenna budowy wewnętrznej obrazu. Cechy tekstury odgrywają ważną rolę w automatycznym rozpoznaniu charakterystycznej informacji obrazowej^{4, 5, 6, 7}. Ponieważ różne tekstury oznaczają różne klasy informacji, zatem cechy tekstury mogą być z powodzeniem zastosowane do klasyfikacji obrazów pisma odręcznego.

Ekstrakcja cech charakterystycznych i integracja są podstawowymi dwoma krokami większości algorytmów segmentacji tekstur i rozpoznawania obrazów. Pierwszym krokiem jest dekompozycja (rozłożenie) oryginalnego obrazu na przefiltrowane obrazy, który każdy zawiera charakterystyczne (szczególne) zmiany częstotliwości i położenia intensywności. Dekompozycja obrazu będzie kompletna jeżeli liczba kanałów będzie wystarczająca. Różnice pomiędzy przestrzenną częstotliwością i położeniem są wykrywane przez zmiany w wartościach szarości w filtrowanych obrazach.

W związku z powyższym analiza tekstury jest ważnym problemem naukowym i praktycznym ze względu na segmentację, detekcje i klasyfikacje obrazów. Tekstura jako obszar lokalny jest generalnie trudna do zdefiniowania w ścisły sposób. Obszar badań strukturalnych skupia się głównie na filtracji wielokanałowej z wykorzystaniem miary w postaci energii tekstualnej, która reprezentuje ilość zmian w analizowanym obszarze.

Idea filtracji wielokanałowej polega na dekompozycji obrazu, na określoną ilość przefiltrowanych obrazów z ustaloną zmianą intensywności w wąskim zakresie częstotliwości i orientacji⁸. Konstruowany jest zestaw filtrów (bank filtrów) o ściśle określonych częstotliwościach i orientacjach, który jest wykorzystywany do analizy tekstur obrazów.

Podejście filtracyjne do analizy obrazu jest podejściem naturalnym, ponieważ dziedzina częstotliwościowa dla różnych tekstur jest różna⁹. Istotną korzyścią filtracji wielokanałowej jest fakt, że można stosować proste statystyki dla filtrowanego obrazu jako miary tekstury oraz z faktu, że przefiltrowany obraz niesie ograniczoną ilość informacji częstotliwościowej.

Obrazy po filtracji filtrem Gabora są najczęściej wykorzystywane do określania wektora cech charakterystycznych dla każdego piksela obrazu, wcześniej jednak z odpowiednim wzmacnianiem informacją tekstualną.

4 T. Niedziela, Kopania Ł.: *Metoda falkowa klasyfikacji obiektów trójwymiarowych z dwuwymiarowych obrazów*, Informatyka XXI wieku - wyzwania i dylematy, Monografia 252, 2020, s. 57-75.

5 J. F. Canny: *A computational approach to edge detection*. IEEE Trans. Pattern Anal. Machine Intell. 8, 6, 1986, pp. 679-697.

6 T. Niedziela: *The method of automatic identification of motor vehicle users*, Journal of Civil Engineering and Transport transEngin, 2, 2, 2020, pp. 63-71.

7 T. Niedziela: *Wavelet method of edge detection in images with high-noise level*, Journal of Civil Engineering and Transport transEngin, 3, 2, 2020, pp. 119-130.

8 D. A. Clausi, M. E. Jernigan: *Designing Gabor filters for optimal texture separability*, Pattern Recognition 33, 2000, pp. 1835-1849.

9 A. K. Jain, F. Farrokhnia: *Unsupervised texture segmentation using Gabor filters*, Pattern Recognition, 16, 1991, pp. 1167-1186.

Filtry i falki Gabora są ważne ze względu na ich rosnącą rolę w wielu zastosowaniach widzenia komputerowego, a także w modelowaniu biologicznym np. wzroku, ponieważ ostatnie dowody neurofizjologiczne z kory wzrokowej mózgow ssaków sugerują, że reakcje filtra w postaci neuronów korowych (zwanymi prostymi komórkami) najlepiej jest modelować jako rodzinę samopodobnych 2D falek Gabora¹⁰.

Ponieważ falki są z natury wielorozdzielcze, stąd mogą być z powodzeniem wykorzystane w modelach analizy tekstur^{11, 12, 13, 14, 15, 16, 17}. Funkcja Gabora może być zaimplementowana jako filtr wielokanałowy, przypominający falekę. Filtr ten posiada właściwości, które czynią go atrakcyjnym nie tylko dla widzenia komputerowego. Te właściwości obejmują zarówno jego atrakcyjność, prostotę, optymalną lokalizację oraz połączenie przestrzenno-częstotliwościowe¹⁸.

Perry i współpracownicy¹⁹ wykorzystali stały zestaw filtrów Gabora np. do segmentacji tekstur. Tumer²⁰ wykorzystał zestaw filtrów Gabora do ekstrakcji tekstur, natomiast Bovik ze współpracownikami²¹ użyli stałego zestawu filtrów Gabora, gdzie rzeczywistą część każdego filtra stanowi parzysto-symetryczny filtr Gabora (tj. $4 = 0$), a część urojona jest nieparzysto-symetrycznym filtrem Gabora (tj. $4 = n/2$). Autorzy zastosowali prosty algorytm znajdowania pików widma obrazu w celu wyznaczenia częstotliwości promieniowych filtrów Gabora. Janin^{22, 23} zaproponował, modelowanie kanału za pomocą zestawu filtrów Gabora, które zachowują prawie wszystkie informacje w obrazie wejściowym.

Ważną właściwością filtrów Gabora jest to, że mogą mieć optymalną lokalizację, lub rozdzielczość, zarówno w przestrzeni jak i w domenie przestrzenno-częstotliwościowej.

-
- 10 T. S. Lee: *Image representation using 2D Gabor wavelets*, IEEE Transactions on pattern analysis and machine intelligence, **18**, 10, October 1996.
 - 11 S. Mallat: *Multifrequency channel decompositions of images and wavelet models*, IEEE Trans. Acoustics, Speech, Signal Processing **37**, 12, 1989, pp. 2091–2110.
 - 12 S. Mallat: *A theory for multiresolution signal decomposition: The wavelet representation*. IEEE Trans. Pattern Anal. Machine Intell. **11**, 12, 1989, pp. 674–693.
 - 13 S. Mallat: *Multiresolution approximation and wavelet orthonormal bases of $L(R^2)$* , Trans. Amer. Math. Soc. **3**, 15, 1989, pp. 69–87.
 - 14 S. Mallat, S. Zhong: *Characterization of signals from multi-scale edges*. IEEE Trans. Pattern Anal Machine Intell. **14**, 7, 1992, pp. 710–732.
 - 15 S. Mallat and W. L Hwang., *Singularity detection and processing with wavelets*, IEEE Trans. Inf. Theory **38**, 1992, pp. 617–641.
 - 16 T. Niedziela, Ł. Kopania: *Zmodyfikowany korelator znormalizowanej transformaty Fouriera*, Informatyka XXI wieku – wyzwania i dylematy, Monografia **252**, 2020, s. 42–56.
 - 17 T. Niedziela, Ł. Kopania: *Metoda falkowa klasyfikacji obiektów trójwymiarowych z dwuwymiarowych obrazów*, Informatyka XXI wieku – wyzwania i dylematy, Monografia **252**, 2020, s. 57–75.
 - 18 J. G. Daugman: *Uncertainty relation for resolution In space, spatial frequency, and orientation optimized by two-dimensional visual cortex flters*, J. Opt. Soc. Am. A **2**, 7, pp. 1160–1169, 1985.
 - 19 A. Perry and D. G. Lowe: *Segmentation of Textured Images*, Proc. CVPR, 1989, pp. 326–332.
 - 20 S. Mallat: *Multiresolution approximation and wavelet orthonormal bases of $L(R^2)$* , Trans. Amer. Math. Soc. **3**, 15, 1989, pp. 69–87.
 - 21 A. C. Bovik, M. Clark and W. S. Geisler: *Multichannel Texture Analysis Using Localized Spatial Filters*, IEEE Trans. PAMI, **12**, 1, 1990, pp. 55–73.
 - 22 A. K. Jain, Farrokhnia F, *Unsupervised Texture Segmentation Using Gabor Filters*.
 - 23 A. K Jain, A. Vailaya: *Image retrieval using color and shape*, Pattern Recognition **29**, 8, 1996, pp. 1233–1244.

Filtr Gabora to liniowy filtr służący do analizy tekstury funkcji obrazowej tzn. analizy wybranych częstotliwości w określonych kierunkach, zlokalizowanych w ustalonych obszarach. Odpowiedź impulsowa filtra Gabora jest określana przez fale harmoniczną modulowaną funkcją Gaussa. Transformatą Fouriera odpowiedzi impulsowej filtra Gabora jest splotem transformaty funkcji harmonicznego (funkcji sinusoidalnej) i transformaty Fouriera funkcji Gaussa.

Filtry Gabora są bezpośrednio powiązane z falkami Gabora, ponieważ mogą być realizowane w praktyce dla wielu skal i obrotów, Zwykle projektowany jest bank filtrów składających się z filtrów Gabora o różnych skalach i obrotach. Bank filtrów Gabora o różnych częstotliwościach i orientacjach był pomocny w ekstrakcji cech charakterystycznych funkcji obrazowej²⁴.

Filtry Gabora są wykorzystywane w problemach związanych z analizą obrazów w procesie segmentacji obiektów, wykrywania krawędzi, ekstrakcji cech charakterystycznych, w procesie automatycznego rozpoznawania obiektów (klasyfikacji, czy identyfikacji np. twarzy lub linii papilarnych).

Zmieniając parametry orientacji i częstotliwości filtra Gabora uzyskujemy możliwość identyfikacji obiektów, posiadających określone cechy tekstualne.

1. Cechy charakterystyczne w identyfikacji pisma odręcznego

Identyfikację w sensie prawnym można uznać za definitywne rozpoznanie: osoby, pisma osoby, obiektu lub czynności na podstawie wcześniejszej wiedzy. W identyfikacji pisma odręcznego ekspert działa pośrednio jako świadek tożsamości osoby. W swojej procedurze pracy czynniki behawioralne (czynniki powodujące postępowanie człowieka w oderwaniu od aktów świadomości) wpływają zarówno na identyfikatora, jak i na jego sposób identyfikacji pisma oraz na samo pismo autora.

W identyfikacji pisma odręcznego osobowość eksperta ma duże znaczenie, zatem konieczne jest prześledzenie jego roli jako świadka tożsamości osoby. Jakie są jego środki identyfikacji na różnych etapach, gdy postrzega, zapamiętuje, porównuje, eliminuje lub rozpoznaje, a ostatecznie kończy z moralną pewnością o tożsamości autorstwa.

Ogólnie rzecz biorąc, istnieją dwie różne metody identyfikacji. Jedną z nich jest analityczna, „obiektywna” metoda porównywania, mierzenia lub analizowania cech fizycznych według ich kształtu, liczby, lokalizacji lub kombinacji, jaką jest przykładowo metoda stosowana w identyfikacji odcisków palców i w analizie chemicznej. Drugą jest metoda nieanalityczna np. metoda identyfikacji twarzy, znanych obiektów, głosów, melodii i innych wzorców poprzez proces szybkiego porównania z wcześniejszymi wrażeniami, zachowanymi jako ślady pamięci.

24 R. Heintz, M. G. Schäfer: *Local invariant object localization based on Gabor feature space*.

Te złożone ślady pamięciowe zapewniają subiektywną pewność określonej wiedzy, zwykle nazywaną znajomością. Połączenie metod obiektywnych i subiektywnych jest również stosowane, przykładowo w diagnostyce medycznej, gdzie techniki laboratoryjne oraz objawy kliniczne lub wrażenia prowadzą do identyfikacji choroby.

Obiektywna metoda identyfikacji opiera się na dwóch czynnikach: na niezmiennym charakterze obserwowanych, klasyfikowalnych właściwości fizycznych oraz na rozumowaniu o prawdopodobieństwach, rodzaju kombinacji i wzorcach identycznych znaków, występujących przypadkowo w dwóch różnych obiektach. W tej metodzie uprzednia wiedza wymagana w dowolnej procedurze identyfikacji jest dostarczana przez standardową procedurę dostępną dla wszystkich, a dowody na tożsamość można zademonstrować każdemu inteligentnemu obserwatorowi, poprzez kwantyfikację lub bezpośrednią inspekcję. Nie tak jest w subiektywnej, nieanalitycznej metodzie identyfikacji, gdzie uprzedniej wiedzy nie można przekazać innym, ponieważ jest ona oparta na pamięci. Proces identyfikacji odbywa się w ośrodkowym układzie nerwowym osoby, która ją zidentyfikowała i która może co najwyżej podać powód swojej znajomości z daną osobą i wskazać szczegóły w jej wyglądzie.

Dla celów kryminalistycznych ta druga forma subiektywnej identyfikacji lub rozpoznania ma pierwszorzędne znaczenie, ponieważ jest to nasz naturalny sposób orientacji w otoczeniu i przenika nasze codzienne życie. Od wczesnego ranka, kiedy identyfikujemy godzinę po położeniu wskazówek zegara lub rozróżniamy prawe i lewe kapcie po ich formie i względnym położeniu, aż do wieczora, kiedy spojrzenie w niebo mówi nam coś o pogodzie następnego dnia, nieustannie interpretujemy obrazy, dźwięki i znaki, które identyfikują dla nas to, co uważamy za fakty. Proces rozpoznania może być świadomym procesem podejmowanym celowo lub może nadejść jako nagła decyzja doświadczana jako nagły wgląd, być może werbalizowany jako „to to”. Ale rozpoznanie może również występować po prostu jako subiektywna pewność, nigdy niepodważalna, ale po prostu wchodząca do naszych umysłów jako fakty w momencie percepcji.

Oczywiście cała procedura prawna, mianowicie zadanie ustalania faktów, w dużej mierze zależy od tego rodzaju decyzji identyfikujących lub eliminujących. Policjant na miejscu przestępstwa lub czytający zeznania świadków, starając się zastosować konkretny paragraf kodeksu karnego do szczegółów czynu, stale postrzega i porównuje obrazy i znaki lub znaczenie słów, aż obie strony przybiorą ostateczny kształt w jego umyśle. Przedmioty, osoby i ich działania są identyfikowane nie tylko przez bezpośrednie rozpoznanie świadków, ale także przez ustny opis lub ślady pozostawione na miejscu przestępstwa. Osoby można zidentyfikować po ich chodzie i innych wzorcach ruchu, a nawet mo-

tywy działania można zidentyfikować na podstawie wnioskowania z jawnego ekspresyjnego zachowania.

Wszyscy członkowie zawodu prawniczego muszą polegać na tym samym typie wrażeń; sędzia może zidentyfikować fałsz po rumieńcu i jąkaniu się świadka, a prawnik może zidentyfikować brak zainteresowania sędziego, który „oczywiście” nie słucha jego argumentów i w rzeczywistości „wygląda, jakby” miał zaraz zasnąć. Podsumowując, w celu identyfikacji faktów w dużym stopniu polegamy na naszych małych znakach, opracowanych przez empatię lub metodzie prób i błędów, znakach tworzących ciekawą mieszankę prawdy, wyobraźni i uprzedzeń.

