

DOI: 10.4467/29567610PIB.26.007.23479

prof. dr hab. inż. Tadeusz Niedziela

Wyższa Szkoła Kształcenia Zawodowego
tadeuszniedziela@poczta.onet.pl

prof. dr hab. Brunon Hołyst

Wyższa Szkoła Kształcenia Zawodowego
holyst.brunon@gmail.com

METODA FALKOWA ROZPOZNAWANIA OBRAZÓW TWARZY

A WAVELET METHOD FOR FACIAL IMAGE RECOGNITION

Streszczenie

Artykuł przedstawia metodę falkową rozpoznawania twarzy, z obrazów zakresu widzialnego dla celów automatycznej identyfikacji sprawców przestępstw. Wykorzystanie metrycznych falkowych cechy współwystąpień poziomów szarości w prezentowanej implementacji może być odpowiednie do celów identyfikacji. Algorytm zastosowany w tej metodzie wykorzystuje nowoczesną technologię do automatycznego rozróżnienia twarzy zlokalizowanych w zakłóconym środowisku, poprzez wyznaczenie nie zachodzących na siebie bloków obrazu, na których jest realizowane porównanie blok po bloku współwystąpień falkowych odcieni szarości. Metoda może znaleźć zastosowanie w identyfikacji twarzy osób poszukiwanych, zaginionych i przestępców. W artykule przedstawiono optoelektroniczny system identyfikacji twarzy bazujący na próbkowaniu obrazów dyfrakcyjnych i identyfikacji przez sztuczną sieć neuronową.

Słowa kluczowe: wykrywanie twarzy, dyskretna transformata falkowa, tekstura, falkowe cechy współwystąpień poziomów szarości, sztuczna sieć neuronowa

Abstract

The article presents a wavelet-based method of face recognition from visible images for the purpose of automatic identification of criminals. The use of wavelet metrics of gray level co-occurrence features in the presented implementation may be suitable for identification purposes. The algorithm used in this method employs modern technology to automatically distinguish faces located in a disturbed environment by determining non-overlapping image blocks on which a block-by-block comparison of wave co-occurrences of gray tones is performed. The method can be used to identify wanted persons, missing persons, and criminals. The article presents an optoelectronic face identification system based on sampling diffraction images and identification by an artificial neural network.

Keywords: face detection, discrete wavelet transform, texture, wavelet features of gray level co-occurrences, artificial neural network

Wprowadzenie

Rozpoznawanie obrazów twarzy stało się popularnym obszarem badań w dziedzinie sztucznej inteligencji. Zazwyczaj jest stosowane w systemach bezpieczeństwa sieciowego oraz w systemach kontroli dostępu, jak również przydatne jest w wielu obszarach przetwarzania informacji multimedialnych. Jednym z jego zastosowań jest identyfikacja twarzy przestępcy. Rejestr karny zazwyczaj zawiera informacje o konkretnej osobie wraz ze zdjęciem. Aby zidentyfikować przestępcę, potrzebujemy pewnej informacji dotyczącej konkretnej osoby lub osób, które są podawane przez naocznych świadków. W większości przypadków jakość i rozdzielczość zarejestrowanych segmentów obrazu jest słaba i trudno jest zidentyfikować twarz.

Identyfikacja twarzy to technika, która jest głównie wykorzystywana do rozpoznawania przestępców. Na podstawie wskazówek od naocznych świadków tworzony jest wizerunek przestępcy, a następnie porównywany jest z obrazami, które zgromadzone są w bazie danych. Aby zidentyfikować przestępcę, musimy mieć informacje, które zazwyczaj zawierają: imię, wiek, lokalizację, poprzednie przestępstwo, płeć, zdjęcie itp. Biorąc pod uwagę posiadane nieruchome lub ruchome obrazy stosowana jest odpowiednia technika identyfikacji jednego lub więcej segmentowanych elementów ekstrahowanych ze sceny, które można zidentyfikować i dopasować. Słowo obraz jest rozumiane tu jako dokładna lub analogiczna reprezentacja osoby lub rzeczy. Obraz monochromatyczny, taki jak czarno-biały, jest reprezentowany jako dwuwymiarowa funkcja natężenia światła $f(x,y)$, gdzie x i y oznaczają współrzędne przestrzenne. Obraz cyfrowy to obraz $f(x,y)$, który został zdigitalizowany zarówno pod względem współrzędnych przestrzennych, jak i jasności. Elementy takiej cyfrowej tablicy nazywane są elementami obrazu, lub pikselami.

Technologie biometryczne¹ zostały rozwinięte jako skuteczne rozwiązanie realizacji bezpiecznej identyfikacji i weryfikacji osobistej. Potrzeba wysoce bezpiecznych technologii identyfikacji i weryfikacji osobistej staje się oczywista w miarę wzrostu poziomu naruszeń bezpieczeństwa i oszustw transakcyjnych. Coraz częstsze stosowanie technologii biometrycznych w aplikacjach o wysokim poziomie bezpieczeństwa i nie tylko, stworzyło wymagania wysoce niezawodnych systemów rozpoznawania twarzy. System rozpoznawania

¹ "Bureau of Justice Statistics, U. S. Department of Justice", April 1990, pp. 43-66; SEARCH Group, Legal and Policy Issues Relating to Biometric Identification Technologies.

twarzy służy do weryfikacji tożsamości osoby poprzez dopasowanie danej twarzy do bazy znanych danych. Stał się on alternatywą dla tradycyjnych metod identyfikacji i uwierzytelniania, takich jak: używanie kluczy, kart identyfikacyjnych i haseł. Rozpoznawanie twarzy obejmuje komputerowe rozpoznawanie tożsamości osobistej na podstawie cech geometrycznych lub statystycznych, które są pochodne z obrazów twarzy². Chociaż człowiek może łatwo wykrywać i identyfikować twarze w scenie, to zbudowanie zautomatyzowanego systemu jest nadal kłopotliwe. Technologia rozpoznawania twarzy może być stosowana w wielu obszarach, np.: w kontroli dostępu do komputerów, nadzorze lotniskowym, prywatnym nadzorze, kryminalistyce, identyfikacji i bezpieczeństwie transakcji w bankomacie. Ponadto, systemy rozpoznawania twarzy zmierzają w kierunku inteligentnego środowiska nowej generacji, w którym komputery są projektowane tak, aby wchodzić w interakcje bardziej niż ludzie. W ostatnich latach poczyniono znaczne postępy w dziedzinie rozpoznawania twarzy dzięki rozwojowi wielu przydatnych technik. Postęp w technologii komputerowej ułatwił rozwój modułów wizji działających w czasie rzeczywistym, które w ostatnich latach wchodzą w interakcje z ludźmi. Przykładów jest wiele, szczególnie w biometrii i interakcji człowiek-komputer, ponieważ informacje zawarte w twarzach muszą zostać przeanalizowane, aby systemy mogły odpowiednio zareagować³. W przypadku systemów biometrycznych, które wykorzystują twarze jako nieinwazyjne moduły wejściowe, konieczne jest zlokalizowanie twarzy w scenie, zanim będzie można zastosować jakikolwiek algorytm rozpoznawania. Inteligentny interfejs użytkownika oparty na wizji powinien być w stanie określić, na czym skupia się użytkownik (tj. gdzie użytkownik patrzy), aby odpowiednio zareagować⁴. Aby dokładnie wykrywać rysy twarzy, twarze muszą zostać najpierw zlokalizowane i zarejestrowane, by ułatwić dalsze przetwarzanie. Oczywiście jest, że wykrywanie twarzy odgrywa ważną i krytyczną rolę dla osiągnięcia wysokiej skuteczności wszelkich systemów przetwarzania. Problem automatycznego wykrywania twarzy jest trudny, ponieważ musi uwzględniać wszystkie możliwe odmiany wyglądu spowodowane zmianą: oświetlenia, cech twarzy, okluzji itp. Ponadto musi wykrywać twarze, które pojawiają się w różnej skali, pozach, z uwzględnieniem obrotów w płaszczyźnie obserwacji. Często rozmiar obrazu jest zbyt

² D. Cai, X. He, J. Han, H.-J. Zhang, *Orthogonal laplacianfaces for face recognition*, "IEEE Transactions On Image Processing", 2006.