W identyfikacji pisma odręcznego ekspert jest wzywany do zidentyfikowania osoby poprzez porównanie nagrań ruchu, to znaczy poprzez wnioskowanie. Jego zadanie nie jest wykonywane poprzez porównywanie i obliczanie stałych cech, jak ma to miejsce w przypadku metody obiektywnej. Nie jest to również subiektywna forma bezpośredniej identyfikacji opartej na śladach pamięciowych stałych cech. W pewnym sensie eksperta od pisma ręcznego można porównać do świadka identyfikującego osobę dobrze mu znaną po jej chodzie, postawie lub innych często postrzeganych złożonych wzorcach ruchu, bez widzenia jej twarzy. Ekspert od pisma odręcznego robi właśnie to: porównuje ślady celowego zachowania motorycznego zapisane w konwencjonalnych symbolach z narzędziem do pisania w różnych sytuacjach. Podczas gdy konwencjonalne symbole nakładają pewne ograniczenia na swobodną ekspresję indywidualnych wzorców ruchu, ekspert od pisma odręcznego jest w nieco lepszej sytuacji niż nasz świadek, który może porównywać tylko ulotne wrażenia, ponieważ może badać zapisy tego ruchu w swoim czasie wolnym. Z drugiej strony, badacz kryminalistyczny mógł nigdy nie widzieć człowieka, którego próbuje zidentyfikować, i musi uzyskać „wcześniejszą wiedzę” z tak zwanego standardowego pisma danej osoby, które uważa się za reprezentatywne dla jej zachowania motorycznego.

Tak więc zadanie badacza dokumentów w postaci pisma odręcznego jest w pewnym stopniu podobne do zadania krytyka sztuki, który próbuje ustalić autorstwo obrazu poprzez zapoznanie się z innymi dziełami tego samego malarza. Wszystkie rozważania dotyczące okresu, stylu, manieryzmu, treści, środków wyrazu, celowych zmian lub fałszerstw, tak ważne w ocenie sztuk twórczych, są widoczne w procesie oceniania utworu literackiego. Aby dotrzeć do przyczyn subiektywnej pewności eksperta a także do źródła możliwych błędów w osądzie, należy mieć na uwadze jego proces umysłowy, związany z porównywaniem obrazów wizualnych.

Rozważmy najpierw dziecko, które zdaje test inteligencji: jego zadaniem jest dopasowanie kwadratowego kawałka drewna do jednego z otworów,

o różnym kształcie w desce przed nim. Porusza oczami tam i z powrotem z łatwością odrzucając okrągłe i trójkątne otwory, wahając się nad prostokątem, a następnie z krzykiem dopasowuje kwadrat do właściwego otworu.

Z kolei obserwujemy świadka – powiedzmy ofiarę napadu – szybko przeglądającego dziesiątki zdjęć podejrzanych na posterunku policji, kiedy z niewiarygodną szybkością odrzuca jedno po drugim, czasami się waha, wybiera kilka do dalszego przyglądania się, a następnie nagle wykrzykuje „to ten mężczyzna!”. To tak, jakby przejrzysty, ale bardzo wyraźny obraz tego mężczyzny został szybko przesuwany nad każdym ze zdjęć, pokazując na pierwszy rzut oka oczywiste różnice lub podobieństwa.

Ale teraz – w przeciwieństwie do przypadku dziecka – osoba postronna nie może ocenić słuszności decyzji, ponieważ jeden z porównywanych obiektów nie jest już widoczny, istnieje tylko jako obraz mentalny w mózgu świadka. Aby zrozumieć aktywność umysłową eksperta od dokumentów, siedzącego przy biurku, badającego i porównującego szczegóły w standardowym piśmie z tymi w kwestionowanym dokumencie, należy uwzględnić procesy fizjologiczne zaangażowane w zatrzymywanie i przypominanie obrazów.

Obrazy są magazynowane poprzez „odciskanie” na wybranych częściach mózgu za pomocą reakcji elektrochemicznych²⁵. Te wzorzyste zmiany elektryczne w mózgu to to, co nazywa się zespołami komórek, składającymi się z komórek nerwowych, naładowanych elektrycznie i połączonych ze sobą w zamkniętym obwodzie. Można przypuszczać, że zespoły komórek związane z wrażeniami wzrokowymi mają „kod”, jeśli nie wzorzec, postrzeganego obrazu, ponieważ są one zdolne do odtworzenia tego obrazu zarówno w umyśle, jak i w pracy malarza, na płótnie. Rzeczywistą obecność anatomicznie ustrukturyzowanych „obrazów mózgowych” można wywnioskować z następujących faktów: jeśli zachowane są wzorzyste ślady pamięciowe z ruchów postrzeganych w otoczeniu, to będą one zapraszać do naśladownictwa i mogą być faktycznie przechowywane jako wzorce działań dla przyszłych zachowań, zdolne do wytwarzania wzorzystych form ruchu, nawet jeśli nigdy wcześniej nie były wykonywane. Na przykład: jeśli mężczyźnie pod hipnozą powiedziano, aby był Napoleonem lub udawał kobietę, będzie wykonywał niezwykle złożone zadania na podstawie obrazu mentalnego i na podstawie rozkazu.

Percepcja obrazu i jego przywołanie mogą podlegać wpływowi świadomej uwagi lub, w kategoriach fizjologicznych, ilości neurohormonów wytwarzanych w danej sytuacji. Hormony pomagają wywierać wrażenie na naszym umyśle z różną intensywnością i nieusuwalnością, a hormony zdają się dawać nam

25 V. Kyrki, J-K. Kamarainen, H. Kälviäinen: *Simple Gabor feature space for invariant object recognition*, Pattern Recognition Letters, 25, 3, 2004, pp. 311–318.

zdolność do wyodrębniania i umieszczania na pierwszym planie wybranych obrazów przez samo skupienie podczas gdy inne ślady pamięciowe zanikają w tle.

Niebezpieczny efekt selektywnej koncentracji umysłowej jest dobrze znany doświadczonemu badaczowi kwestionowanych dokumentów. Jeśli badacz po zidentyfikowaniu fragmentu pisma próbuje dla kontroli udowodnić sobie, że się myli, zwracając uwagę na różnice, może odkryć ze zgrozą, że w wielu przypadkach zobaczy teraz silne dowody na rzecz braku tożsamości. Możemy tylko wnioskować, że oczekiwania wpływają na percepcję, że obrazy, na których się skupiamy, szukają własnego podobieństwa.

Wielu badaczy wskazywało na proces selektywnej percepcji w badaniu pisma ręcznego, podczas gdy szkoły psychologii i różne metody testowania projekcyjnego rozwijają i wykorzystują to zjawisko rutynowo. Podsumowując, można powiedzieć, że integracja mentalna obrazów, prowadząca do identyfikacji lub eliminacji jest uwarunkowana czynnikami behawioralnymi zmieniającymi skupienie i poziom uwagi, „nastawienie” badającego w momencie porównywania, wpływając w ten sposób na jego ostateczną decyzję.

Podstawa identyfikacji badającego jest uzyskiwana ze standardowego pisma, które uważa się za reprezentatywne dla zachowania motorycznego jednej osoby, z wykluczeniem wszystkich innych osób. Po porównaniu z kwestionowanym dokumentem, tożsamość autorstwa jest wnioskowana, w przypadku braku spójnych różnic, na podstawie podobieństw, które ze względu na ich kombinację i/lub rzadkość są uważane za unikalne. Brak tożsamości lub eliminacja jest wnioskowana na podstawie znacznych różnic, których nie można wyjaśnić żadnymi rozsądnymi czynnikami okolicznościowymi.

Przeważające wyjaśnienie niepowtarzalności pisma odręcznego danej osoby opiera się na niepowtarzalności osobowości, a mianowicie na indywidualnych nawykach jej ruchów na papierze powodujących znaczne odchylenia od standardów podręcznikowych. Problemy wynikające z tak zmiennego medium identyfikacji są zwykle rozwiązywane na kilka sposobów. Niektórzy zalecają wyrażanie tylko różnych stopni prawdopodobieństwa identyfikacji. Inni, w próbie osiągnięcia obiektywnych metod, po prostu dodają jedną cechę więcej do swojej listy cech charakterystycznych, które mają być porównywane, a mianowicie stopień „naturalnej zmienności”. Znów inni próbowali statystycznie obliczyć wartość dowodową grupowań lub kombinacji izolowanych cech występujących razem. I oczywiście istnieje powszechna zgoda, że wszystkie czynniki w sytuacjach pisania, takie jak materiał pisarski, cel, szybkość, treść itp. muszą być badane i brane pod uwagę.

Podczas gdy koncepcja naturalnej zmienności i informacje o sytuacji pisania pomogą wyjaśnić różnice między standardem a kwestionowanym pismem to można ostatecznie udowodnić tożsamość poprzez obecność tych samych

rodzajów odmienności. Główny problem pozostaje: jak badacz może bezpiecznie odróżnić indywidualne cechy charakterystyczne od cech klasowych wspólnych dla grupy ludzi, takiej jak naród lub grupa społeczna, zawodowa lub kulturowa? Odpowiedź brzmi: tylko poprzez długie doświadczenie z wieloma przykładami z danej grupy. Każdy ekspert próbujący zastosować obiektywne metody, pomiary lub jakąkolwiek inną formę kwantyfikacji, w badaniu pisma ręcznego znajdzie je pomocne w klasyfikacji i eliminacji autorstwa. Jest to jednak niemal truizm, że w celach identyfikacyjnych pismo ręczne nie może być traktowane na zasadach zapożyczonych z identyfikacji odcisku palca i że wszystkie takie schematy już zawiodły dotychczas, gdy ktokolwiek próbował je zastosować..

Stąd też jest rzeczą naturalną, że badacze kwestionowanych dokumentów starają się zwiększyć swoją wiedzę, poszukując nowych aspektów tego problemu. Jednym z takich aspektów jest np. aspekt medyczny. Próbę zbadania problemu pisma ręcznego z medycznego lub punktu widzenia neurofizjologicznego podejmowali psychiatrzy i ustalili pięć elementów, które wpływają na pisanie:

1. Wrodzone formy ruchu.
2. Nabyte wzorce ruchu.
3. Zmieniające się napięcie nerwowo-mięśniowe.
4. Wyobrażenie mentalne pożądaney formy pisma.
5. Czynniki behawioralne w konkretnej sytuacji pisania.

Od dawna wiadomo, że postawa, ruchy, gesty i mimika osoby, krótko mówiąc, cała jej aktywność mięśniowa – w tym pismo odręczne – należą do wzorców jej zachowań. Uważa się je za jeden ze stosunkowo trwałych aspektów osobowości, ponieważ rozwinęły się w ciągu życia człowieka na podstawie dziedziczności, poprzez próby i błędy, w stałym kontakcie ze środowiskiem. Jakości tych wzorców ruchu nie da się zmierzyć ilościowo, mimo że są i przenikają całe zachowanie jednostki.

Wykazano względną spójność tych ustrukturyzowanych wzorców ruchu, które są możliwe do rozpoznania w wielu działaniach jednostki osobowej, w wykonywaniu przez nią małych zadań i w innych ruchach, nawet jeśli są nieco zamaskowane. W piśmie odręcznym te wzorce ruchu modyfikują standardy książkowe pod względem formy, rozmieszczenia w przestrzeni i innych cech, takich jak cieniowanie, rozmiar, stopień połączenia, pochylenie i cały zestaw dobrze znanych cech stosowanych do klasyfikacji i identyfikacji pisma odręcznego. Te indywidualne wzorce ruchu w rzeczywistości powodują odstępstwa od standardów książkowych, które są podstawą pracy eksperta.

Odchylenia pisma odręcznego od standardów książkowych są również wykorzystywane przez grafologów w celu oceny osobowości, ale istnieje

ogromna luka w komunikacji między tymi dwoma dziedzinami badań nad pismem odręcznym. Chociaż istnieje ogólna zgoda co do tego, że indywidualność pisma odręcznego jest w jakiś sposób związana z unikalnym wzorcem osobowości piszącego, ekspert kryminalistyki nadal unika wykorzystywania ustaleń bardziej istotnych dla pisma odręcznego jako medium oceny osobowości.

Nie ma powodu, dla którego pismo odręczne nie powinno być wykorzystywane jak każdy inny test osobowości, ponieważ jest bardzo złożoną odpowiedzią na miniaturową sytuację życiową, która jest również podstawą wszelkich testów psychologicznych. Wydajność, motywacja, komunikacja społeczna, ekspresja i autoprojekcja są zawarte w tej konkretnej formie zachowania motorycznego. Ekspert nie musi akceptować żadnych twierdzeń opartych na naiwnych spekulacjach amatorów grafologii, ale powinien zwrócić uwagę na naukową metodologię opracowaną przez psychiatrów i fizjologów w tej dziedzinie. Na pewno przydatne byłoby, na przykład, rozróżnienie między naturalnymi ruchami ekspresyjnymi a „odgrywaniem ról” w piśmie odręcznym, co można lepiej osiągnąć przy pomocy poważnej grafologii. Względna spójność naturalnych ruchów ekspresyjnych jest mniej zależna od sytuacyjnych lub celowych zmian niż przyjęta poza. Dlatego pewna wiedza na temat zmian, których należy się spodziewać w danej sytuacji, może dostarczyć badaczowi przydatnych informacji na temat „niezawodności” konkretnego okazu pisma odręcznego, i pomóc mu wyjaśnić oczywiste różnice między innymi bardzo podobnymi pismami odręcznymi.

Napięcie mięśniowe, jest jednym z najbardziej spójnych czynników wpływających na pismo odręczne i jednocześnie najważniejszym pojedynczym czynnikiem, determinującym zmiany w piśmie odręcznym. Istnieje wiele dowodów klinicznych i eksperymentalnych, które pokazują, że stałe napięcie mięśniowe nie jest tylko cechą danej osoby, ale że zmieniające się napięcie mięśniowe stanowi część nawykowych mechanizmów obronnych danej osoby. Poziom napięcia wzrasta, gdy dana osoba jest wyczulona na niebezpieczeństwo, ale także, gdy po prostu staje w obliczu problemów z dostosowaniem. Wzrasta w obliczu frustracji, konfliktu i wszystkich innych sytuacji stresowych. Osoba może nie być świadoma swoich wysiłków adaptacyjnych w zmieniających się sytuacjach życiowych, a mimo to jej napięcie mięśniowe wzrośnie. Ruchy fizyczne służą jako ciągłe mechanizmy rozładowywania napięcia i łagodzenia lęku. Badania nad gniewem pokazują, że najczęstszymi impulsami reakcji na gniew są słowna odpowiedź, atak fizyczny, ucieczka lub krzyk.

Wszystkie te działania wymagają przygotowania napięcia mięśniowego, które narastają w odpowiednich ośrodkach motorycznych i powodują drobne skurcze w tych mięśniach, które w rzeczywistości wywołują oczekiwaną aktywność. Jeśli uwolnienie tego napięcia jest uniemożliwione przez wewnętrzne lub

zewewnętrzne ograniczenie, napięcie będzie utrzymywać się nie tylko w tych mięśniach, które są głównie zaangażowane w oczekiwaną akcję, ale także jako uogólnione napięcie we wszystkich tych mięśniach, które służą do stabilizacji i unieruchomienia ciała.

Cechy charakterystyczne pisma odręcznego, badane są głównie w przestrzeni obrazowej. Pismo odręczne to nie tylko pismo sporządzone ręką, lecz przede wszystkim jest to wytwór mózgu. Na obraz pisma ma wpływ zarówno wczesne, jak i aktualne środowisko. Pewność siebie, szczerłość, beztroška czy wyobraźnia mają wpływ na cechy charakterystyczne pisma. Niecierpliwość i pośpiech, chęć najszybszego wrażenia myśli wyrażają na piśmie cechy indywidualne, umożliwiające identyfikację jej autora. W każdym piśmie odręcznym występuje kombinacja cech grupowych i indywidualnych, różna u różnych osób. Cechy grupowe obejmują konstrukcje liter, ich łączenie oraz wpływy systemowe. Różne cechy grupowe osób lewo i praworęcznych niekiedy stają się cechami indywidualnymi. Charakterystyczną cechą grupową jest kierunek pisania od strony lewej do prawej lub od prawej do lewej, wiązania liter i wiersza, spacjowanie międzyliterowe, odstępy elementów liter itd. Cechy indywidualne pisma odręcznego są mniej lub bardziej szczegółowe dla określonej osoby, stanowiące podstawę do jej identyfikacji. Indywidualność pisma wiąże się z poziomem „dojrzałości graficznej” piszącego²⁶.

Indywidualne cechy charakterystyczne pisma są związane z budową liter i ich liniami wstępnymi i kończącymi, sposobem łączenia, proporcjami międzyliniowymi, nawykami stacjonowania, stosunkami wielkościowymi. Ale nie wszystkie pisma zawierają wysoki stopień indywidualności. Przykładowo podpisy, choć bywają nieczytelne są jednak z reguły indywidualistyczne. Pismo czytelne nie oznacza, że brak mu cech indywidualnych umożliwiających identyfikację wykonawcy. Tekst obszerny zwykle zawiera więcej cech identyfikacyjnych. Wiele cech indywidualnych występuje w liniach wiążących litery, w zależności od ich położenia, w środku wyrazu, na początku lub na końcu. Różne warianty łączy stanowią znaczne trudności w interpretacji łączy. Linie wstępne i końcowe stanowią szczególnie wyraz indywidualności piszącego. W jaki sposób piszący konstruuje litery jest jego cechą indywidualną zwłaszcza jeśli są upiększane. Upiększanie takie występują najczęściej na podpisach i są na tyle bogate, że trudno ustalić, czy odnoszą się do pierwszej litery imienia czy nazwiska. Cechy charakterystyczne pisma odręcznego odzwierciedlają nawyki pisarskie autora i stanowią podstawę do identyfikacji piszącego.

Na jakość pisma mają wpływ: szybkość kreślenia i rytmiczność, łagodność lub ostrość wiązań linii liter, podłoże i biegłość w pisaniu. Jakość albo klasa

26 B. Hołyst, *Kryminalistyka*, Warszawa 2023.

pisma może decydować o wykluczeniu autorstwa podejrzanego o niskiej klasie pisania ale nigdy odwrotnie. W badaniach porównawczych pisma odręcznego istotne znaczenie ma jakość samej linii pisania. Szczególnie wymowne są różnice pomiędzy obrazem linii w podpisach fałszowanych przez naśladownictwo a wzorem oryginalnego podpisu. Z uwagi, że nie każdy pisze z taką samą szybkością, uwzględnienie tej cechy pisma stanowi ważny element identyfikacji. O szybkości pisania świadczy impuls, w przypadku dużej szybkości jest on wyrazowy a nawet zdaniowy. Podczas szybkiego pisania myśl i koncepcja piszącego wbiegają zdecydowanie naprzód. Pisanie nie nadąża za myśleniem, litery stają się coraz szersze, a elementy nad i podlinijne wykazują tendencje do pochylania się zgodnie z kierunkiem pisania.

Ważnym są względne proporcje liter jak: pochylenie, spacjowanie, ułożenie liter i wyrazów w stosunku do nakreślonej lub wyobrażonej linii podłoża, proporcje wielkości liter, długości elementów nad i podlinijnych. Proporcje w piśmie ręcznym są rzadko zmieniane. Trudno to jest uczynić stale bez utraty płynności pisania. Pochylenie pisania wiąże się z wyuczonymi wzorcami, jednak osoba dojrzała wprowadza do pisma cechy indywidualne, również jego pochylenie. Pochylenie pisania jest najczęściej zmieniane w celu zamaskowania własnego pisma. Jest to jednak zabieg najłatwiejszy, ale też łatwy do wykrycia wskutek zmian symetrii linii krzywych w piśmie, jednorodności liter czy znacznego pogorszenia jakości pisania. Ważne cechy charakterystyczne pisma odręcznego to proporcje między wysokimi literami oraz między wysokością a szerokością poszczególnych liter. Każda litera w każdym wyrazie wykazuje określone proporcje do innych liter w odniesieniu do wysokości i szerokości. Różne osoby rozpoczynają i różnie kończą pisma oraz prowadzą wiersze. Duże znaczenie mają akapity i wielkości marginesów, wykorzystanie przestrzeni oraz inne cechy nawykowe. Eksperci kryminalistyki biorą te cechy (topograficzne) pod uwagę w badaniach porównawczych. Kształt linii wstępnych i końcowych jest indywidualną cechą piszącego. Ich obecność lub brak stanowią o naturalnych odmianach liter. Linie wstępne i końcowe odzwierciedlają niezbywalne nawyki piszącego. Długość i kierunek linii wstępnych oraz końcowych jest niepowtarzalną cechą piszącego, zwłaszcza gdy zawierają one różne upiększenia. Kierunek linii pisma jest ważny w ustalaniu czy dana osoba sporządziła dokument dowodowy.

Rozpoznawanie pisma odręcznego jest jak widać sztuką i nauką. Jest nauką, gdyż jest dziedziną wiedzy lub badań zajmujących się zbiorem faktów lub uporządkowanych prawd. Identyfikacja pisma odręcznego jest niczym innym jak wiedzą i stosowaniem metod badawczych.

Główne zasady identyfikacji pisma odręcznego to:

- nie ma dwóch osób, które pisałyby jednakowo,

- nikt po raz drugi nie napisze tego samego tekstu w identyczny sposób,
- piszący nie jest w stanie naśladować wszystkich cech pisma innej osoby, jeżeli pisze z taką samą szybkością i bieginością, jak ten, którego pismo chce naśladować,
- w pismach, w których skutecznie zmienia się własne nawyki pisania albo naśladuje się czyjeś inne nawyki, a nie posiada własnych, zidentyfikowanie naśladowcy jest niezmiernie trudne

Znaczenie cech pisma jako dowodów identyczności lub nieidentyczności wiąże się z cechami rzadkości, względnej szybkości i naturalności, z jaką pisano. Prawdopodobieństwo takiej zgodności pisma dwóch osób, że każda cecha byłaby zgodna, jest praktycznie niemożliwa.

Aktualnie proces identyfikacji pisma odręcznego polega na określeniu cech charakterystycznych wspólnych i różnych w przestrzeni obrazowej i poddawaniu ich ocenie.

Autorzy niniejszego artykułu proponują metodę automatycznego rozpoznawania (klasyfikowania) cech charakterystycznych pisma odręcznego nie w przestrzeni obrazowej ale w przestrzeni transformaty falkowej, która pozwala te cechy łatwo zlokalizować i uwypuklić oraz zautomatyzować ich klasyfikacje.

2. Filtry i falki Gabora

Filtry Gabora proponowane są najczęściej do lokalnej ekstrakcji cech charakterystycznych obrazu. Czyniono próby zastosowania tej techniki do segmentacji tekstury oraz rozpoznawania obrazów^[27]^[28]. Uzyskiwano wektor cech charakterystycznych jako rezultat splotu (konwolucji) funkcji obrazu z dwuwymiarowym filtrem Gabora.

Dwuwymiarowy filtr Gabora w dziedzinie przestrzeni można opisać następującym równaniem:

$$g_w(x, y, \sigma_x, \sigma_y, \theta_n, \omega_k) = \frac{1}{\sigma_x \sigma_y} \cdot \exp \left[-\frac{1}{2} \left(\frac{\eta^2}{\sigma_x^2} + \frac{\vartheta^2}{\sigma_y^2} \right) \right] \cdot \exp \{ [i \omega_k \cdot (\eta + \vartheta)] \} \quad (1)$$

gdzie:

$$\eta = \begin{matrix} \cos \theta_n & \sin \theta_n & x \\ \sin \theta_n & \cos \theta_n & y \end{matrix}$$

27 S. E. Grigorescu, N. Petkov, P. Kruizinga: *Comparison of texture features based on Gabor filters*, IEEE Transactions on Image Processing, **11**, 10, 2002, pp. 142-147.

28 Wenyi Feng, Yingbai Yan, Geaogui Hang, Guofain Jin, *Micro-optical multiwavelet element for hybrid texture segmentation processor*, Optical Engineering, **37**, 1, 1998.

Funkcja filtrowania $g_w(x, y, \sigma_x, \sigma_y, \theta_n, \omega_k)$ nazywana filtrem Gabora może być rozumiana jako modulacja zespolonej fali płaskiej przez funkcje Gaussa w kierunku θ_n i częstotliwości ω_k w przestrzeni $x, y(\sigma_x, \sigma_y)$.

Zatem dyfrakcję (dekompozycje rotacji i częstotliwości dla $\sigma_x = \sigma_y = \sigma = \frac{\pi}{\omega_k}$) fali płaskiej na filtrze Gabora można przedstawić w postaci funkcji:

$$g_w(x, y, \theta_n, \omega_k) = \frac{\omega_k^2}{\pi^2} \cdot \exp \left\{ \left[\frac{\omega_k^2 [(x \cos \theta_n + y \sin \theta_n)^2 + (-x \sin \theta_n + y \cos \theta_n)^2]}{2\pi^2} \right] \right\} \cdot \exp \left\{ i \omega_k [(x \cos \theta_n + y \sin \theta_n) + (-x \sin \theta_n + y \cos \theta_n)] \right\} \quad (2)$$

Z powyższego widać, że ekstrakcja cech charakterystycznych w przestrzeni θ_n, ω_k może być realizowana w prosty sposób w oparciu o filtr Gabora. Rezultat splotu (konwolucji) funkcji obrazu z dwu wymiarowym filtrem Gabora (bankiem filtrów Gabora) można przedstawić w postaci macierzy strumienia intensywności $J(x, y)$.

$$J(x, y) = \begin{bmatrix} g_w(x, y, \omega_1, \theta_1) & g_w(x, y, \omega_1, \theta_2) & \dots & g_w(x, y, \omega_1, \theta_N) \\ g_w(x, y, \omega_2, \theta_1) & g_w(x, y, \omega_2, \theta_2) & \dots & g_w(x, y, \omega_2, \theta_N) \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ g_w(x, y, \omega_M, \theta_1) & g_w(x, y, \omega_M, \theta_2) & \dots & g_w(x, y, \omega_M, \theta_N) \end{bmatrix} \quad (3)$$

gdzie: θ - kąt zorientowania tekstury, ω - częstotliwość w obszarze tekstury, σ_x, σ_y - analizowany obszar tekstury

$$\omega_k = 2 \frac{m}{2} \quad z \quad m = \{1, \dots, M\} \quad M > 1$$

$$\theta_n = n \frac{\pi}{N} \quad z \quad n = \{1, \dots, N\} \quad N > 1$$

Matryca przefiltrowanego strumienia intensywności ($J(x, y)$) pikseli funkcji obrazu może być zatem interpretowana jako dyskretyzacja przestrzeni cech tekstury (θ_n, ω_k). Ta przestrzeń znana jest jako przestrzeń Gabora.

Naturalnym narzędziem przeniesienia funkcji z dziedziny przestrzeni do dziedziny częstotliwości jest transformata Fouriera. Dyskretna transformata Fouriera macierzy strumienia intensywności $\mathfrak{J}_{x,y}$ w przestrzeni Gabora definiowana jest jako:

$$\mathfrak{J}_{x,y} \equiv F_{\mu,\nu} = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N J_{x,y} \cdot \exp \left\{ -i2\pi \left[\frac{(u-1) \cdot (x-1)}{M} + \frac{(v-1) \cdot (y-1)}{N} \right] \right\} \quad (4)$$

Zaś amplituda transformaty Fouriera $|F_{\mu,v}|$ strumienia intensywności $J(x,y)$ w przestrzeni Gabora może być wyrażona w prosty sposób przez stałą przesunięcia $G_{\mu,v}$ macierzy $J(x,y)$.

$$G_{\mu,v} = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N J_{x+a,y+b} \cdot \exp\left\{-i2\pi\left(\frac{(u-1)\cdot(x-1)}{M} + \frac{(v-1)\cdot(y-1)}{N}\right)\right\} =$$

$$\frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N J_{x',y'} \cdot \exp\left\{-i2\pi\left[\frac{(u-1)\cdot(x-a-1)}{M} + \frac{(v-1)\cdot(y'-b-1)}{N}\right]\right\}$$

Dokonajmy prostego przekształcenia

$$G_{\mu,v} = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N J_{x',y'} \cdot e^{-i2\pi\left(\frac{(u-1)\cdot(x'-1)}{M} + \frac{(v-1)\cdot(y'-1)}{N}\right)} \cdot e^{i2\pi\left(\frac{(u-1)\cdot(a)}{M} + \frac{(v-1)\cdot(b)}{N}\right)} =$$

$$= F_{\mu,v} \cdot e^{i2\pi\left(\frac{(u-1)\cdot(a)}{M} + \frac{(v-1)\cdot(b)}{N}\right)}$$

stąd

$$|F_{\mu,v}| = \left| \Im f(x,y) \cdot \exp\left\{i2\pi\left(\frac{(u-1)\cdot(a)}{M} + \frac{(v-1)\cdot(b)}{N}\right)\right\} \right| = |G_{\mu,v}| \quad (5)$$

Oznacza to, że moduł amplitudy transformaty Fouriera ($|F_{\mu,v}|$) macierzy strumienia intensywności ($J(x,y)$) w przestrzeni Gabora może być obliczony jako moduł wartości stałego przesunięcia ($|G_{\mu,v}|$) macierzy intensywności, która to jest podstawową niezmienniczą cechą transformaty Fouriera.

Falki Gabora

Elementarną funkcję Gabora $G(x,y)$ można wykorzystać jako falkę matkę do generowania rodziny falek Gabora. Matematycznie 2D funkcję Gabora dla przypadku ogólnego można przedstawić w postaci^{29,30,31,32}:

$$G(x,y) = \frac{1}{2\pi\sigma\beta} \exp\left\{-\pi\left[\frac{(x-x_o)^2}{\sigma^2} + \frac{(y-y_o)^2}{\beta^2}\right]\right\} \cdot \exp[i(\xi_o x + \nu_o y)] \quad (6)$$

gdzie (x_o, y_o) jest środkiem analizowanego pola w przestrzeni domeny (σ, β) , ξ_o, ν_o są częstotliwościami przestrzennymi filtra w kierunku x i y , natomiast σ, β jest standardowym

- 29 T. S. Lee: *Image representation using 2D Gabor wavelets*, IEEE Transactions on pattern analysis and machine intelligence, **18**, 10, October 1996.
- 30 J. G. Daugman: *Uncertainty relation for resolution In space, spatial frequency, and orientation optimized by two-dimensional visual cortex filters*, J. Opt. Soc. Am. A **2**, 7, 1985, pp. 1160-1169.
- 31 J. G. Daugman: *Two-Dimensional Spectral Analysis of Cortical Receptive Field Profile*, Vision Research, **20**, 1980, pp. 847-856.
- 32 R. J. Duffin and A. C. Schaeffer: *A Class of Nonharmonic Fourier Series*, Trans. Am. Math. Soc., **72**, 1952, pp. 341-366.

odchyleniem eliptycznym funkcji Gaussa wzdłuż osi x i y . Zatem 2D funkcja Gabora jest iloczynem eliptycznej funkcji Gaussa i zespolonej fali płaskiej. 2D funkcja Gabora osiąga granicę rozdzielczości tylko w jej zespolonej postaci. Zespolona funkcja Gabora zawiera parzysty symetryczny składnik cosinusowy i nieparzysty składnik sinusoidalny.