³ W. Caarls, P. Jonker, H. Corporaal, *Smartcam: Devices for embedded intelligent cameras, in PROGRESS 2002, 3rd seminar on embedded systems*, Proceedings (Utrecht, The Netherlands).

⁴ M. Cole, *Algorithmic skeletons: structured management of parallel computations, in Research Monographs in Parallel and Distributed Computing*, Massachusetts 1989.

duży, a czas przetwarzania musi być krótki i zwykle muszą być spełnione ograniczenia czasu rzeczywistego.

Dlatego w ciągu ostatnich dekad wzrosło zainteresowanie rozwojem i wykorzystaniem algorytmów równoległych w przetwarzaniu obrazu. Wykrywanie twarzy jest powiązane z pytaniem, czy na danym obrazie (zwykle w skali szarości) znajdują się jakieś twarze i jeśli są obecne, ustalenie lokalizacji obrazu i zawartości każdej twarzy. To pierwszy krok każdego w pełni automatycznego systemu, który analizuje informacje zawarte w twarzach (np.: tożsamość, płeć, wyraz twarzy, wiek, rasę i pozę).

Problem automatycznego rozpoznania twarzy związany jest z przetwarzaniem obrazów w celu wykrycia, sklasyfikowania i śledzenia najczęściej w obecności szumów. Badacze pracujący nad komputerowym widzeniem, przez wiele lat usiłują zamodelować podstawowe komponenty ludzkiego układu wizyjnego, aby zrozumieć naszą umiejętność widzenia. Kroki niezbędne do udanej implementacji zadania automatycznego rozpoznania wymagają automatycznego wykrycia, sklasyfikowania i śledzenia twarzy zlokalizowanej w obrazie. Aktualnie szybko rozwijane są metody wykorzystujące transformaty falkowe do rozpoznawania obrazów.

Transformata falkowa jest techniką wielorozdzielczościową, która może zostać zaimplementowana jako struktura typu piramidy lub drzewa i jest podobna do dekompozycji podpasmowej. Dyskretna transformata falkowa posiada własności, które czynią ją idealnym narzędziem do przetwarzania sygnałów 2D (obrazów) w zastosowaniach służących do rozpoznania. Właściwości takie jak: szybkie przetwarzanie, naturalna zdolność do przystosowania się do zmieniających się statystyk obrazu, skuteczna reprezentacja nagłych zmian i precyzyjna informacja o położeniu oraz zdolność do przystosowania się do silnych zakłóceń tła i niepewności dotyczących właściwości obiektu oraz względna niezależność od odległości pomiędzy obiektem a czujnikiem czynią ją niezwykle atrakcyjną w automatycznym rozpoznaniu.

Przetwarzanie obrazu

Klasyczne przetwarzanie obrazu odnosi się do przetwarzania funkcji dwuwymiarowej przez komputer. Obraz może być uważany za zawierający podobrazy, czasami określany jest jako obszar zainteresowania (ROI – *Regions-of-Interest*) lub po prostu region. Ta koncepcja odzwierciedla fakt, że obrazy często zawierają zbiory obiektów, z których każdy może być podstawą dla regionu. W zaawansowanym systemie przetwarzania obrazu powinno być możliwe zastosowanie określonych operacji przetwarzania dla wybranych regionów. W ten sposób jedna część obrazu (region) może być przetwarzana

w celu tłumienia rozmycia ruchu, podczas gdy inna część może być przetwarzana w celu poprawy odwzorowania kolorów. Zatem operacje przetwarzania obrazu można klasyfikować jako *niskiego poziomu*, *średniego poziomu* i *wysokiego poziomu*. Na tej podstawie klasyfikacji możliwe jest zdefiniowanie biblioteki szkieletowej dla operacji przetwarzania obrazu w celu przeprowadzenia operacji rozpoznawania.

- Operacje niskiego poziomu.
- Operacje średniego poziomu.
- Operacje wysokiego poziomu.

Możliwe jest użycie paradygmatu paralelizmu danych z podejściem master-slave do operacji przetwarzania obrazów niskiego, średniego i wysokiego poziomu⁵.

Procesor główny jest wykorzystywany do dzielenia i dystrybucji danych do procesorów podrzędnych. Procesor nadrzędny może również przetwarzać część obrazu (dane). Każdy procesor podrzędny przetwarza otrzymaną część obrazu (dane), a następnie procesor nadrzędny zbiera i składa obraz (dane) z powrotem. Przed przejściem do głównego przetwarzania obrazu jest on konwertowany do postaci cyfrowej. Digitalizacja obejmuje próbkowanie obrazu i kwantyzację próbkowanych wartości. Po przekształceniu obrazu na informacje bitowe, jest wykonywane właściwe przetwarzanie. Techniki przetwarzania to: *wstępne przetwarzanie obrazu*, *polepszanie obrazu*, *rekonstrukcja obrazu* i *kompresja obrazu*.

Funkcje wstępnego przetwarzania obrazu obejmują operacje, które są zwykle wymagane przed główną analizą danych oraz ekstrakcją informacji i są zazwyczaj grupowane jako poprawki radiometryczne lub geometryczne.

Polepszanie obrazu odnosi się do akcentowania lub wyostrenia cech obrazu, takich jak granice lub kontrast, aby uczynić wyświetlacz graficzny bardziej użytecznym do wyświetlania i analizy. Ten proces nie zwiększa wewnętrznej zawartości informacji w danych. Obejmuje on manipulację poziomem szarości i kontrastem, redukcję szumów, wyostrenie krawędzi, filtrowanie, interpolację i powiększenie, pseudo kolorowanie itd.

Rekonstrukcja obrazu dotyczy filtrowania obserwowanego obrazu w celu zminimalizowania wpływu degradacji. Skuteczność rekonstrukcji obrazu zależy od zakresu i dokładności wiedzy na temat procesu degradacji, a także od konstrukcji filtra. Przywracanie obrazu różni się od polepszania obrazu tym, że to drugie polega na większej ekstrakcji lub akcentowaniu cech obrazu.

⁵ C. Nicolescu, P. Jonker, *A data and task parallel image processing environment*, "Lecture Notes in Computer Science", vol. 2131/2001, pp. 393–408.

Kompresja obrazu ma na celu zminimalizowanie liczby bitów wymaganych do przedstawienia obrazu. Kompresja jest stosowana w telekonferencji, faksowaniu, transmisji, w dokumentach edukacyjnych i biznesowych, obrazach medycznych, które powstają w tomografii komputerowej, obrazowaniu rezonansu magnetycznego i cyfrowej radiologii, ruchu, zdjęciach, obrazach satelitarnych, mapach pogodowych, badaniach geologicznych itd.