Ostatnio filtry Gabora, ponownie zostały odkryte i uogólnione, i są obecnie stosowane w różnych wizyjnych systemach komputerowych. Zauważono, że zespół prostych komórek najlepiej jest modelować jako rodzinę 2D falek Gabora z próbkowaniem domeny częstotliwości w sposób logarytmiczny. Ten klasyczny filtr jest równoważny rodzinie wygenerowanych afinicznie stanów koherentnych przez grupę afiniczną. Dekompozycje obrazu w stany koherentne nazywamy *transformatą falkową* $Wf(a, \theta, x_o, y_o)$ funkcji obrazu $f(x, y)$:

$$Wf(a, \theta, x_o, y_o) = \|a\|^{-1} \iint f(x, y) \cdot \psi_\theta \left(\frac{x - x_o}{a}, \frac{y - y_o}{b} \right) dx dy \quad (7)$$

gdzie: a jest parametrem dylatacji związanym z σ , β , natomiast x_o i y_o to parametry translacji przestrzennej, θ parametr orientacji falki (ψ_θ)

$$\psi_\theta(a, x, y, x_o, y_o) = \|a\|^{-1} \psi_\theta \left(\frac{x - x_o}{a}, \frac{y - y_o}{b} \right). \quad (8)$$

Elementarna 2D funkcja falkowa obracana jest o kąt θ . Przykłady 2D falek Gabora przedstawiono w pracy [10].

Przeanalizujemy rodzinę 2D falek Gabora $\psi(x, y, \omega_o, \theta)$ spełniających teorię falkową i ich ogólne falkowe ograniczenia.

$$\psi(x, y, \omega_o, \theta) = \frac{\omega_o}{\sqrt{2\pi k}} \exp - \left\{ \frac{\omega_o^2}{8k^2} \left[4(x \cos \theta + y \sin \theta)^2 + (-x \sin \theta + y \cos \theta)^2 \right] \right\}. \quad (9)$$

$$\exp[i(\omega_o x \cos \theta + \omega_o y \sin \theta)]$$

gdzie: ω_o jest częstotliwością promieniową w radianach na jednostkę długości, θ jest orientacją falki w radianach. Falka Gabora jest wyśrodkowana gdy $x = 0$, $y = 0$ i normalizowana w taki sposób aby współczynniki falkowe spełniały relacje $\langle \psi, \psi \rangle = 1$, tj. aby to była falka w przestrzeni L^2 , k jest stałą, gdzie $k = \pi$ dla pasma częstotliwości jednej oktawy i $k = 2,5\pi$ dla pasma o częstotliwości 1,5 oktawy. Dyskretny zbiór o szerokości pasma 1,5 oktawy generuje rodzinę [16] falek Gabora.

Transformata falkowa $Wf(a, \theta, x_o, y_o)$ jest rzutem funkcji obrazu na falkę Gabora, która jest iloczynem funkcji obrazu f z polem wyśrodkowanym na (x_o, y_o) ,

$$Wf(a, \theta, x_o, y_o) = \langle \psi(x, y, \omega_o, \theta), f \rangle = \iint \psi(x_o - x, y_o - y, \omega_o, \theta) \cdot f(x, y) dx dy \quad (10)$$

Rozpocznijmy dyskusję od najbardziej ogólnego przypadku zespolonej falki Gabora $\psi(x, y, \xi_o, v_o, x_o, y_o, \theta, \sigma, \beta)$, tzn. tak aby była $L^2 < \psi, \psi \rangle = 1$.

$$\begin{aligned} \psi(x, y, \xi_o, v_o, x_o, y_o, \theta, \sigma, \beta) = \\ \frac{\omega_o}{\sqrt{\pi\sigma\beta}} \exp - \left\{ \left[\frac{[(x-x_o)\cos\theta + (y-y_o)\sin\theta]^2}{2\sigma^2} \right] + \frac{[-(x-x_o)\sin\theta + (y-y_o)\cos\theta]^2}{2\beta^2} \right\} \cdot \\ \exp[i(\xi_o(x-x_o) + v_o(y-y_o))] \end{aligned} \quad (11)$$

Filtr falkowy jest wyśrodkowany gdy $x = x_o$, $y = y_o$ w domenie przestrzennej σ, β o częstotliwości $\xi = \xi_o$, $v = v_o$, które są eliptycznymi odchyleniami standardowymi funkcji Gaussa wzdłuż osi x i y . Kąt θ jest orientacją filtra, obróconego przeciwnie do ruchu wskazówek zegara wokół początku układu współrzędnych. Istnieje zatem siedem stopni swobody ogólnej funkcji Gabora: $\xi_o, v_o, x_o, y_o, \theta, \sigma, \beta$. Ponieważ tylko moduł mocy ma istotne znaczenie w praktycznych zastosowaniach, zatem możemy uprościć *falkowy filtr Gabora*, ustalając położenie przestrzenne środka filtra w punkcie $x_o = 0, y_o = 0$.

Transformata Fouriera $\Im\psi(\xi, v, \xi_o, v_o, \theta, \sigma, \beta)$ *uproszczonego filtra Gabora* o wartościach zespolonych to funkcja:

$$\begin{aligned} \Im\psi(\xi, v, \xi_o, v_o, \theta, \sigma, \beta) = 2\sqrt{\pi\sigma\beta} \cdot \left\{ \exp - \frac{1}{2} \left[(\xi - \xi_o)\cos\theta + (v - v_o)\sin\theta \right]^2 \sigma^2 \right\} + \\ 2\sqrt{\pi\sigma\beta} \cdot \left\{ \exp - \frac{1}{2} \left[-(\xi - \xi_o)\sin\theta + (v - v_o)\cos\theta \right]^2 \beta^2 \right\} \end{aligned} \quad (12)$$

gdzie ξ, v są częstotliwościami przestrzennymi w radianach na jednostkę długość wzdłuż osi x i y . Możemy jeszcze bardziej ograniczyć liczbę stopni swobody filtra. Zauważmy, że θ, σ, β są ograniczone w zastosowaniach praktycznych przez ξ_o, v_o dla przestrzennego pasma częstotliwości do zakresu od 0,5 do 2,5 oktawy. przy obwiedni Gaussa z eliptycznym stosunkiem 1,5-2,0 [29] przy propagacji fali płaskiej wzdłuż krótkiej osi eliptycznej Gaussa.

Ograniczenie 1. Współczynnik proporcji σ/β koperty eliptycznej Gaussa z reguły wynosi 2:1 [18].

Ograniczenie 2. Fala płaska o częstotliwości (ξ_o, v_o) ma wyróżniony „kierunek propagacji” wzdłuż krótkiej osi eliptycznej Gaussa. Kiedy fala płaska obraca się, to również Gaussian eliptyczny obraca się odpowiednio. Oznacza to wyróżnioną częstotliwość środkową.

Przyjmijmy konwencje transformaty Fouriera $\mathfrak{F}(f(x,y))$ definiowaną w postaci:

$$\begin{aligned}\mathfrak{F}(f(x,y)) &= \iint f(x,y) \cdot \exp(-i\xi x) \cdot \exp(-iv y) dx dy \\ f(x,y) &= \frac{1}{4\pi^2} \iint \mathfrak{F}(f(x,y)) \cdot \exp(i\xi x) \cdot \exp(iv y) d\xi dv\end{aligned}\quad (13)$$

Częstotliwości (ξ_o, v_o) filtra są związane z kątem obrotu θ i są modulowane Gaussianem postaci $\xi_o = \omega_o \cos \theta$ i $v_o = \omega_o \sin \theta$, gdzie częstotliwość promieniowa $\omega_o = \sqrt{\xi_o^2 + v_o^2}$.

Dzięki tym dwóm ograniczeniom *uproszczony filtr falkowy Gabora* $\psi(x,y,\xi_o,v_o,\sigma)$ staje się:

$$\begin{aligned}\psi(x,y,\xi_o,v_o,\sigma) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{8\sigma^2} \left[4 \left[x \frac{\xi_o}{|\omega_o|} + y \frac{v_o}{|\omega_o|} \right]^2 + \left[-x \frac{v_o}{|\omega_o|} + y \frac{\xi_o}{|\omega_o|} \right]^2 \right] \right\} \\ & [i(\xi_o x + v_o y)]\end{aligned}\quad (14)$$

Orientacja filtra jest wyrównana zgodnie z długą osią eliptyczną Gaussa.

Ograniczenie 3. Szerokość pasma dla amplitudy o połowie częstotliwości odpowiedzi wynosi od 1 do 1,5 oktawy wzdłuż optymalnej orientacji [30-32]. Związek między σ i ω_o można uznać jako,

$$\sigma = \frac{k}{\omega_o} \quad k = \sqrt{2} \ln \frac{2^\theta + 1}{2^\theta - 1} \quad (15)$$

gdzie θ jest pasmem w oktawach. Dla $\theta = 1$ oktawy, $\sigma = \frac{\pi}{\omega_o} = \frac{\lambda}{2}$. Dla $\theta = 1,5$ oktawy, $\sigma = \frac{2,5}{\omega_o}$. Nakładając to ograniczenie, otrzymujemy *rodzinę samopodobnych filtrów falkowych Gabora* $\psi(x,y,\omega_o,\theta)$ w zależności od czterech zmiennych,

$$\begin{aligned}\psi(x,y,\omega_o,\theta) &= \frac{\omega_o}{\sqrt{2\pi}k} \exp\left\{-\frac{\omega_o^2}{8k^2} \left[4 \left[x \cos \theta + y \sin \theta \right]^2 + \left[-x \sin \theta + y \cos \theta \right]^2 \right] \right\} \\ & \exp[i(x\omega_o \cos \theta + y\omega_o \sin \theta)]\end{aligned}\quad (16)$$

gdzie $\theta = \arctan \frac{v_o}{\xi_o}$, a k jest ustalone dla falek Gabora dla konkretnej przepustowości. Całą rodzinę falek można określić dla dowolnej pozycji przestrzennej (x_o, y_o) . Aby zdefiniować filtry Gabora dla dopuszczalnych falek, musimy wprowadzić kolejne ograniczenie.

Ograniczenie 4. Dopuszczalne falki to funkcje mające ograniczoną wartość. Rozważmy zbiór falek $\psi(f) : L^2(\mathbb{R}^2), L^2(\mathbb{R}^4)$. Falka Gabora jest dopuszczalna tylko wtedy, gdy jej normą jest wartość skończona, tj.

$$\|\psi(f)\|_{L^2(\mathbb{R}^4)} = \int_R \int_R \int_o^{2\pi} \int_o^\infty |\psi(f(\omega,\theta,x,y))|^2 \frac{d\omega}{\omega} d\theta dx dy = \|\mathfrak{F}f\|_{L^2(\mathbb{R}^2)} \cdot \|\mathfrak{F}\psi\|_{L^2\left(\frac{d\omega}{\omega} d\theta\right)} < \infty \quad (17)$$

Oznacza to, że $\|\Im\psi(\omega) = 0\|$ jest warunkiem koniecznym. W przeciwnym razie norma stanie się nieskończona w miarę $\frac{d\omega}{\omega}$ gdy $\omega \rightarrow 0$.

Składnik sinusoidalny filtra Gabora o wartości zespolonej ma zerową wartość średnią, ale jego składowa cosinusoidalna ma niezerową wartość średnią (odpowiedź DC). Odpowiedź DC można obliczyć z transformata Fouriera, dla $\xi = 0$ i $v = 0$.

$$\Im(\psi(\xi = 0, v = 0; \xi_o, v_o)) = \sqrt{8\pi}\sigma \exp\left(-\frac{k^2}{2}\right) \quad (18)$$

Stąd można uzyskać *rodzinę dopuszczalnych 2D falek Gabora* $\psi(x, y, \omega_o, \theta)$ odejmując odpowiedź DC od filtra Gabora,

$$\psi(x, y, \omega_o, \theta) = \frac{\omega_o}{\sqrt{2\pi}k} \exp\left(-\frac{\omega_o^2}{8k^2} \left\{4\left[x(\cos\theta + y\sin\theta)^2\right] + \left[-x(\sin\theta + y\cos\theta)^2\right]\right\}\right) \cdot \left[\exp i(x\omega_o \cos\theta + y\omega_o \sin\theta) - \exp\left(-\frac{k^2}{2}\right) \right] \quad (19)$$

z transformatą Fouriera

$$\Im\psi(\xi, v, \xi_o, v_o) = \frac{\sqrt{8\pi}k}{\omega_o} \cdot \left\{ \exp\left(-\frac{k^2}{2\omega_o^2} \left[(\xi - \xi_o)\cos\theta + (v - v_o)\sin\theta \right]^2 + \left[4(-(\xi - \xi_o)\sin\theta + (v - v_o)\cos\theta) \right]^2 \right) \right\} - \exp\left(-\frac{k^2}{2\omega_o^2} \left\{ \left[(\xi \cos\theta + v \sin\theta)^2 + 4(-\xi \sin\theta + v \cos\theta)^2 + \omega_o^2 \right] \right\} \right) \quad (20)$$

gdzie: $\xi\xi_o + vv_o = 0$

Każda z tych dwóch rodzin falek Gabora może zostać wygenerowana przez rotację i dylatację (grupa afiniczna) z falki matki Gabora $\psi(x, y)$,

$$\psi(x, y) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{8}(4x^2 + y^2)\right] \cdot \left[\exp(ikx) - \exp\left(-\frac{k^2}{2}\right) \right] \quad (21)$$

gdzie transformata Fouriera

$$\Im\psi(x, y) = \sqrt{8\pi} \left\{ \exp\left[-\frac{1}{2}(\xi - k)^2 + 4v^2\right] - \left[\exp\left(-\frac{1}{2}[\xi^2 + 4v^2 + k^2]\right) \right] \right\} \quad (22)$$

Uzyskaliśmy odpowiednio sparametryzowane równanie falkowe Gabora i możemy rozwiązać problem reprezentacji zupełności. Mówi się, że przekształcenie zapewnia kompletna reprezentacja, jeśli możemy zrekonstruować funkcje na drodze numerycznie stabilnym od przekształcenia funkcji f lub alternatywnie, jeśli w ogóle funkcję f można zapisać jako superpozycję funkcji przekształcenia podstawowego.

Zauważmy, że ciągła transformata falkowa $Wf(a, \theta, x_o, y_o)$ jest definiowana jako:

$$Wf(a, \theta, x_o, y_o) = \|a\|^{-1} \iint f(x, y) \cdot \psi_{\theta} \left(\frac{x-x_o}{a}, \frac{y-y_o}{a} \right) dx dy \quad (23)$$

gdzie: a jest parametrem dylatacji, x_o i y_o przestrzenne parametry translacji, θ parametr orientacji falki.