Operacje klasyfikacji i analizy obrazu są wykorzystywane do cyfrowej identyfikacji i klasyfikacji pikseli w danych. Klasyfikacja jest zwykle wykonywana na wielokanałowych zestawach danych, a proces ten przypisuje każdy piksel w obrazie do określonej klasy lub motywu w oparciu o statystyczne cechy wartości jasności pikseli. Istnieje wiele podejść do wykonywania klasyfikacji cyfrowej. Dwa podstawowe podejścia są najczęściej stosowane, to: *klasyfikacja nadzorowana* i *nienadzorowana*.

Rozpoznawanie twarzy to termin obejmujący kilka podproblemów. Technika rozpoznawania twarzy składa się z trzech etapów: *wykrywania twarzy*, *ekstrakcji cech charakterystycznych* i *identyfikacja twarzy*.

Wykrywanie twarzy jest definiowane jako proces ekstrakcji twarzy z obrazów i scen. Tak więc system pozytywnie wykrywa pewien obszar obrazu rozumiany jako twarz. Ta procedura ma wiele zastosowań, takich jak: śledzenie twarzy, szacowanie pozy lub kompresja.

Ekstrakcja cech twarzy obejmuje uzyskiwanie odpowiednich cech charakterystycznych twarzy z danych. Cechy te mogą być określonymi obszarami twarzy, wariacjami, kątami lub miarami, które mogą być istotne dla osoby (np. oczy odstępny) lub nie. Ta faza też ma inne zastosowania, takie jak: śledzenie cech twarzy lub rozpoznawanie emocji.

Identyfikacja twarzy to proces dopasowywania jeden do wielu, który porównuje obraz twarzy zapytania ze wszystkimi obrazami szablonowymi w bazie danych, aby określić tożsamość twarzy zapytania. Identyfikacja obrazu testowego odbywa się poprzez zlokalizowanie w bazie danych obrazu, który ma największe podobieństwo do obrazu testowego. Proces identyfikacji to „zamknięty” test, co oznacza, że układ detekcji wykonuje obserwację osoby, o której wiadomo, że znajduje się w bazie danych. Podstawowe rozpoznawanie obejmuje albo *cechy własne*, albo *twarze własne*. Gdy system identyfikacji cech twarzy wykorzystuje *podejście twarzy własnej*, oznacza to, że interpretuje każdy obraz twarzy jako dwuwymiarowy zestaw jasnych i ciemnych obszarów w określonym wzorze. To właśnie te jasne i ciemne obszary są uważane za *twarze własne*. Wzory jasnych i ciemnych obszarów są następnie konwertowane i przedstawiane jako algorytm, który jest z kolei tymczasowo przechowywany jako kombinacja twarzy własnych. Ostatecznie, bieżąca kombinacja twarzy własnych jest skanowana przez

system identyfikacji twarzy i jest używana do porównywania z zapisanymi *twarzami własnymi* w bazie danych. Jednak podejście systemu cech własnych dąży do określenia odległości między takimi samymi cechami twarzy, jak: nos, oczy, struktura kości, usta i brwi. Istota tej metody polega na tym, że system identyfikacji twarzy przechwytuje twarz, a następnie wyodrębnia pewne *cechy własne*. Systemy cech twarzy można sklasyfikować jako *ciągłe* lub *wyzwalane*. *Systemy ciągłe* są zawsze aktywne i stale skanują obrazy twarzy. *Systemy wyzwalane* muszą zostać w jakiś sposób aktywowane, aby system mógł zeskanować twarz.

Aby zidentyfikować przestępcę, identyfikacja twarzy jest bardzo ważna, ponieważ gdy konkretna osoba popełni przestępstwo, można ją łatwo zidentyfikować za pomocą konkretnej techniki. Specjaliści kryminalistyki dysponują bazą danych dotyczącą istniejących przestępców zmagazynowaną w rejestrach. Zatem takie podejście może być łatwym zadaniem w zbieraniu informacji dotyczących osoby, która ma kartotekę kryminalną i jest podejrzana w konkretnej sprawie. Analizując konkretną sytuację i mając naocznego świadka, można prześledzić podejrzany obraz i upewnić się, że twarz wykonana jest podobna do twarzy, którą widziano podczas konkretnego przestępstwa. Aby porównać twarz z twarzami istniejącymi w bazie danych, najpierw proponuje się podzielić twarz na kilka równych części, takich jak: włosy, głowa, oczy, nos, usta i broda. Zatem dzieląc dane, staje się łatwiejsze zadanie identyfikacji twarzy. W większości przypadków byłaby ona rozpoznawana, gdyby brano pod uwagę podział obrazów.

Istnieją trzy głównie ważne metody konstruowania twarzy: *poprzez użycie funkcji naocznego świadka, dodawanie szczegółów oraz przycinanie obrazu*. W rezultacie otrzymujemy twarz jako ostateczny parametr identyfikacyjny, aby stwierdzić, kto popełnił przestępstwo.

Jeśli obraz twarzy jest dzielony na podobrazy (części), to staje się łatwe dopasowanie przestępcy do bazy danych. Używając tych podobrazów, możemy skonstruować *nowe twarze* zgodnie z przewidywaniami naocznego świadka. W celu skonstruowania nowego obrazu, możemy połączyć dowolne wycinki twarzy z dowolnym innym wycinkiem, aby skonstruować *nową twarz*. Teraz, w oparciu o te nowo skonstruowane twarze, możemy porównać te nowe twarze z wcześniej zapisanymi obrazami w bazie danych i rozpocząć dopasowywanie obrazu do innych obrazów, które mają pewne podobieństwa, aby móc uzyskać najlepsze dopasowanie z dostępnej bazy danych. Na końcu na ekranie wyświetlany jest wynikowy obraz konkretnej osoby, którą można uznać za najbardziej odpowiedniego podejznanego. Dzięki temu możemy zweryfikować (zidentyfikować) istniejące informacje o przestępstwie, w odniesieniu do których można łatwo zidentyfikować przestępców. Tak więc, może okazać się

to dobrym podejściem w identyfikacji przestępcy, dzięki algorytmowi bazowemu proponowanemu w pracy⁶. Ogólnie rzecz biorąc, obraz przedstawiany przez naocznego świadka jest obrazem namalowanym ręcznie. Korzystając z konkretnego namalowanego obrazu, można stworzyć dokładny obraz przestępcy^{7,8}. Jeśli konkretna osoba na obrazie nie jest dostępna w bazie danych, wówczas zespół śledczy może rozpocząć działanie w celu znalezienia przestępcy, uwzględniając szczegóły twarzy przedstawiane przez świadka. W przypadku gdy szczegóły twarzy zostaną dopasowane do bazy danych, wówczas poprzez podanie danych wejściowych w postaci obrazu twarzy, wyświetla się pełne dane biograficzne przestępcy, z zastosowaniem metody proponowanej w pracy⁹. Takie podejście pomaga zespołowi śledczemu w uzyskaniu innych szczegółów, które pomocne są w wyśledzeniu przestępcy.

Śledczy dysponują rejestrem karnym, który zazwyczaj zawiera dane osobowe o konkretnej osobie wraz ze zdjęciem. Aby w praktyce zidentyfikować przestępcę, potrzebują szczegółowych danych identyfikacyjnych dotyczących osoby, które są przekazywane przez naocznych świadków¹⁰. Na podstawie szczegółów podawanych przez naocznych świadków, przestępca, który popełnił przestępstwo, zostaje najczęściej zidentyfikowany „ręcznie”.

Ale praktycznie w istniejących systemach identyfikacji pojawiają się często następujące problemy:

- W większości przypadków jakość i rozdzielczość zarejestrowanych segmentów obrazu jest zaszumiona i trudno zidentyfikować twarz.
- Jeśli naoczny świadek obserwuje przestępcę tylko z jednego wybranego kierunku, to może nie być możliwe jego rozpoznanie.

⁶ J. Nautiyal, S. Gahlot, P. Kumar Mishra, *An Automated Technique for Criminal Face Identification Using Biometric Approach*, Conference on Advances in Communication and Control Systems 2013 (CAC2S 2013).

⁷ H. Fatemi, H.E. Malek, R. Kleihorst, H. Corporaal, P. Jonker, *Real-time face recognition on a mixed SIMD VLIW architecture*, in *PROGRESS 2003, 4th seminar on embedded systems*, Proceedings (Nieuwegein, The Netherlands), 22 October 2008.

⁸ S. Anila, N. Devarajan, *Global Journal of Computer Science and Technology Graphics & Vision Preprocessing Technique for Face Recognition Applications under Varying Illumination Conditions*, 2012.

⁹ D. Serot, J. Derutin, *Skipper: A skeleton-based programming environment for image processing applications*, in *Proceeding of the Fifth International Conference on Parallel Computing Technologies*, 1999.

¹⁰ J. Haddadnia, K. Faez, P. Moallem, *Human face recognition with moment invariants based on shape information*, in *Proceedings of the International Conference on Information Systems, Analysis and Synthesis*, vol. 20 (Orlando, Florida USA), International Institute of Informatics and Systemics (ISAS), 2001.

- Zdjęcie, które jest kopią papierową, może nie być podzielone lub rozbite na podobrazy. Dlatego bardzo trudno je znaleźć, chyba że zdobędziemy wszystkie szczegóły.
- Czasami naoczny świadek może nie być w stanie odtworzyć twarzy przestępcy.
- Czasami, jeśli przechowujemy „ręcznie” szczegóły przestępstwa to po pewnym czasie zdjęcia i inne szczegóły mogą się zatrzeć.

W ogólności ekstrakcja cech charakterystycznych wnosi mnóstwo zbiorów uogólnionych cech twarzy z kilku obrazów tej samej twarzy. Każdy obraz twarzy jest przetwarzany i cechy są ekstrahowane, a zbiór cech jest analizowany i łączony w jeden uogólniony zbiór cech, który jest zapisywany w bazie danych.

Twarz jest głównym celem uwagi w stosunkach społecznych, odgrywa główną rolę w przekazywaniu identyfikacji i emocji.

Chociaż zdolność do wnioskowania o inteligencji lub charakterze na podstawie wyglądu twarzy jest odrębnym zgadywaniem, to należy zauważyć, że mimo to ludzka zdolność do rozpoznawania twarzy jest niezwykła.

W niniejszej pracy, wykrycie twarzy jest proponowane poprzez obliczanie macierzy współwystąpień falkowych odcieni szarości ze szczegółów podpasm przetworzonych przy użyciu dyskretnej transformaty falkowej, nie zachodzących na siebie, ale przyległych podbloków różnych rozmiarów, zależnych od docelowego obrazu.

Z otrzymanych obliczeń, podblok z maksymalnymi połączonymi wartościami własności współwystąpień falkowych odcieni szarości jest określany jako ziarno okna. Następnie poprzez zastosowanie algorytmu obszaru wzrostu, podbloki lub regiony są grupowane w większy blok lub region w oparciu o kilka z góry zdefiniowanych kryteriów. Wtedy obiekt zainteresowania jest oznaczany poprzez narysowanie wokół niego prostokąta.

Celem niniejszego artykułu jest propozycja nowego narzędzia informatycznego w postaci falkowych cech współwystąpień poziomów szarości dla automatycznego rozpoznania twarzy przestępców.

Analiza tekstur

Sukces większości problemów związanych z komputerowym widzeniem zależy od tego jak skutecznie tekstura jest reprezentowana ilościowo. Bez znaczenia jest to, czy aplikacją jest wykrywanie obiektów, rozpoznanie, segmentacja tekstury czy wykrywanie krawędzi. System musi być zdolny do rozpoznawania i oznaczenia jednorodnych regionów tekstur w obszarze obrazu

oraz odróżnić każdy odrębny obszar^{11, 12}. W związku z powyższym, analiza tekstur jest jedną z najważniejszych technik wykorzystywanych w badaniu i interpretacji obrazów, składających się z powtarzających się lub quasi powtarzających się podstawowych elementów obrazu.

Analiza tekstur w celu segmentacji, klasyfikacji i rozpoznania wymaga identyfikacji odpowiednich atrybutów lub cech charakterystycznych, które wyróżniają teksturę w obrazie. Przyjmuje się, że cechy charakterystyczne są jednorodne wewnątrz obszarów zawierających te same tekstury. Początkowo, analiza tekstur była oparta na metodach statystycznych pierwszego lub drugiego rzędu^{13, 14, 15}.

Następnie w celu scharakteryzowania tekstur, zostały zaproponowane modele losowego pola Gaussiana Markov'a i losowego pola Gibbs'a. Adaptacyjny anizotropowy parametr oszacowania w modelu słabej membrany i adaptacyjnego systemu rozpoznania w celu segmentacji sceny zostały zaproponowane w pracy¹⁶. Później, lokalne przekształcenia liniowe zostały użyte do obliczenia cech charakterystycznych tekstury¹⁷. Następnie dla analizy tekstur została zaproponowana technika widma tekstury¹⁸. Powyższe tradycyjne podejścia statystyczne wykorzystywano do analizy tekstur, takich jak matryce współwystąpień odcieni szarości, statystyki drugiego rzędu, lokalne przekształcenia liniowe i widmo tekstury, które są ograniczone do analizy wzajemnych oddziaływań przestrzennych na relatywnie małych obszarach w stałej skali. W konsekwencji, ich wydajność jest wystarczająca jedynie przy analizie mikrotekstur.

¹¹ F. Espinal, T.L. Huntsberger, B.D. Jawerth, T. Kubota, *Wavelet-based fractal signature analysis for automatic target recognition*, "Optical Engineering", vol. 37, no. 1, pp. 166–174, 1998.

¹² P.P. Raghu, B. Yegnanarayana, *Segmentation of Gabor-filtered textures using deterministic relaxation*, IEEE Trans. Image Processing, vol. 5, no. 12, pp. 1625–1636, 1996.

¹³ R.M. Haralick, K. Shanmugam, I. Dinstein, *Textural features for image classification*, "IEEE Trans. Systems, Man, and Cybernetics", vol. 3, no. 6, pp. 610–621, 1973.

¹⁴ J. Sklansky, *Image segmentation and feature extraction*, "IEEE Trans. Systems, Man, and Cybernetics", vol. 8, no. 4, pp. 237–247, 1978. Automatic Target Detection Using Wavelet Transform 2673

¹⁵ P.C. Chen, T. Pavlidis, *Segmentation by texture using correlation*, "IEEE Trans. on Pattern Analysis and Machine Intelligence", vol. 5, no. 1, pp. 64–69, 1983.

¹⁶ T. Kubota, T.L. Huntsberger, *Adaptive pattern recognition system for scene segmentation*, "Optical Engineering", vol. 37, no. 3, pp. 829–835, 1998, Special section on advances in pattern recognition.