Falka $\psi_{\theta}(a, x, y, x_o, y_o) = \|a\|^{-1} \psi_{\theta} \left(\frac{x-x_o}{a}, \frac{y-y_o}{b} \right)$ jest ementarną funkcją falkową, obróconą o θ . Funkcję zawsze można zrekonstruować z jej ciągłej transformaty falkowej korzystając z wzoru tożsamościowego, pod warunkiem, że falki są dopuszczalne

$$f = \frac{1}{C_{\psi}} \int_{O}^{\infty} \frac{da}{a^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy \int_0^{2\pi} \langle f, \psi(a, x, y, \theta) \rangle \psi(a, x, y, \theta) d\theta \quad (24)$$

gdzie:

$$C_{\psi} = 4\pi^2 \int_0^{\infty} \frac{d\omega}{\omega} \int_0^{2\pi} |\Im \psi(\omega \cos \theta, \omega \sin \theta)|^2 d\theta \quad (25)$$

$$\omega = \sqrt{\xi^2 + \nu^2}, \quad \psi \in L^1(R), \quad C_{\psi} < \infty, \quad \Im \psi(O, O) = O$$

a, x, y, θ przyjmują wartości dyskretne tj.

$$a = a_o^{-m}, \quad a_o > 1, \quad \theta = \theta_l = l\theta_o, \quad x = nb_o, \quad y = kb_o a_o n$$

m i n, l, k są liczbami całkowitymi. Zatem transformata falkowa

$$Wf(m, n, k, l) = a_o^{-1} \iint f(x, y) \cdot \psi_{\theta_l} \left(a_o^{-m} x - nb_o, a_o^{-m} y - kb_o \right) dx dy \quad (26)$$

nazywa się *dyskretną transformatą falkową* $Wf(m, n, k, l)$, gdzie ψ_{θ_l} to wersja obróconej falki macierzystej $\psi(x, y)$. $a_o > 1, \theta_o > 0$ i $b_o > 0$

W tym schemacie falka przestrzennie węższa zapewnia drobniejsze kroki, a szersza falka zapewnia większe kroki. W rezultacie zdyskretyzowane falki przy każdym poziomie m „pokrywają” domenę przestrzenną w ten sam sposób jak $\psi(x - nb_o, y - kb_o)$ na poziomie $m = 0$.

Falki dyskretne są generowane w postaci

$$\begin{aligned} \psi_{m,n,k,l}(x,y) &= a_o^{-m} \psi_{\theta_l} \left(a_o^{-m} \left(x - nb_o, a_o^m \right), a_o^{-m} \left(y - kb_o, a_o^m \right) \right) = \\ &= a_o^{-m} \psi_{\theta_l} \left(a_o^{-m} x - nb_o, a_o^{-m} y - kb_o \right) \end{aligned} \tag{27}$$

Zauważmy, że chociaż parametryzacja jest dyskretna, to każda elementarna funkcja falkowa jest funkcją ciągłej zmiennej x i y . Niektóre dyskretne falki 1D, na przykład jednowymiarowe bazy Harra, fale Meyera i falki Battle-Lemarie tworzą bazy ortonormalne. W tym przypadku funkcja może być zrekonstruowana przez liniową superpozycję bazy to znaczy przez *współczynniki falkowe*,

$$f = \sum_{m,n} \langle f, \psi_{m,n} \rangle \psi_{m,n} \tag{28}$$

gdzie $\langle a_1, a_2 \rangle$ oznacza iloczyn skalarny a_1 i a_2 . Jednak falki Gabora 1D i 2D nie tworzą ortonormalnej bazy. Nazywa się je *falkami nieortogonalnymi*. W tym przypadku interesuje nas, czy zbiór dyskretny fal nieortogonalnych tworzy ramkę. Jeśli fale tworzą ramkę, wtedy można jednoznacznie scharakteryzować funkcję f znając $W(f)$ i można zrekonstruować f w postaci numerycznej na drodze stabilnej od $W(f)$ z podwójną ramką $\Im\psi$ [33],

$$f = \sum_{m,n,k,l} \langle f, \psi_{m,n,k,l} \rangle \Im\psi_{m,n} = \sum_{m,n,k,l} \langle f, \Im\psi_{m,n,k,l} \rangle \psi_{m,n,k,l} \tag{29}$$

Koncepcja ram została po raz pierwszy wprowadzona przez Duffina i Schaffera [29] w kontekście nieharmonicznych szeregów Fouriera.

3. Funkcje skalujące i falki w dziedzinie częstotliwości

Funkcje skalujące i falki mogą być wyrażone w dziedzinie częstotliwości, czyli w przestrzeni Fouriera [10].

Transformatę Fouriera funkcji skalującej $\phi(\omega)$ definiujemy jako:

$$\phi(\omega) = \int_{-\infty}^{\infty} \phi(t) \cdot e^{-i\omega t} dt \tag{30}$$

Mnożąc obie strony równania rozszerzającego $\phi(t) = \sqrt{2} \sum_k h_k \phi(2t - k)$ przez $e^{-i\omega t}$ i całkując otrzymujemy:

$$\begin{aligned} \hat{\phi}(\omega) &= \int_{-\infty}^{\infty} \phi(t) e^{-i\omega t} dt = \sum_k \sqrt{2} h_k \int_{-\infty}^{\infty} \phi(2t - k) e^{-i\omega t} dt \\ &= \left(\sum_k \frac{h_k}{\sqrt{2}} e^{-i k \omega / 2} \right) \int_{-\infty}^{\infty} \phi(2t - k) e^{-i(2t-k)\frac{\omega}{2}} d(2t) \\ &= \left(\sum_k \frac{h_k}{\sqrt{2}} e^{-i k \omega / 2} \right) \hat{\phi} \left(\frac{\omega}{2} \right). \end{aligned}$$

Definiując następnie:

$$H(\omega) = \left(\sum_k \frac{h_k}{\sqrt{2}} e^{-i k \omega} \right), \quad (31)$$

otrzymujemy równanie rozszerzające funkcję skalującą w przestrzeni Fouriera:

$$\hat{\phi}(\omega) = H\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right). \quad (32)$$

Analogicznie postępując, transformatę Fouriera falki podstawowej wyrażamy w postaci $\hat{\psi}(\omega) = \int_{-\infty}^{\infty} \psi(t) e^{-i\omega t} dt$ oraz definiując:

$$G(\omega) = \sum_k \frac{g_k}{\sqrt{2}} e^{-i k \omega}, \quad (33)$$

otrzymujemy równanie rozszerzające falkę w przestrzeni Fouriera w postaci:

$$\hat{\psi}(\omega) = G\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right). \quad (34)$$

Jeżeli $\phi(t)$ jest ortonormalna do jej przesunięcia $\phi(t - k)$

$$\begin{aligned} \delta_{k,0} &= \int_{-\infty}^{\infty} \phi(t) \overline{\phi(t-k)} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\phi}(\omega) \overline{\hat{\phi}(\omega)} e^{i\omega k} d\omega \\ &= \frac{1}{2\pi} \int_0^{2\pi} \sum_{l=-\infty}^{\infty} |\hat{\phi}(\omega + 2\pi l)|^2 e^{i\omega k} d\omega. \end{aligned} \quad (35)$$

to otrzymujemy współczynniki a_k w szeregach Fouriera funkcji okresowej: $A(\omega) = \sum_{l=-\infty}^{\infty} |\hat{\phi}(\omega + 2\pi l)|^2$. Inaczej mówiąc,

$$\sum_{l=-\infty}^{\infty} |\hat{\phi}(\omega + 2\pi l)|^2 = \sum_k a_k e^{-i k \omega} = 1. \quad (36)$$

Zatem ortonormalność funkcji skalującej $\{\phi(t - k)\}$ może być wyrażona w dziedzinie częstotliwości, jako:

$$\sum_{l=-\infty}^{\infty} |\hat{\phi}(\omega + 2\pi l)|^2 = \sum_k a_k e^{-i k \omega} = 1. \quad (37)$$

Jeżeli zbiór funkcji skalujących $\{\phi(t - k)\}$ nie jest ortonormalny, to można je łatwo ortogonalizować w przestrzeni Fouriera. Należy tylko podzielić $\hat{\phi}(\omega)$ przez pierwiastek kwadratowy z wyrażenia $\sum_{l=-\infty}^{\infty} |\hat{\phi}(\omega + 2\pi l)|^2$,

$$\hat{\phi}_{orth}(\omega) = \frac{\hat{\phi}(\omega)}{\sqrt{\sum_{l=-\infty}^{\infty} |\hat{\phi}(\omega + 2\pi l)|^2}} \quad (38)$$

stąd

$$\sum_{k=-\infty}^{\infty} |\hat{\phi}_{orth}(\omega + 2\pi k)|^2 = \sum_k \left| \frac{\hat{\phi}(\omega + 2\pi k)}{\sqrt{\sum_{l=-\infty}^{\infty} |\hat{\phi}(\omega + 2\pi l)|^2}} \right|^2 = 1. \quad (39)$$

Zatem teraz $\hat{\phi}_{orth}(\omega)$ spełnia warunek ortonormalności (37) i staje się $\{\phi_{orth}(t - k)\}$ ortogonalna. Korzystając z równania rozszerzającego, możemy napisać *warunek ortonormalności*, jako warunek na $H(\omega)$: w dziedzinie częstotliwości:

$$\begin{aligned} \sum_{l=-\infty}^{\infty} |\hat{\phi}(\omega + 2\pi l)|^2 &= \sum_{l=-\infty}^{\infty} |H(\omega + 2\pi l)|^2 |\hat{\phi}(\omega + 2\pi l)|^2 \\ &= \sum_{l=-\infty}^{\infty} |H(\omega' + \pi l)|^2 |\hat{\phi}(\omega' + \pi l)|^2 = 1. \end{aligned}$$

Ponieważ powyższe równanie jest prawdziwe dla wszystkich $\omega = \omega/2$, stąd możemy ω' po prostu zastąpić przez ω . Rozdzielając sumę na parzyste ($l=2k$) i nieparzyste ($l=2k+1$) składniki, otrzymujemy:

$$\sum_{k=-\infty}^{\infty} |H(\omega + 2k\pi)|^2 |\hat{\phi}(\omega + 2k\pi)|^2 + \sum_{k=-\infty}^{\infty} |H(\omega + (2k+1)\pi)|^2 |\hat{\phi}(\omega + (2k+1)\pi)|^2 = 1.$$

Ponownie wykorzystując fakt, iż funkcja $H(\omega)$ jest okresowa ((2π)), powyższe równanie redukuje się do równania:

$$|H(\omega)|^2 \sum_{k=-\infty}^{\infty} |\hat{\phi}(\omega + 2k\pi)|^2 + |H(\omega + \pi)|^2 \sum_{k=-\infty}^{\infty} |\hat{\phi}((\omega + \pi) + 2k\pi)|^2 = 1.$$

Stąd otrzymujemy *warunek ortogonalności dla $H(\omega)$ w dziedzinie częstotliwości*

$$|H(\omega)|^2 + |H(\omega + \pi)|^2 = 1, \quad (40)$$

tzn. warunek na $\{h_k\}$ dla zbioru funkcji $\{\phi(t - k)\}$ aby był ortogonalny.

Warunki ortogonalności dla falek w dziedzinach częstotliwości mogą być wyprowadzone w analogiczny sposób. Przykładowo dla zbioru falek $\{\psi(t - k)\}$ aby był on ortonormalny musi być spełniony warunek:

$$\sum_{l=-\infty}^{\infty} |\hat{\psi}(\omega + 2\pi l)|^2 = 1, \quad (41)$$

lub

$$|G(\omega)|^2 + |G(\omega + \pi)|^2 = 1.$$

Możemy także wyprowadzić warunek w przestrzeni częstotliwości taki, że $\{\psi(t-k)\}$ oraz $\{\phi(t-k)\}$ będą ortogonalne,

$$\begin{aligned} 0 &= \int_{-\infty}^{\infty} \phi(t) \overline{\psi(t-k)} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\phi}(\omega) \hat{\psi}(\omega) e^{i\omega k} d\omega \\ &= \frac{1}{2\pi} \int_0^{2\pi} \sum_{l=-\infty}^{\infty} \hat{\phi}(\omega + 2\pi l) \overline{\hat{\psi}(\omega + 2\pi l)} e^{i\omega k} d\omega. \end{aligned}$$

Jako, że współczynniki Fouriera funkcji $\sum_{l=-\infty}^{\infty} \hat{\phi}(\omega + 2\pi l) \overline{\hat{\psi}(\omega + 2\pi l)}$ są równe zeru, można wywnioskować, że

$$\sum_{l=-\infty}^{\infty} \hat{\phi}(\omega + 2\pi l) \overline{\hat{\psi}(\omega + 2\pi l)} = 0.$$

Podobnie, wykorzystując równanie funkcji skalującej $\hat{\phi}(\omega) = H\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right)$ i równania falki $\hat{\psi}(\omega) = G\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right)$ otrzymujemy:

$$H(\omega) \overline{G(\omega)} + H(\omega + \pi) \overline{G(\omega + \pi)} = 0. \quad (42)$$

Jest to warunek, który musi być spełniony, aby $\{\psi(t-k)\}$ oraz $\{\phi(t-k)\}$ były ortogonalne. To równanie wiąże współczynniki występujące w równaniu falki $\{g_k\}$ z współczynnikami występującymi w równaniu funkcji skalującej $\{h_k\}$. Aby znaleźć wyraźne związki pomiędzy dwoma zbiorami współczynników należy rozwiązać powyższe równanie. Oczywiście jest, iż powyższe równanie jest spełnione jeżeli $G(\omega)$ jest wybrane tak, aby spełniać relacje:

$$\overline{G(\omega)} + \lambda(\omega) \overline{H(\omega + \pi)},$$

gdzie $\lambda(\omega) = -\lambda(\omega + \pi)$ jest funkcją okresową (2π) oraz $|\lambda(\omega)|^2 = 1$. Warunki ortogonalności nie precyzują jednoznacznie G lub falkę, mamy zatem swobodę wyboru funkcji $\lambda(\omega)$. Zastosujmy zwykłą konwencję wyboru $\lambda(\omega) = -e^{-i\omega}$

W związku z powyższym ustaliliśmy (określiśmy) związki pomiędzy współczynnikami funkcji skalującej, a współczynnikami falki. Są one powiązane przez:

$$G(\omega) = -e^{-i\omega} H(\omega + \pi) \quad (43)$$

lub

$$G(\omega) = -e^{-i\omega} \sum_k \frac{h_k}{\sqrt{2}} e^{i(\omega+\pi)k} = \sum_k \frac{h_k}{\sqrt{2}} (-1)^{1-k} e^{-i\omega(1-k)} = \sum_n \frac{(-1)^n h_{1-n}}{\sqrt{2}} e^{-i\omega n}.$$

Zatem dochodzimy do wniosku, że

$$g_n = (-1)^n h_{1-n}. \quad (44)$$

Można więc zauważyć, że gdy funkcja skalująca ϕ jest znana, współczynniki $\{h_n\}$ są otrzymywane przez obliczenie iloczynów ϕ i ϕ_{1k} . Następnie współczynniki $\{g_n\}$ są obliczane z (44) i wykorzystując równanie falki, otrzymujemy falkę. Z drugiej strony, jeżeli znamy współczynniki $\{h_n\}$, możemy znaleźć funkcję skalującą poprzez rozwiązanie równania rozszerzającego.

Zauważmy, że naszym celem jest znaleźć ortogonalną funkcję skalującą i falkę, które są nośnikami zwartymi i równe są określonym momentom znikającym. Wskaźniki te są wskaźnikami w rozszerzonym równaniu oraz po odpowiedniej transformacji w równaniu falkowym. Te równania określają funkcję skalującą i falkę matkę.