¹⁷ M. Unser, *Local linear transforms for texture measurements*, "Signal Processing", vol. 11, no. 1, pp. 61–79, 1986.

¹⁸ D.-C. He, L. Wang, *Texture unit, texture spectrum, and texture analysis*, "IEEE Transactions on Geoscience and Remote Sensing", vol. 28, no. 4, pp. 509–512, 1990.

Ostatnio, dużą uwagę zwrócono na metody oparte na analizach wielorozdzielczościowych lub wielokanałowych, takich jak filtry Gabor'a i transformata falkowa¹⁹. Niestety, wyjścia filtra Gabor'a nie są wzajemnie prostokątne, co może skutkować istotnym powiązaniem pomiędzy cechami charakterystycznymi tekstur. Ostatecznie, przekształcenia te są zwykle nieodwracalne, co ogranicza ich przydatność do syntezy tekstur. Większości z tych problemów można uniknąć, jeżeli użyje się transformaty falkowej, która zapewnia precyzyjny i ujednoczony szkielet do analizy i charakteryzacji sygnału w różnych skalach. Inną korzyścią transformaty falkowej w porównaniu z filtrami Gabor'a jest to, iż filtry dolnoprzepustowe i górnoprzepustowe użyte w transformacie falkowej pozostają takie same pomiędzy dwoma kolejnymi skalami, podczas gdy podejście Gabor'a wymaga filtrów o różnych parametrach. Innymi słowy, filtry Gabor'a wymagają odpowiedniego dostrojenia swoich parametrów dla różnych skal²⁰. Statystyczne cechy falkowe zostały zaproponowane i skutecznie użyte do scharakteryzowania i klasyfikacji tekstur w pracach^{21, 22}.

Wykrywanie obrazów obiektów

Ostatnio zaproponowano system wizyjny z ekstraktorem cech charakterystycznych czasu rzeczywistego przy użyciu techniki wielorozdzielczościowej²³. Algorytm wykrywający granice przy użyciu dipola krawędzi i pola krawędzi został zaprezentowany w pracy²⁴. Do wykrywania konturów zostało zaproponowane podejście adaptacyjne, oparte na pikselach

¹⁹ M. Unser, *Texture classification and segmentation using wavelet frames*, "IEEE Trans. Image Processing", vol. 4, no. 11, pp. 1549–1560, 1995.

²⁰ T. Chang and C.-C. J. Kuo, *Texture analysis and classification with tree-structured wavelet transform*, IEEE Trans. Image Processing, vol. 2, no. 4, pp. 429–441, 1993.

²¹ G.M. Haley, B.S. Manjunath, *Rotation-invariant texture classification using modified Gabor filters*, in Proc. IEEE International Conference on Image Processing (ICIP '95), pp. 262–265, Washington, DC, USA, October 1995; G. Van de Wouwer, P. Scheunders, D. Van Dyck, *Statistical texture characterization from discrete wavelet representations*, "IEEE Trans. Image Processing", vol. 8, no. 4, pp. 592–598, 1999.

²² T. Kubota, T.L. Huntsberger, C.O. Alford, *A vision system with real-time feature extractor and relaxation network*, "International Journal of Pattern Recognition and Artificial Intelligence", vol. 12/1998, no. 3, pp. 335–354.

²³ T. Kubota, T.L. Huntsberger, C.O. Alford, *A vision system with real-time feature extractor and relaxation network*, "International Journal of Pattern Recognition and Artificial Intelligence", vol. 12/1998, no. 3, pp. 335–354.

²⁴ T. Kubota, T.L. Huntsberger, *Edge dipole and edge field for boundary detection*, "Signal, Image and Video Processing" VI, vol. 3389 of Proceedings of SPIE, pp. 179–189, Orlando, Fla, USA, April 1998.

połączenia danych²⁵. W pracach^{26,27} zaproponowano techniki oparte na falkach do automatycznego wykrycia i rozpoznania obiektów. Podjęto również próbę zastosowania falkowego wymiaru fraktalnego dla automatycznego rozpoznawania obiektów²⁸.

W pracy²⁹ zostały zaproponowane kontury Chernoff'a dla automatycznego rozpoznawania obiektów ze skompresowanych danych. Uporządkowany zespolony algorytm optycznego przepływu został użyty do wykrywania ruchomych obiektów na obrazach np. z zakresu podczerwieni³⁰. W kolejnej publikacji³¹ do wykrywania oraz klasyfikacji obiektów użyto statystycznej analizy spektralnej i falkowego współczynnika charakteryzacji. Następnie falki zostały wykorzystane do detekcji obiektów³². Howard i współpracownicy zaproponowali grupowanie poprzedzone bezpośrednią nominalną analizą składową dla inteligentnej detekcji obiektu w czasie rzeczywistym³³.

²⁵ T. Kubota, *Adaptive pixel-based data fusion for boundary detection*, in *Proc. Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR '99)*, Lecture Notes in Computer Science, pp. 173–188, Springer-Verlag, York, UK, July 1999.

²⁶ T.L. Huntsberger, B.D. Jawerth, *Wavelet based automatic target detection and recognition*, Annual Tech. Rep., University Research Initiative Program for Combat Readiness, University of South Carolina, Columbia, SC, USA, 1998.

²⁷ T.L. Huntsberger, B.D. Jawerth, *Wavelet based automatic target detection and recognition*, Annual Tech. Rep., University Research Initiative Program for Combat Readiness, University of South Carolina, Columbia, SC, USA, 1998.

²⁸ A. Jain, P. Moulin, K. Ramchandran, *Chernoff bounds for automatic target recognition from compressed data*, in *Proc. IEEE Information Theory Workshop on Detection, Estimation, Classification and Imaging (DECI '99)*, p. 57, Santa Fe, NM, USA, February 1999. 2674 EURASIP Journal on Applied Signal Processing.

²⁹ A. Jain, P. Moulin, K. Ramchandran, *Chernoff bounds for automatic target recognition from compressed data*, in *Proc. IEEE Information Theory Workshop on Detection, Estimation, Classification and Imaging (DECI '99)*, p. 57, Santa Fe, NM, USA, February 1999. 2674 EURASIP Journal on Applied Signal Processing.

³⁰ G. Castellano, J. Boyce, M. Sandler, *Moving target detection in infrared imagery using a regularized CDWT optical flow*, in *Proc. IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications (CVBVS '99)*, pp. 13–22, Fort Collins, Colo, USA, June 1999.

³¹ Y. Tian, H. Qi, X. Wang, *Target detection and classification using seismic signal processing in unattended ground sensor systems*, in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP '02)*, vol. 4, p. 4172, Orlando, Fla, USA, May 2002.

³² G. Boccignone, A. Chianese, A. Picariello, *Using Renyi's information and wavelets for target detection: an application to mammograms*, "Pattern Analysis and Applications", vol. 3, no. 4, pp. 303–313, 2000.

³³ A. Howard, C. Padgett, K. Brown, *Real time intelligent target detection and analysis with machine vision*, in *Proc. 3rd International Symposium on Intelligent Automation and Control*, World Automation Congress (ISIAC-WAC '00), Maui, Hawaii, June 2000.

Kolejną koncepcją, zaimplementowaną w publikacji³⁴ dla zadań wykrywania obiektów naturalnych w zaszumionych obrazach jest mechanizm przeduważnej selekcji (*preattentive selection*) oparty na architekturze systemu wizyjnego. Kubota i współpracownicy zaproponowali probabilistyczne relacje (*probabilistic relaxation*) bazujące na krawędziach dla ekstrakcji konturów subpikseli, które są użyteczną techniką przy wykrywaniu obiektu^{35, 36}.

Pomimo, że zostały już podjęte znaczne wysiłki związane z analizą tekstur i wykrywaniem obiektów, ich przydatność w zadaniach wykrywania twarzy z obrazów zaszumionych, jest nadal ograniczona. Proponowane podejście wykorzystuje cechy współwystąpień falkowych odcieni szarości, pochodzące ze szczegółów podpasem obrazów przekształconych przy użyciu falek dyskretnych do wykrywania twarzy z obrazów zaszumionych.

Dyskretne przekształcenia falkowe

Sygnały jednowymiarowe

Falki są to funkcje wygenerowane z pojedynczej funkcji ψ przez skalowanie (j) i translację (k).

$$\psi_{j,k}(t) = 2^{-j/2} \cdot \psi(2^{-j}t - k); \dots j, \dots k \in Z \quad (1)$$

Funkcja lub sygnał może być uważany jako złożony z „gładkiego tła” oraz wahań lub szczegółów (detali). Różnica między „gładką częścią” a szczegółami jest określana przez rozdzielczość (2^{-j}) która wzrasta gdy j dąży do $-\infty$, to znaczy przez skalę (2^j), poniżej której szczegóły lub sygnał nie mogą być dostrzeżone. Na danym poziomie rozdzielczości (2^{-j}) sygnał jest aproksymowany przez ignorowanie wszystkich wahań poniżej danej skali. Możemy wyobrazić sobie stopniowy wzrost rozdzielczości, przykładowo na każdym stopniu wzrostu rozdzielczości dodawane są drobne szczegóły do zgrubnego opisu, dostarczając kolejno lepsze aproksymacje dla sygnału. Ostatecznie, gdy rozdzielczość dąży do nieskończoności odzyskujemy sygnał oryginalny. Schemat poniższy (rys. 1) przedstawia intuicyjnie wielorozdzielczą reprezentację

³⁴ L. Itti, C. Gold, C. Koch, *Visual attention and target detection in cluttered natural scenes*, "Optical Engineering", vol. 40, no. 9, pp. 1784–1793, 2001.

³⁵ T. Kubota, T.L. Huntsberger, J.T. Martin, *Edge based probabilistic relaxation for sub-pixel contour extraction*, in Proc. 3rd International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR '01), pp. 328–343, INRIA, Sophia Antipolis, France, September 2001.

³⁶ T. Niedziela, *Metody falkowe analizy sygnałów*, „Logistyka” 6/2014, s. 7788–7796.

sygnału jednowymiarowego (poziomy reprezentacji sygnałów). Sygnał S na pierwszym poziomie rozdzielczości (reprezentacji) jest rozkładany na część zgrubną – aproksymację $A1$ oraz szczegóły $D1$, później proces jest powtarzany dla aproksymacji sygnału $A1$. Następnie rozkładany jest odpowiednio na aproksymację sygnału $A2$ i szczegół $D2$.

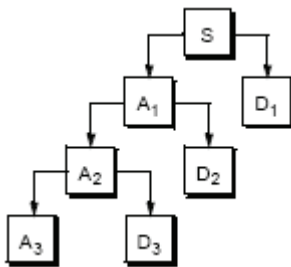
Oznacza to, że po rozłożeniu na aproksymacje i szczegóły sygnał może być prezentowany na różnych poziomach rozdzielczości (reprezentacji) np.:

$$S = A1 + D1$$

lub

$$S = A2 + D2 + D1 \dots\dots \text{itd.}$$

Rys. 1. Intuicyjny algorytm dekompozycji sygnału



$$S = A_1 + D_1$$

$$= A_2 + D_2 + D_1$$

$$= A_3 + D_3 + D_2 + D_1$$

Powyższy intuicyjny opis sygnału może być przedstawiony bardziej precyzyjnie w następujący sposób. Etykietujemy poziom rozdzielczości przez liczbę całkowitą j . Skala związana jest z poziomem od $j = 0$ do $j = 1$ na poziomie j z $1/2^j$.

Rozważmy przykładową funkcję $f(t)$. Na poziomie rozdzielczości j jest ona aproksymowana przez $f^j(t)$. Na następnym poziomie rozdzielczości $j+1$, szczegóły są oznaczone przez $d^j(t)$. Mamy, zatem aproksymację $f(t)$ na nowym poziomie rozdzielczości $f^{j+1}(t) = f^j(t) + d^j(t)$. Oryginalna funkcja $f(t)$ jest odzyskiwana, gdy rozdzielczość dąży do nieskończoności.

$$f(t) = f_j(t) + \sum_{k=1}^{k=\infty} d_k(t) \tag{2}$$

Analiza wielorozdzielcza odnosi się do równoczesnej obecności różnych rozdzielczości. Powyższe równanie reprezentuje jedną z dróg dekompozycji funkcji $f(t)$ na gładką część oraz część szczegółową. Podobnie możemy przedstawić przestrzeń funkcji całkownych z kwadratem $L^2(R)$, jako złożoną z ciągu podprzestrzeni W^k oraz podprzestrzeni V^j . W ten sposób

aproxymacja $f(t)$ na poziomie rozdzielczości j rozumiana będzie jako $f^j(t)$ w przestrzeni V^j a szczegóły $d^k(t)$ w przestrzeni W^k .

W dyskretnej transformacji falkowej (DWT – *Discrete Wavelet Transform*) generalnie używamy falek $\psi(t)$, które są powiązane z funkcją skalującą $\phi(t)$.

W analizie wielorozdzielczościowej funkcja falkowa $\psi(t)$ i funkcja skalująca $\phi(t)$ powiązane są równaniami:

$$\psi(t) = \sum_{k \in Z} g_k \cdot \phi(2t - k) \quad (3)$$

$$\phi(t) = \sum_{k \in Z} h_k \cdot \phi(2t - k) \quad (4)$$

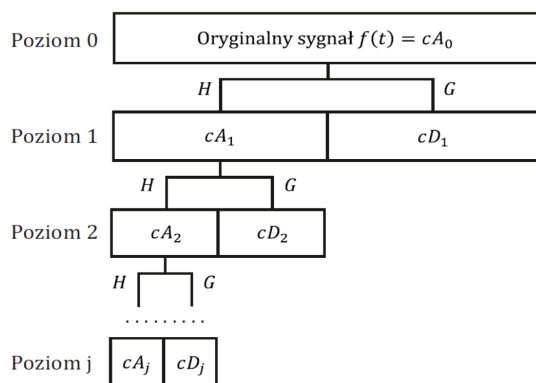
gdzie: h_k i g_k są odpowiednio współczynnikami dwóch filtrów, dolnoprzepustowego H oraz górnoprzepustowego G . Relacje pomiędzy tymi współczynnikami są dane równaniem:

$$g_k = -(-1)^k \cdot h_{L-k}; \dots k = 0, \dots, L-1 \quad (5)$$

gdzie: L jest długością filtrów: H oraz G (filtry mają taką samą długość).

Filtry te są używane w rekurencyjnej dekompozycji sygnału (algorytmu zaproponowanego przez Mallat'a). Algorytm ten jest również znany jako „algorytm piramidowy” lub „algorytm drzewa” i oferuje hierarchiczną, wielorozdzielczościową reprezentację sygnału (poziomy reprezentacji sygnału).