Obecnie skupimy się na znalezieniu rozwiązania równania rozszerzającego i równań falkowych dających współczynniki filtrów. Jednoczesne wymagania ortogonalności i zwartości nośnika generalnie zabezpieczają nas od poszukiwania analitycznych zależności na $\phi(t)$ i $\psi(t)$, więc zadowolamy się znalezieniem ich numerycznie i przedstawienie graficznie.

Funkcja skalująca nośnika zwartego spełnia równanie rozszerzenia:

$$\phi(t) = \sum_{k=0}^n \sqrt{2} h_k \phi(2t - k) = \sum_{k=0}^n c_k \phi(2t - k) \quad (45)$$

gdzie nośnik $\phi(t)$ jest z przedziału $[0, n]$. Generalnie nośnik ϕ może być w przedziale $[n_1, n_2]$ dla $n_1 \neq 0$; w tym znaczeniu sumowanie po k rozszerzone jest od n_1 do n_2 . W (45) nie ma straty uniwersalności kiedy ϕ osiągnięte jest przez sumowanie po k od n_1 do n_2 . Jest ono po prostu powiązane z sumowaniem k od 0 do $n_2 = n_2 - n_2$ przez przesunięcie n_1 .

Jest kilka sposobów na obliczenie funkcji skalującej $\phi(t)$ z (45). Przykładowo pierwszą metodą jest rozszerzenie równania $\phi(t) = \sum_{k=0}^n c_k \phi_{j-1}(2t - k)$ zaczynając od funkcji $\phi(t)$, do czasu aż funkcja się zbiegnie. Kolejną metodą jest przedstawienie iteracji w obszarze Fouriera. Rozszerzenie równania w dziedzinie częstotliwości ma postać:

$$\hat{\phi}(\omega) = H(\omega/2) \hat{\phi}(\omega/2)$$

co może być wykorzystane do obliczenia:

$$\hat{\phi}(\omega) = \prod_{j=1}^{\infty} H\left(\frac{\omega}{2^j}\right), \quad (46)$$

czego skutkiem, jest funkcja skalująca $\phi(t)$. (zauważmy, że $\phi(0)=1$ pochodzi z normalizacji $\int_{-\infty}^{\infty} \phi(t) = 1$).

Inną metodą obliczenia funkcji skalującej jest rekursja na której. skoncentrujemy się i postaramy się ją wyjaśnić.

Zauważmy, że dwie skale w równaniu rozszerzającym to t i $2t$. Wartości funkcji skalującej w połowie liczb całkowitych są określone przez odpowiadające im liczby całkowite; wartości w ćwiartce liczb całkowitych przez odpowiadające im wartości połowy liczb całkowitych i tak dalej. Więc, jeżeli znajdziemy wartość $\phi(t)$ dla liczby całkowitej, korzystając z równania rozszerzonego możemy otrzymać wszystkie wartości $\phi(t)$ w podwójnych punktach $t=k/2^j$, gdzie $k, j=1, 2, \dots$. Jeśli jakieś rzeczywiste t może być aproksymowane przez podwójne punkty formuły $k/2^j$ dla danej dokładności, $\phi(t)$ w każdym punkcie t może w zasadzie, być obliczane dla każdej ustalonej dokładności tak długo, jak $\phi(t)$ jest ciągłe w t . W praktyce zazwyczaj obliczamy wartości $\phi(t)$ z $t=k\Delta$, gdzie $\Delta=2^{-j \max}$ i interpolujemy wartość pomiędzy.

Wartości funkcji skalującej $\phi(t)$ dla liczby całkowitej rozwiązuje problem wartości własnej. Z rozszerzenia równania

$$\phi(t) = \sum_{k=0}^n \sqrt{2} h_k \phi(2t - k) = \sum_{k=0}^n c_k \phi(2t - k),$$

mamy,

$$\phi(i) = \sqrt{2} \sum_k h_k \phi(2i - k) = \sum_k c_k \phi(2i - k) = \sum_j c_{2i-j} \phi(j) \quad (47)$$

kiedy t przybiera wartości liczby całkowitej i .

Jest to jest równoważne wartości własnej

$$\mathbf{C}\bar{\phi} = \bar{\phi} \quad (48)$$

gdzie jest macierzą \mathbf{C} ($n \times n$) posiadającą elementy $(\mathbf{C})_{ij} = c_{2i-j}$ oraz $\bar{\phi} = (\phi(0), \phi(1), \dots, \phi(n-1))'$ gdzie ' oznacza przesunięcie macierzy \mathbf{W} ,

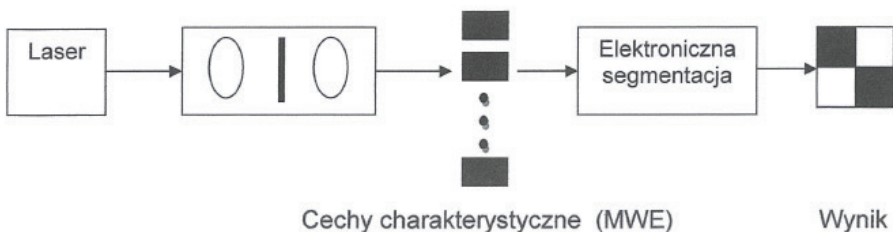
Zatem wartość funkcji skalującej ϕ w funkcji liczby całkowitej są dane przez wektor własny \mathbf{C} nawiązujący do wartości własnej równej 1. Następnie normalizujemy wektor własny $\sum_k \phi(k)=1$. co zagwarantuje nam, że całka z ϕ równa się 1. Z kolei możemy przystąpić do znalezienia wartości ϕ , jako połowę liczby całkowitej, ćwiartkę liczby całkowitej itd.

Funkcja skalująca generuje zbiór punktów $\{k/2^{j_{max}}, \phi(k/2^{j_{max}})\}$, gdzie k i j_{max} są liczbami całkowitymi oraz $k/2^{j_{max}}$ osiąga wartości wewnątrz nośnika ϕ . Wartości funkcji skalującej są osiągane przez znalezienie wektora własnego, który rozwiązuje problem wartości własnej z danym filtrem, następnie rekurencyjnie znajdujemy wartości pomiędzy aż do skali $2^{-j_{max}}$.

4. Optoelektroniczny system segmentacji tekstury obrazów pisma odręcznego

Rozwój teorii falek oferuje nowe metody do analizowania tekstur obrazów [11–15]. Zmiana rozmiaru okna zmienia się zgodnie z częstotliwością a dekompozycja falek ma optymalną wspólną rozdzielczość w przestrzennej i przestrzenno-częstotliwościowej dziedzinie. W drugim kroku algorytm grupowania (nadzorowany lub nienadzorowany) łączy cechy charakterystyczne tekstury i tworzy segmentację. Rysunek 1 ilustruje schemat ideowy optoelektronicznego (hybrydowego) systemu z filtrem multifalkowym do segmentowania tekstur opartym na transformacie falkowej. Multirozdzielczość jest niezbędna do dokładnej dekompozycji obrazu. W tym artykule, proponuje się ekstrakcję cech charakterystycznych z wykorzystaniem transformaty falkowej Gabora.

Rys. 1. Schemat ideowy optoelektronicznego (hybrydowego) systemu do segmentacji tekstur



Propozycja transformaty falkowej, która realizowana jest w czasie rzeczywistym jest ważna dla reprezentacji falkowej obrazu. Optyczna transformata falkowa (przy wykorzystaniu układu optycznego 4-f) jest atrakcyjna, ponieważ jej zaletami jest duża prędkość działania, szybka odpowiedź oraz równoległość obliczeń w porównaniu do procesu elektronicznego, którego działanie jest szczególnie czasochłonne i zawile. Chociaż multiplexery elementów filtra w czasie i przestrzeni zostały zbudowane do badania efektywności falek w obrazach 2-D, to dotychczas zaproponowano kilka metod i zrealizowano je aby rozwiązać problem segmentacji obrazu wejściowego. Zaproponowano

multiodniesienie dopasowanych filtrów z technikami holograficznymi, które miały niską wydajność dyfrakcji z powodu typu amplitudy oraz bliźniaczego obrazu. Bliźniacze obrazy zredukowano kanałem przepustowości w płaszczyźnie wyjściowej. Proponowano wytworzenie fazy dyfrakcji filtru multifalkowego z techniki kodowania dyfrakcji optycznej przez komputer, ale jest to kłopotliwe w fizycznej realizacji.

W niniejszym artykule proponowany jest microoptyczny filtr multifalkowy z funkcjami dekompozycji, filtrowania oraz zobrazowania do segmentacji obrazów pisma odręcznego. Poniżej zostanie zilustrowane zastosowanie microoptycznego filtra multifalkowego do analizy tekstury obrazów pisma.

4.1 Idea microoptycznego multifalkowego filtra

Transformata falkową $Wf(a, \theta, x_0, y_0)$ funkcji obrazu $f(x, y)$ w przestrzeni 2-D definiowana jest równaniem (7), gdzie falka córka, generowana jest z falki matki (8) przez rozciągnięcie ($\mathbf{a} = (a_x, a_y)$) i przesunięcie ($\mathbf{b} = (b_x, b_y)$):

Odpowiednio transformata falkowa $W(a, b)$ w dziedzinie częstotliwości (częstości) definiowana jest w postaci:

$$W(a, b) = (a_x a_y)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \{F(\xi, v) \cdot H^*(a_x \xi + a_y, v) \cdot \exp[i2\pi(b_x \xi + b_y v)]\} d\xi dv \quad (49)$$

Transformata falkowa w dziedzinie częstotliwości $W(a, b)$ jak widać z powyższego r-nia jest splotem pomiędzy widmem Fouriera $F(\xi, v) = \mathfrak{F}f(\xi, v)$ funkcji obrazu $f(x, y)$ oraz transformatą Fouriera $H_{a,b}(\xi, v) = \mathfrak{F}\psi(\xi, v)$ falek, rozumianych jako filtrów transformaty falkowej. Zatem wygląda to tak jakby obraz 2-D generował obraz 4-D. Transformatę falkową w przestrzeni częstotliwości proponuje się w tym artykule zaimplementować w postaci optycznych multifalkanałowych elementów jako filtrów górnoprzepustowych. Z powodu niezmienności właściwości optycznych, współrzędne w wiązaniu wyjściowym każdego kanału będą stale przesuwane o wektor \mathbf{b} i niezmiennie rozciągane o wektor \mathbf{a} , które wykorzystane zostaną w projekcie filtra transformaty falkowej ale z wektorem przesunięcia \mathbf{b} i zerowym wektorem rozciągania $\mathbf{a}=\mathbf{0}$. Aby uniknąć efektu krzyżowania falkowego oraz umożliwić rejestrowanie przez standardowy np. aparat fotograficzny, konieczne jest odpowiednie zobrazowanie w płaszczyźnie wyjściowej. Każdy element filtra jest swoistym transformatorem poprzez fazę klinową ze ściśle zdefiniowanymi klinami: pod ustalonymi kątami i ustalonymi kierunkami. W ten sposób każdy filtrowany obraz pisma ręcznego jest kierowany do pewnych wyseperowanych obszarów w płaszczyźnie wyjściowej. Pomijając odwrotną transformatę Fouriera soczewki i aby uprościć optyczny system, należy funkcję fazy soczewki dodawać oddzielnie do każdego filtra.

Optyczny filtr multifalkowy ($MWE(\xi, \nu)$) w dziedzinie częstotliwości można matematycznie opisać równaniem:

$$MWE(\xi, \nu) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} H(a_{x,m,n} \xi + a_{y,m,n} \nu) \cdot \exp\left\{i \frac{2\pi}{\lambda} \left[\alpha_{m,n} \xi + \beta_{m,n} \nu - \frac{1}{2f} (\xi^2 + \nu^2) \right]\right\} \quad (50)$$

gdzie: $\alpha_{m,n}$ i $\beta_{m,n}$ są odpowiednimi kierunkami cosinusów klinów, a f jest ogniskową soczewki. W związku z powyższym optyczny filtr multifalkowy (MWE) posiada funkcje dekompozycji, filtrowania i zobrazowania.

Typową falką do zastosowań może być *uproszczona falka Gabora* $\psi(\xi, \nu)$ z impulsową odpowiedzią:

$$\psi(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \cdot \cos(2\pi\xi_0 x) \quad (51)$$

Z powyższego widać, że fala harmoniczna modulowana jest funkcją Gauss'a, gdzie: ξ_0 jest częstotliwością na osi x a σ_x i σ_y są stałymi przestrzeni koperty Gauss'a na osi x i osi y .

Transformatę Fouriera *uproszczonej falki Gabora* (51) $H(\xi, \nu) = \mathfrak{F}\psi(\xi, \nu)$ można przedstawić w postaci:

$$H(\xi, \nu) = \psi(\xi, \nu) = \frac{1}{2} \exp\left\{2\pi^2\left[\left(\xi - \xi_0\right)^2 \sigma_x^2 + \nu^2 \sigma_x^2\right]\right\} + \frac{1}{2} \exp\left\{2\pi^2\left[\left(\xi + \xi_0\right)^2 \sigma_x^2 + \nu^2 \sigma_x^2\right]\right\} \quad (52)$$

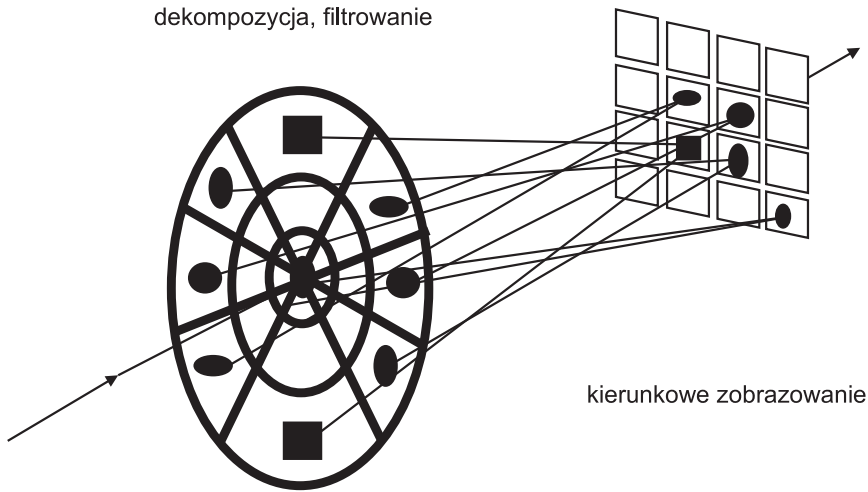
Granicami są dwie elipsy oddzielone o $2\xi_0$ i o stosunku osi dużej do małej σ_y/σ_x . Oczywiście jest, że dwuwymiarowy filtr Gabora w obydwu dziedzinach: przestrzennej jak i przestrzenno-częstotliwościowej jest funkcją typu funkcji Gauss'a. Skuteczne długości filtra w przestrzeni (szerokość pasma) i w dziedzinie przestrzenno-częstotliwościowej są powiązane, a to ma optymalnie wspólną rozdzielczość w obydwu dziedzinach. Filtry Gabora są przystosowane do każdego zakresu częstotliwości a kierunek przechodzi przez wybrane ξ_0 , σ_x i σ_y . Upraszczając zrozumienie optyki, filtry Gabora są kwantyzowane w formie dziur lub par dziur Gauss'a. Szum kwantyzacji może być zignorowany w segmentacji obrazu. W ten sposób optyczny filtr multifalkowy ($MWE(\xi, \nu)$) matematycznie reprezentowany wzorem (50) może być tylko filtrem fazowym.

Kanały o rozmiarze 4×4 są wystarczające dla celów praktycznych do rozszerzenia falki Gabora. Filtr multifalkowy w proponowanym eksperymencie reprezentowany jest przez szesnaście falek Gabora filtrowanych czterema skalami rozdzielonymi przez czynniki dwu i cztero kierunkowe rozłożone co 45 stopni. Realizowane funkcje filtra $MWE(\xi, \nu)$ są zilustrowane na rys. 2.

Rozdzielenie filtrowanych obrazów daje w przestrzeni macierz 4 na 4 , która jest wygodna do rejestracji typowym aparatem fotograficznym. Całkowity

rozkład z transformatą falkową może być zrealizowany równocześnie przez system optyczny z $MWE(\xi, \nu)$.

Rys. 2. Ilustracja realizacji funkcji optycznego filtra multifalkowego



4.2 Optoelektroniczny proces segmentacji tekstur pisma odręcznego

Propozycje laboratoryjnego układu optoelektronicznego systemu segmentacji tekstury zilustrowano na rys. 3. Bazuje on na laserze, dwóch soczewkach Fourier'a (L_1, L_2) oraz optycznym filtrze multifalkowym ($MWE(\xi, \nu)$). Faza filtra jest kodowana binarnie jako czterokanałowa z dwoma generowanymi komputerowo znakami - rys. 2 (dwa czarne prostokąty). Element ze znakami, może być wytwarzany techniką np. trawienia jonowego.

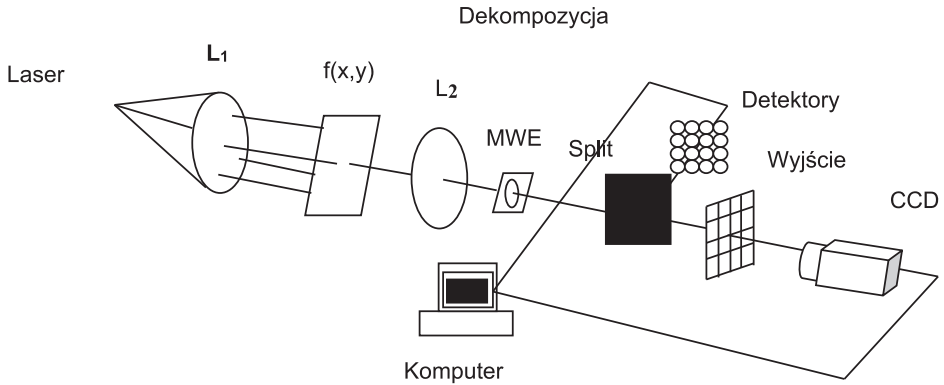
Obraz wejściowy ($f(x, y)$) jest dekomponowany (rozkładany) na serie filtrowanych obrazów po optycznym filtrowaniu falką Gabora. Jeżeli liczba pikseli w każdym filtrowanym obrazie jest taka sama jak w obrazie wejściowym, to każdy piksel w wyjściowym obrazie jest przedstawiany przez wektor cech charakterystycznych $O(i, j)$ ³³

$$O(i, j) = \{E_k(i, j), \quad k = 1, 2, \dots, N \times N\}, \quad i, j = 1, 2, \dots, M \quad (53)$$

gdzie $N \times N$ jest liczba kanałów, $M \times M$ jest sumaryczną liczbą pikseli w filtrowanym obrazie, natomiast $E_k(i, j)$ są wartościami szarości zlokalizowanego piksela w punkcie (i, j) w kanale k .

33 Wenyi Feng, Yingbai Yan, Geaogui Hang, Guofain Jin, *Micro-optical multiwavelet element for hybrid texture segmentation processor*, Optical Engineering, 37, 1, 1998.

Rys. 3. Optoelektroniczny system segmentacji tekstur z optycznym filtrem multifal-kowym ($MWE(\xi, \nu)$)



Segmentacja obrazu wejściowego jest sprowadzona do problemu klasyfikacji wektorów cech charakterystycznych. Sieć neuronowa albo rozmyty algorytm łączący wartości *c-średnie* jest proponowany przy realizacji segmentacji. Obraz wejściowy $f(x,y)$ jest dekomponowany (rozkładany) na 4×4 filtrowane obrazy z różnymi częstotliwościami i kierunkami obszarów. W związku z powyższym otrzymujemy szesnaście filtrowanych obrazów jako cechy charakterystyczne do segmentacji. Tam gdzie obrazy są filtrowane z małą ilości informacji; to obrazy te są nieistotne dla segmentacji i mogą być pomijane zmniejszając tym samym wymagania obliczeniowe. Jak ilustruje rys. 3, 4×4 matryca detektorów, umieszczona jest w płaszczyźnie ogniskowej zintegrowanego systemu, która ustala kandydatów do odrzucenia. Punkty centralne (4×4) w płaszczyźnie ogniskowej (pokazane na rys. 2) są rejestrowane przez detektory. Wartości energii są przenoszone do komputera i normalizowane. Niska wartość energii jest równoważna małej ilości informacji w filtrowanym obrazie. Próg T może być ustawiany ręcznie. Filtrowane obrazy, których wartość energii jest mniejsza niż T - są odrzucane. W proponowanym układzie w środku $MWE(\xi, \nu)$ cztery kanały obrazów mogą być wyeliminowane poprzez wygenerowanie czarnej maski. Dwanaście filtrowanych obrazów w postaci tablicy może być wykorzystane do segmentacji.

Wyniki badań

Wiele uwagi poświęcono badaniom eksperymentalnym rozpoznawania pisma odręcznego w przestrzeni obrazowej^{34,35,36,37,38,39,40,41,42,43}, natomiast niewiele w przestrzeni dyfrakcyjnej. Autorzy niniejszej pracy zasymulowali w komputerze nowoczesny optoelektroniczny układ rozpoznawania obrazów w przestrzeni fourierowskiej (rys. 4).

Holograficzny detektor pierścieniowo-klinowy (HDPK), pełniący funkcje ekstraktora cech, jest okrągłym elementem złożonym z półpierścieni i klinów (zwanymi obszarami) pokrytych prostokątną siatką dyfrakcyjną o różnych częstościach przestrzennych i orientacji. Jeśli taki element (HDPK) zostanie umieszczony w tylnej płaszczyźnie ogniskowej soczewki nazywanej płaszczyzną Fourier'a, to realizowane będzie próbkowanie obrazów dyfrakcyjnych Fraunhofer'a - próbkowanie mocy widmowej, która jest zogniskowana na pierścieniach i klinach. Dzięki temu, każdy obszar ekstraktora cech (HDPK) generuje dokładnie jedną cechę charakterystyczną obrazu, obliczaną w postaci całki z widma mocy nad danym obszarem. Jeśli dyfrakcyjny obraz Fraunhofer'a przykładowo jest wyrażony we współrzędnych biegunowych (ω, ϕ) , to cechy odpowiadające pierścieniom R_i i cechy odpowiadające klinom W_j są rozumiane jako:

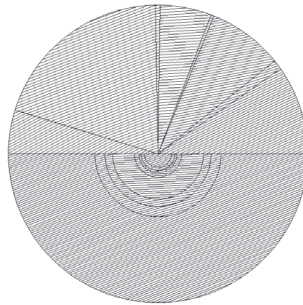
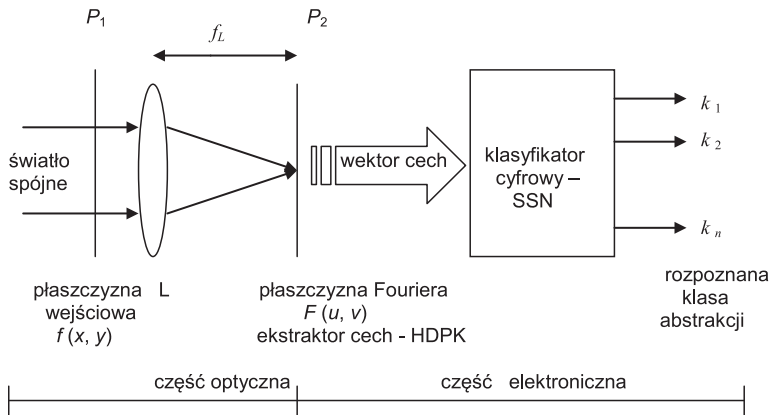
$$R_i = \int_0^{\pi} \left(\int_{r_i}^{r_{i+1}} F^2(\omega, \phi) d\omega \right) d\phi, \quad W_j = \int_0^R \left(\int_{\phi_j}^{\phi_{j+1}} F^2(\omega, \phi) d\phi \right) d\omega \quad (54)$$

gdzie: R jest cechą charakterystyczną HDPK, r_i ($i = 1, \dots, N_r$) są promieniami pierścieni i ϕ_j ($j = 1, \dots, N_w$) są kątami klinów, podczas gdy N_r i N_w są, odpowiednio, całkowitą ilością pierścieni i klinów. Z dobrze znanych właściwości fizycznych transformaty Fouriera i charakterystycznych kształtów obszarów HDPK jest oczywiste, iż ekstrahowany wektor cech

-
- 34 A. Smith, Hong-Phuong Tran, Eric Dimla, *Offline Handwritten Text Recognition using Convolutional Recurrent Neural Network* International Conference on Advanced Computing and Applications 2019 (ACOMP).
 - 35 R. Reeve Ingle, Yasuhisa Fujii, Thomas Deselaers, Jonathan Baccash, Ashok C. Popat, *A Scalable Handwritten Text Recognition System*, International Conference on Document Analysis and Recognition 2019 (ICDAR).
 - 36 A. Gupta, M. Srivastava, C. Mahanta, *Offline Handwritten Character Recognition Using Neural Network* International Conference on Computer Application and Industrial Electronics 2011 (ICCAIE-2011).
 - 37 J. Yang, P. Ren, X. Kong, *Handwriting Text Recognition Based on Faster R-CNN* IEEE xplore 2020.
 - 38 S. España-Boquera, Maria Jose Castro-Bleda, Jorge Gorbe-Moya, and Francisco Zamora-Martinez, *Improving Offline Handwritten Text Recognition with Hybrid HMM/ANN Models* IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 33, NO. 4, APRIL 2011.
 - 39 R. Vaidya, D. Trivedi, S. Satra, *Handwritten Character Recognition Using Deep-Learning* 2nd International Conference on Inventive Communication and Computational Technologies 2018 (ICICCT-2018).
 - 40 B. Balci, D. Saadati, D. Shiferaw, *Handwritten Text Recognition using Deep Learning*.
 - 41 R. Vaidya, D. Trivedi, S. Satra, *Handwritten Character Recognition Using Deep-Learning* 2nd International Conference on Inventive Communication and Computational Technologies 2018 (ICICCT-2018).
 - 42 V. Pham, T. Bluche, Christopher Kermorvant, and Jérôme Louradour, *Dropout improves Recurrent Neural Networks for Handwriting Recognition* 14th International Conference on Frontiers in Handwriting Recognition 2014.
 - 43 Sushmitha, Sushmitha B Poojary, Varshitha, Vidya K C, Shilpa, *Handwritten text recognition system*, International Journal of Creative Research Thoughts (IJCRT), Volume 9, Issue 7 July 2021.

charakterystycznych zawiera niezmiennie informacje względem przesunięcia i rotacji ale zależne od skalowania (w odniesieniu do pierścieni) oraz niezmiennie informacje względem przesunięcia i skali ale zależne od rotacji (w odniesieniu do klinów).

Rys. 4. Schemat optoelektronicznego układu do automatycznego rozpoznawania pisma odręcznego z ekstraktorem cech charakterystycznym w postaci komputerowo wgenerowanego hologramu (HDPK) oraz klasyfikatora w postaci sztucznej sieci neuronowej (SSN)



HDPK - ekstraktor cech

Wiązki światła koherentnego w rezultacie dyfrakcji w obszarach komputerowo generowanego hologramu kierowane są na detektor w postaci matrycy fotodetektorów gdzie realizowana jest konwersja sygnałów optycznych na elektryczne. Wysoka sprawność takiego optycznego ekstraktora cech jest gwarantowana jako wynik optymalizacji elementu dyfrakcyjnego w komputerze symulującym system rozpoznawania obrazów.

Rozpoznawanie (rozdzielanie) pisma odręcznego zostało przebadane eksperymentalnie jako przykład ważnej klasy problemów w których pragnę-

liśmy zweryfikować proponowaną metodę z zachowaniem dużej jej trafności. Przykładowo dwie osoby przepisały po dwie strony tekstu. Jedna strona pisma posłużyła jako próbka ucząca dla sztucznej sieci neuronowej jako klasyfikatora a druga strona pisma została użyta do testowania systemu. Rozpoznawano jedynie niewielkie fragmenty tekstów (załączone poniżej). Celem przeprowadzonego eksperymentu była ekstrakcja i identyfikacja cech charakterystycznych pisma odręcznego oraz identyfikacja fragmentów tekstów.

Pismo 1

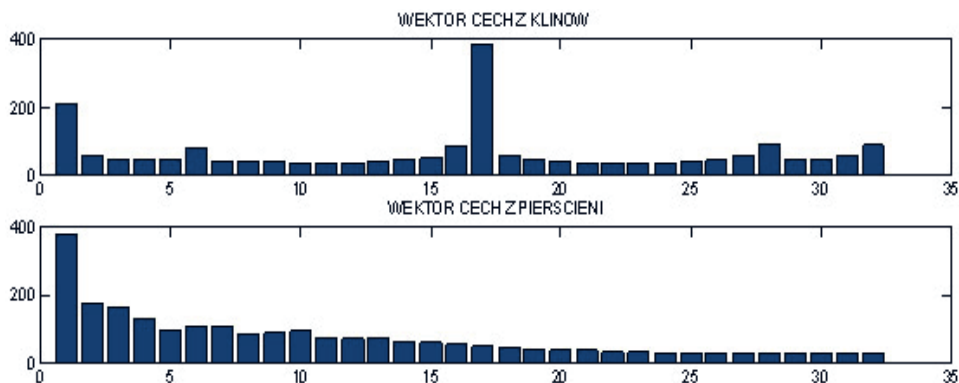
Ludzie są najważniejszym zasobem w każdym badaniu kryminalistycznym i aby umożliwić personelowi efektywną i wydajną pracę, każda osoba zaangażowana w ten proces musi rozumieć charakter zadań i posiadać odpowiednie cechy osobowe wymagane do ich wykonania. Przyjmuje się, że rekrutowanie ekspertów kryminalistycznych pisma ręcznego jest zgodne z wymaganiami danej jednostki organizacyjnej i może obejmować względy prawne, kwalifikacje akademickie lub doświadczenie zawodowe. W związku z tym zakłada się, że eksperci kryminalistyczni pisma ręcznego będą mieć szeroki zakres doświadczenia, przeszkolenia i wiedzy.

Anna

Pismo 2

Ludzie są najważniejszym zasobem w każdym badaniu kryminalistycznym i aby umożliwić personelowi efektywną i wydajną pracę, każda osoba zaangażowana w ten proces musi rozumieć charakter zadań i posiadać odpowiednie cechy osobowe wymagane do ich wykonania. Przyjmuje się, że rekrutowanie ekspertów kryminalistycznych pisma ręcznego jest zgodne z wymaganiami danej jednostki organizacyjnej i może obejmować względy prawne, kwalifikacje akademickie lub doświadczenie zawodowe. W związku z tym zakłada się, że eksperci kryminalistyczni pisma ręcznego będą mieć szeroki zakres doświadczenia, przeszkolenia i wiedzy.