Rys. 2. Rekurencyjny algorytm dekompozycji sygnału (algorytm drzewo)



W algorytmie „drzewie”, sygnał o długości N jest przepuszczany przez dolno- (H) i górno-przepustowe (G) filtry. Poprzez down-sampling na wyjściu filtrów mamy dwa zbiory $n/2$ współczynników.

Filtr górnoprzepustowy (G) filtruje dane reprezentujące wysoko częstotliwościowe składowe oryginalnego sygnału. Filtr dolnoprzepustowy (H) filtruje dane reprezentujące nisko częstotliwościowe składowe oryginalnego sygnału. Na pierwszym poziomie rozdzielczości (poziomie reprezentacji), nazywamy ten poziom szczegółowymi współczynnikami. Przybliżenie współczynnika może być użyte wtedy jako dane wejściowe dla następnego kroku analizy, zgodnie z tym co ilustruje rys. 2 (poziomy reprezentacji sygnału).

Możemy również zrekonstruować oryginalny sygnał z falkowej dekompozycji. Aby tego dokonać potrzebujemy filtry rekonstrukcyjne wywodzące się z filtrów rekonstrukcji H i G , które są postaci:

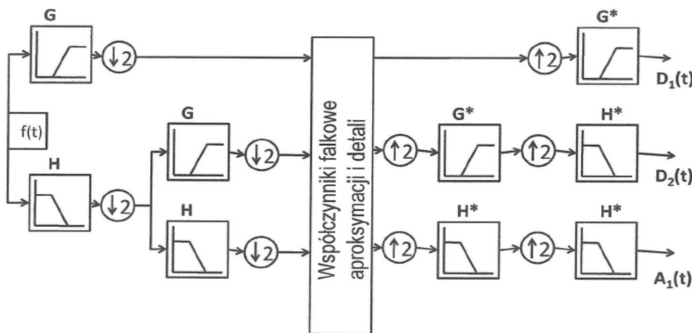
$$h \cdot r_k = h_{L-k}; \dots k = 0, \dots, L - 1 \tag{6}$$

$$g \cdot r_k = g_{L-k}; \dots k = 0, \dots, L - 1 \tag{7}$$

gdzie: $h \cdot r_k$ są współczynnikami dolnoprzepustowego H^* filtra rekonstrukcyjnego, natomiast $g \cdot r_k$ są współczynnikami górnoprzepustowego H^* filtra rekonstrukcyjnego.

Używając tych filtrów rekonstrukcyjnych możemy zrekonstruować przybliżenia $A_n(t)$ i szczegóły sygnału $D_{n-1}(t)$, $D_n(t)$ na różnym poziomie dekompozycji, tak jak zilustrowano na rys. 3.

Rys. 3. Dwu poziomowa dekompozycja falkowa sygnału $f(t)$ i rekonstrukcja przybliżenia $A_1(t)$ oraz szczegółów $f(t)$ $D_1(t)$, $D_2(t)$ sygnału



Z powyższego rysunku możemy prześledzić down-sampling po każdym etapie dekompozycji, i up-sampling przed każdym etapem rekonstrukcji. Oryginalny sygnał na pierwszym poziomie rozdzielczości może być obliczony jako:

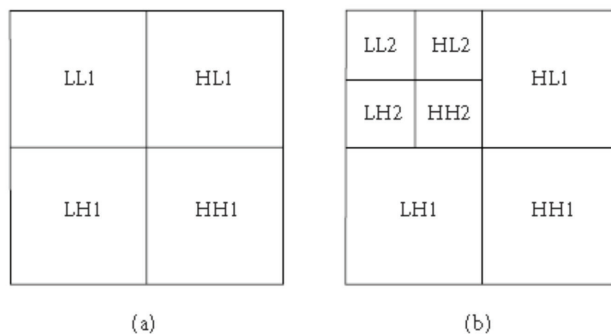
$$f(t) = A_2(t) + D_1(t) + D_2(t) \quad (8)$$

Dowolna superpozycja falkowa rozkłada określoną funkcję na różne poziomy, gdzie każdy poziom jest w znacznej mierze uwarunkowany rozdzielczością dostosowaną do ustalonego poziomu.

Sygnały dwuwymiarowe

Dyskretna transformata falkowa jest identyczna jak hierarchiczny system podpasma, gdzie podpasma są logarytmicznie rozmieszczone w przestrzeni częstotliwości i reprezentują rozkład oktaowy. W rezultacie takiego podejścia, sygnał dwuwymiarowy (obraz) jest w rzeczywistości podzielony, czyli rozłożony na cztery podpasma (LL1, HL1, LH1, HH1) i dyskretnie podpróbkowany, tak jak zilustrowano na rys. 4a.

Rys. 4. Schemat dekompozycji wielorozdzielczościowej obrazu: (a) jeden poziom dekompozycji, (b) dwa poziomy dekompozycji

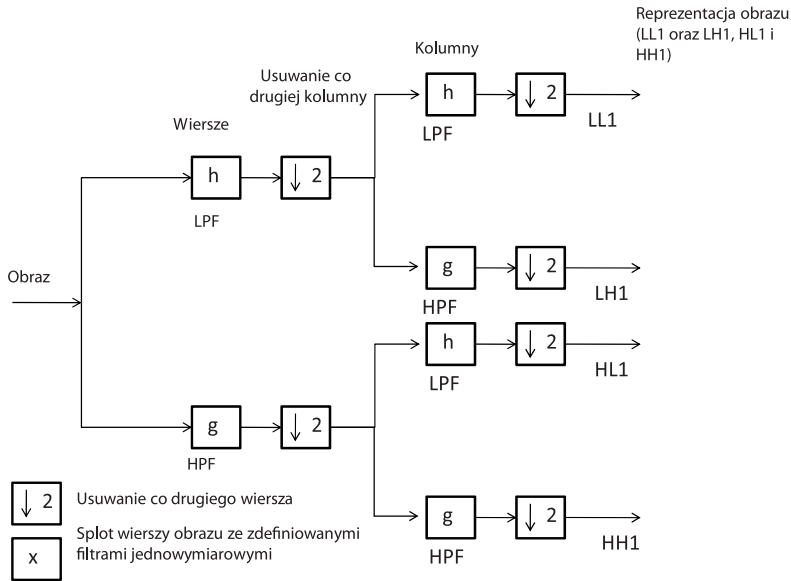


Te cztery podpasma powstały z oddzielnego zastosowania pionowych i poziomych filtrów tak jak zilustrowano na rys. 5.

Filtry h i g pokazane na rys. 5 są odpowiednio jednowymiarowymi filtrami dolnoprzepustowymi (LPF) i górnoprzepustowymi (HPF). A zatem, dekompozycja obrazu tworzy podpasma odpowiadające różnym poziomom rozdzielczości i orientacji.

Podpasma oznaczone jako LH1, HL1 i HH1 reprezentują najwyższą skalę współczynnika falkowego, który jest obrazem szczegółów, podczas gdy podpasmo LL1 odpowiada za współczynnik poziomu szumów, który jest obrazem przybliżonym. Aby uzyskać następny poziom szumu współczynnika falkowego, tylko podpasmo LL1 zostanie dekomponowane i dyskretnie próbkowane przy użyciu podobnego zespołu filtrów jak zilustrowano na rys. 5.

Rys. 5. Algorytm jednopoziomowej dekompozycji falkowej obrazu 2D



Wynikiem tego jest dwupoziomowa dekompozycja falkowa tak jak zilustrowano na rys. 4b. Podobnie, dla uzyskania dalszej dekompozycji, zostanie wykorzystane tylko podpasmo LL2. Proces ten trwa dopóki nie zostanie osiągnięta ustalona skala docelowa dekompozycji.