Klana



Przykład otrzymanego wektora cech

Interesującą właściwość, którą ujawnił przeprowadzony eksperyment jest możliwość rejestracji wektora cech charakterystycznych, kojarzonych zarówno z pierścieniami jak i z klinami. Ponieważ rozpoznawanie pisma odręcznego jest oczywistą sytuacją, w której kąтова orientacja odgrywa ważną rolę zatem w eksperymencie poświęcono uwagę głównie informacji zawartej w części klinowej ekstraktora cech. Klinowa intensywność danych z obrazów pisma odręcznego została wykorzystana, w opracowaniu oprogramowania, do rozpoznawania obrazów z informacji otrzymanych z ekstraktora cech charakterystycznych w postaci komputerowo generowanego hologramu.

Autorzy proponują algorytm sumy różnicy uporządkowania pierścieniowego dyfrakcyjnych obrazów pisma do oceny podobieństwa obrazów.

Algorytm różnicy uporządkowania pierścieniowego obrazów dyfrakcyjnych

Autorzy w przeprowadzonym eksperymencie zastosowali prosty algorytm różnicy uporządkowania pierścieniowego do rozpoznawania (porównywania podobieństwa) obrazów dyfrakcyjnych dwu tekstów pisma odręcznego. Algorytm ten porównuje wartości wektorów cech charakterystycznych obrazów rzeczywistych otrzymanych z klinów z wektorem cech obrazu wzorcowego. Wzorcowy obraz charakteryzuje się niepowtarzalną wartością i odpowiednimi proporcjami danych w wektorze cech pochodzących z klinów.

Obraz tekstu potraktowany został jak obiekt rzeczywisty, należący do tej samej klasy co tekst wzorcowy, który posiada podobną wartość wektora cech jednak różniący się jego ułożeniem, poziomem poszczególnych składowych oraz zakłóceniami związanymi z akwizycją obrazu.

Zdefiniujmy wektor cech charakterystycznych obrazu tekstu wzorcowego jako:

$$r^1 = (r_1^1, r_1^1, \dots, r_n^1) \quad (55)$$

oraz wektor cech charakterystycznych obrazu rzeczywistego (amplitudy sygnałów dla obrazu dyfrakcyjnego rozpoznawanego) jako:

$$r^2 = (r_1^2, r_1^2, \dots, r_n^2) \quad (56)$$

Następnie dokonajmy zamiany uszeregowania amplitudowego. Przyjmijmy kolejność uszeregowania amplitudowego od największej do najmniejszej amplitudy dla obrazu dyfrakcyjnego wzorcowego:

$$s^1 = (r_1^1, r_1^1, \dots, r_n^1) \quad (57)$$

Analogicznie dokonajmy zamiany kolejność uszeregowania amplitudowego od największej do najmniejszej amplitudy dla rozpoznawanego obrazu dyfrakcyjnego:

$$s^2 = (r_1^2, r_1^2, \dots, r_n^2) \quad (58)$$

Zdefiniujmy następnie miarę podobieństwa dwóch obrazów dyfrakcyjnych jako:

$$S(s^1, s^2) = \sum_{w=1}^n |r_w^1 - r_w^2| \quad (59)$$

gdzie:

r - klin;

1, 2 - indeks obiektu wzorcowego (1) i rzeczywistego (2);

w - indeks klina;

Σ - suma różnic uporządkowania w klinach ($w=1, \dots, n$)

Minimalna odległość pomiędzy wektorem wzorcowym i badanym to:

$$S(s^1, s^2) = \min S(s^1 \cdot s^2) \quad (60)$$

Aby znaleźć minimalną odległość pomiędzy wektorami należy obracać jeden z wektorów co pomiar o jeden klin. Można dokonywać tego zmieniając indeksy wektora cech obrazu rzeczywistego. Każdy indeks oznaczający kolejną pozycję wektora należy zmniejszyć o jeden:

$$x_1 = (m_1^1, m_2^1, \dots, m_n^1) \tag{61}$$

$$m_k^1 = \begin{cases} k = n \text{ to } k = k - n + 1 \\ w \text{ przeciwnym razie } k = k - 1 \end{cases} \tag{62}$$

wynosic będzie $(180/n)^\circ$

gdzie: k - wartość indeksu opisującego numer elementu w wektorze cech charakterystycznych.

Na rys. 5 i rys. 6 zilustrowano algorytm różnicy uporządkowania pierścieniowego obrazów w płaszczyźnie fourierowskiej

Rys. 5 Wartości wzorcowego i rzeczywistego wektora cech charakterystycznych uzyskane z detektora klinowego

Przykładowy detektor klinowy	Klin	Wzorzec		Obraz rzeczywisty	
		Wartość	Kolejność wg wartości	Wartość	Kolejność wg wartości
	1	400	3	2000	1
	2	50	6	150	5
	3	1000	1	400	4
	4	75	5	1000	2
	5	200	4	800	3
	6	500	2	100	6

Rys. 6 Schemat algorytmu automatycznego rozpoznawania obrazów pisma odręcznego z wykorzystaniem detektora klinowego (HDPK)

ALGORYTM	I	II	Δ	Rotacja wektora cech obrazu rzeczywistego o 3 dwa kliny	I	II	Δ
	3	1	2		3	3	0
	6	5	1		6	6	0
	1	4	3		1	1	0
	5	2	3		5	5	0
	4	3	1		4	4	0
	2	6	4	2	2	0	Duże podobieństwo
		Σ	14		Σ	0	

Algorytm klinowo-różnicowej sumy eksperymentalnie udowodnił wysoką jakość w przeprowadzonym eksperymencie i może być z powodzeniem stosowany do obrazów o złożonej strukturze wewnętrznej (w kryminalistyce w rozpoznawaniu np. pisma, jak również odcisków palców, lub twarzy). Algorytm jest wygodny i skuteczny, ponieważ porównuje w prosty sposób wektory cech charakterystycznych. Wektory te zawierają dużo mniejszą liczbę danych, niż obrazy dyfrakcyjne, więc metoda ta jest bardzo szybka obliczeniowo.

Dwie natychmiastowe korzyści płynące z zastosowania powyższej metody rozpoznawania optoelektronicznego to:

1. proste zgrubne próbkowanie płaszczyzny transformaty fourierowskiej znakomicie przyspiesza obliczenia (redukuje czas obliczeń)
2. nieliniowe obliczenia są tu standardem. Klinowo - pierścieniowa suma różnicowa jest algorytmem, który udowodnił swoją eksperymentalną potęgę obliczeniową (zauważmy, że jest silnie nieliniowy, w przybliżeniu jest identyczny do wzajemnej korelacji dwóch klinów ekstraktora).

Proponowany algorytm do pomiaru podobieństwa obrazów dyfrakcyjnych pisma odręcznego jest niewrażliwy na wpływ: orientacji kątowej obrazu wejściowego, jasności obrazu oraz zakłóceń powstałych w czasie akwizycji obrazu.

Podsumowanie

W pracy zaprezentowano ideę dekompozycji przestrzeni Gabora oraz jej niezmienniczą cechę charakterystyczną. Przedstawiono technikę generacji ortogonalnych funkcji skalujących i falek w dziedzinie częstotliwości. Zaproponowano optoelektroniczny (hybrydowy) układ laboratoryjny, bazujący na transformacie falkowej Gabora do przetwarzania informacji o teksturze obrazów pisma odręcznego. Łączy on równoległość i szybkość optyki oraz dojrzałość elektroniki. Optyczny filtr multifalkowy, posiada funkcje dekompozycji, filtrowania i kierunkowej segmentacji. Proponowany układ laboratoryjny jest prosty i stały względem elementu filtrującego oraz może być modelowany cyfrowo. Propozycja optycznego filtra multifalkowego jest innowacyjna dla hybrydowego (optoelektronicznego) systemu przetwarzania informacji dla celów kryminalistyki. Zaproponowano prosty algorytm rozpoznawania obrazów w przestrzeni dyfrakcyjnej dla celów kryminalistyki.

Bibliografia

1. Andrysiak T., Chora M., *Image retrieval based on hierarchical gabor filters*, Int. J. Appl. Math. Comput. Sci., **15**, 4, 2005.
2. Arivazhagan S., Ganesan L., *Texture segmentation using wavelet transform*, Pattern Recognition Letters **24**, 2003.
3. Balci B., Saadati D., Shiferaw D., *Handwritten Text Recognition using Deep Learning*.
4. Bovik A. C., Clark M. and Geisler W.S. r: *Multichannel Texture Analysis Using Localized Spatial Filters*, IEEE Trans. PAMI, **12**, 1, 1990.
5. Canny J. F., *A computational approach to edge detection*. IEEE Trans. Pattern Anal. Machine Intell. **8**, 6, 1986.
6. Clausi D. A., Jernigan M. E., *Designing Gabor filters for optimal texture separability*, Pattern Recognition **33**, 1835–1849, 2000.
7. Daubechies, *The Wavelet Transform, Time-Frequency Localization and Signal Analysis*, IEEE Trans. Information Theory, **36**, 5, 1990. Grigorescu S. E., Petkov N., Kruižinga P., *Comparison of texture features based on Gabor filters*, IEEE Transactions on Image Processing, **11**, 10, 2002.
8. Daugman J.G., *Uncertainty relation for resolution In space, spatial frequency, and orientation optimized by two-dimensional visual cortex flters*, J. Opt. Soc. Am. A **2**, 7, 1985.
9. Daugman J. G., *Two-Dimensional Spectral Analysis of Cortical Receptive Field Profile*, Vision Research, **20**, 1980.
10. Daugman J. G., *Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression*, IEEE Trans. Acoustics, Speech, and Signal Processing, **36**, 7, 1988.
11. Duffin R. J. and Schaeffer A. C., *A Class of Nonharmonic Fourier Series*, Trans. Am. Math. Soc., **72**, 1952.
12. España-Boquera S., Jose Castro-Bleda M., Gorbe-Moya J., and Zamora-Martinez F., *Improving Offline Handwritten Text Recognition with Hybrid HMM/ANN Models* IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 33, NO. 4, APRIL 2011.
13. Gupta A., Srivastava M., ChitralakhaMahanta, *Offline Handwritten Character Recognition Using Neural Network* International Conference on Computer Application and Industrial Electronics 2011 (ICCAIE-2011).
14. Heintz R., Schäfer M. G., *Local invariant object localization based on Gabor feature space*.
15. Hołyst B. Kryminalistyka, Warszawa 2023. Jain A. K., Farrokhnia F., *Unsupervised texture segmentation using Gabor filters*, Pattern Recognition, **16**, 1167–1186, 1991.
16. Jain A. K., Farrokhnia F, *Unsupervised Texture Segmentation Using Gabor Filters*.
17. Jain A. K., Vailaya A., *Image retrieval using color and shape*, Pattern Recognition **29**, 8, 1996.
18. Kulikowski J.J. and P. O. Bishop: *Fourier Analysis and Spatial Representation in the Visual Cortex*, Experientia, **37**, 1981.
19. Kyrki V., Kamarainen J-K, Kälviäinen H., *Simple Gabor feature space for invariant object recognition*, Pattern Recognition Letters, **25**, 3, 2004.
20. Lee T. S., *Image representation using 2D Gabor wavelets*, IEEE Transactions on pattern analysis and machine intelligence, **18**, 10, October 1996.
21. Niedziela T., Kopania Ł., *Metoda falkowa klasyfikacji obiektów trójwymiarowych z dwuwymiarowych obrazów*, Informatyka XXI wieku – wyzwania i dylematy, Monografia 252, 2020.

22. Niedziela T., *The method of automatic identification of motor vehicle users*, Journal of civil engineering and transport transEngin, **2**, 2, 2020.
23. Niedziela T., *Wavelet method of edge detection in images with high-noise level*, Journal of civil engineering and transport transEngin, **3**, 2, 2020.
24. Niedziela T., Kopania Ł., *Zmodyfikowany korelator znormalizowanej transformaty Fouriera*, Informatyka XXI wieku – wyzwania i dylematy, Monografia **252**, 2020.
25. Niedziela T., Kopania Ł., *Metoda falkowa klasyfikacji obiektów trójwymiarowych z dwuwymiarowych obrazów*, Informatyka XXI wieku – wyzwania i dylematy, Monografia **252**, 2020.
26. Naftali A., *Behavior factors in handwriting identification*, The Journal of criminal law, criminology and police science Vol 56, No. 4, 528–539, 1965
27. Mallat S., *Multifrequency channel decompositions of images and wavelet models*, IEEE Trans. Acoustics, Speech. Signal Processing **37**, 12, 1989.
28. Mallat S., *A theory for multiresolution signal decomposition: The wavelet representation*. IEEE Trans. Pattern Anal. Machine Intell. **11**, 12, 1989.
29. Mallat S., *Multiresolution approximation and wavelet orthonormal bases of $L(R^2)$* , Trans. Amer. Math. Soc. **3**, 15, 1989.
30. Mallat S., Zhong S., *Characterization of signals from multi-scale edges*. IEEE Trans. Pattern Anal Machine Intell. **14**, 7, 1992.
31. Mallat S. and Hwang W. L., *Singularity detection and processing with wavelets*, IEEE Trans. Inf. Theory **38**, 1992.
32. Marmol U., Lenda G., *Filtry teksturalne w procesie automatycznej klasyfikacji obiektów*, Archiwum Fotogrametrii, Kartografii i Teledetekcji, **21**, 2010.
33. Perry A. and Lowe D. G., *Segmentation of Textured Images*, Proc. CVPR, 1989.
34. Pham V., Bluche T., Kermorvant C., and Louradour J., *Dropout improves Recurrent Neural Networks for Handwriting Recognition* 14th International Conference on Frontiers in Handwriting Recognition 2014.
35. Poojary S. B., Varshitha, Vidya K C, Shilpa, *Handwritten text recognition system*, International Journal of Creative Research Thoughts (IJCRT), Volume 9, Issue 7 July 2021.
36. Reeve Ingle R., Fujii Y., Deselaers T., Baccash J., Ashok C. Papat, *A Scalable Handwritten Text Recognition System*, International Conference on Document Analysis and Recognition 2019 (ICDAR).
37. Smith A., Hong-Phuong T., Dimla E., *Offline Handwritten Text Recognition using Convolutional Recurrent Neural Network* International Conference on Advanced Computing and Applications 2019 (ACOMP).
38. Tumer M. R., *Texture Discrimination by Gabor Functions*, Biol. Cybern., **55**, 1986.
39. Wenyi Feng, Yingbai Yan, Geogui Hang, Guofain Jin, *Micro-optical multiwavelet element for hybrid texture segmentation processor*, Optical Engineering, **37**, 1, 1998.
40. Webster M. A. Valois De R. L., *Relationship Between Spatial Frequency and Orientation Tuning of Striate Cortex Cells*, J. Optical Soc. Am., **A2**, 7, July 1985.
41. Webster M. A. and Valois De R. L., *Relationship Between Spatial Frequency and Orientation Tuning of Striate Cortex Cells*, J. Optical Soc. Am., **A2**, 7, July 1985.
42. Yang J., Ren P., Kong X., *Handwriting Text Recognition Based on Faster R-CNN* IEEE xplore 2020.

43. Vaidya R., Trivedi D., Satra S., *Handwritten Character Recognition Using Deep-Learning* 2nd International Conference on Inventive Communication and Computational Technologies 2018 (ICICCT-2018).
44. Vaidya R., Trivedi D., Satra S., *Handwritten Character Recognition Using Deep-Learning* 2nd International Conference on Inventive Communication and Computational Technologies 2018 (ICICCT-2018).