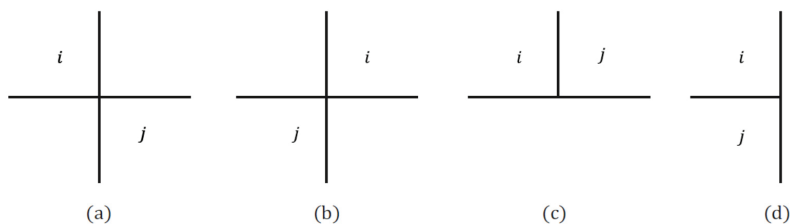
Wartości lub przekształcone współczynniki falkowe w obrazie przybliżonym i w obrazach szczegółowych (obrazów podpasm) są cechami charakterystycznymi, przydatnymi do rozróżniania tekstur i segmentacji. Tekstury, zarówno w skali mikro jak i makro, mają niejednorodne zmiany poziomów szarości, zatem są one statystycznie charakteryzowane przez wartości obrazów przekształconych przy użyciu podpasm dyskretnej transformaty falkowej lub cech charakterystycznych uzyskanych z tych obrazów lub ich kombinacji. Innymi słowy cechy charakterystyczne uzyskane z obrazów przybliżonych i szczegółowych podpasm niepowtarzalnie charakteryzujących teksturę. Cechy charakterystyczne uzyskane z obrazów przekształconych przy użyciu

dyskretnej transformaty falkowej są w niniejszej pracy zilustrowane jako użyteczne do wykrywania obrazów twarzy.

Macierz współwystąpień poziomów szarości

Metoda współwystąpień poziomów szarości do opisu tekstur jest oparta na powtarzających się wystąpieniach określonej konfiguracji poziomów szarości w teksturze, gdzie konfiguracja ta zmienia się szybko wraz z odległością w teksturach dobrej jakości, natomiast powoli w teksturach zaszumionych. Załóżmy, że część teksturowanego obrazu, która ma zostać poddana analizie ma rozmiar $M \times N$. W macierzy współwystąpień poziomów szarości, wystąpienie jednej z konfiguracji poziomów szarości jest opisane przez macierz częstości względnych $C\theta, d(i, j)$ opisujących jak często dwa piksele o poziomach szarości i, j pojawiają się w oknie oddzielnym przez wektor przesunięcia d w kierunku θ .

Rys. 6. Macierz orientacji wektora przemieszczeń współwystąpień poziomów szarości



$$(a) \ d = (1, 1) = 135^\circ, \quad (b) \ d = (1, -1) = 45^\circ, \quad (c) \ d = (0, 1) = 0^\circ, \\ (d) \ d = (1, 0) = 90^\circ,$$

Przykładowo, jeżeli wektor przemieszczenia współwystąpień poziomów szarości (rys. 6) jest ustalony jako $(1, 1)$, wtedy posiada on następującą interpretację: jeden piksel poniżej i jeden piksel na lewo, w kierunku 135° , tak jak to pokazano na rys. 6a, natomiast jeżeli jest ustalony jako $(1, -1)$ to posiada interpretację: jeden piksel poniżej i jeden piksel na prawo w kierunku 45° tak jak to pokazano na rys. 6b. Podobnie, wektor przemieszczenia $(0, 1)$ posiada interpretację: zero pikseli poniżej i jeden piksel na prawo, czyli w kierunku 0° , tak jak to pokazano na rys. 6c i wektor przemieszczenia $(1, 0)$ ma interpretację: jeden piksel poniżej i zero pikseli w lewo, czyli w kierunku 90° tak jak to pokazano na rys. 6d.

Macierze współwystąpień poziomów szarości są symetryczne jeżeli są zdefiniowane zgodnie z poniższymi wskazówkami. Jednakże, może zostać użyta definicja asymetryczna, gdzie wartości matrycy są zależne od kierunku współwystąpień poziomów szarości. Nieznormalizowane częstotliwości współwystąpień poziomów szarości w funkcji kąta i odległości mogą być zdefiniowane jako:

$$C_{0^0,d}(i,j) = \left\{ \left[\begin{array}{l} [(k,l), (m,n)] \in D : k-m=0, \\ |l-n|=d, \quad f(k,l)=i, \quad f(m,n)=j \end{array} \right] \right\} \quad (9)$$

$$C_{45^0,d}(i,j) = \left\{ \left[\begin{array}{l} [(k,l), (m,n)] \in D : (k-m=d, l-n=-d) \\ \text{lub } (k-m=-d, l-n=d), \quad f(k,l)=i, \quad f(m,n)=j \end{array} \right] \right\} \quad (10)$$

$$C_{90^0,d}(i,j) = \left\{ \left[\begin{array}{l} [(k,l), (m,n)] \in D : |k-m|=d, l-n=0 \\ f(k,l)=i, \quad f(m,n)=j \end{array} \right] \right\} \quad (11)$$

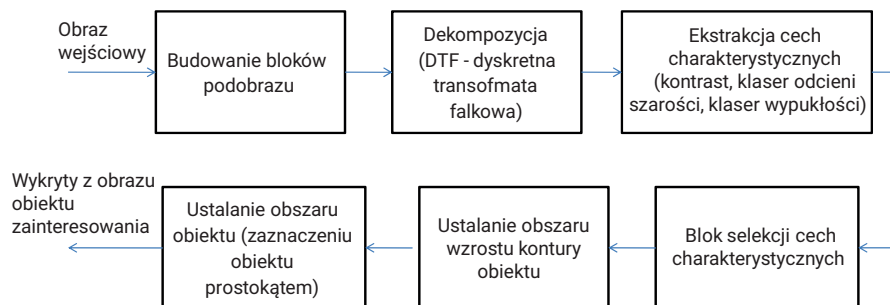
$$C_{135^0,d}(i,j) = \left\{ \left[\begin{array}{l} [(k,l), (m,n)] \in D : (k-m=d, l-n=d) \\ \text{lub } (k-m=-d, l-n=-d), \quad f(k,l)=i, \quad f(m,n)=j \end{array} \right] \right\} \quad (12)$$

gdzie $\{ \cdot \cdot \cdot \}$ odnosi się do ustalonej mocy matrycy zbioru $D = (M \times N) \times (M \times N)$. Macierz współwystąpień poziomów szarości $C(i,j)$ może zostać uzyskana poprzez zliczenie wszystkich par pikseli mających poziomy szarości i oraz j , rozdzielonych zadanym wektorem przemieszczenia d w zadanym kierunku.

System wykrywania twarzy

Kroki niezbędne w procesie wykrywania twarzy (obiektów) zostały zilustrowane na rys. 7.

Rys. 7. Schemat ideowy systemu wykrywania twarzy (obiektów) z obrazów zakresu widzialnego



Algorytm wykrywania twarzy z obrazów zakresu widzialnego

Wejście: Obraz wyjściowy o rozmiarze $N \times N$

Wyjście: Obraz wykrytego obiektu zainteresowania

- (1) Wczytaj obraz wejściowy.
- (2) Uzyskaj 32×32 lub 16×16 bloki podobrazu, rozpoczynając z górnego lewego rogu.
- (3) Dokonaj dekompozycji bloku podobrazu używając *2D-dyskretnej transformaty falkowej*.
- (4) Określ macierz współwystąpień poziomów szarości dla podpasm szczegółów bloku podobrazu zdekomponowanego przy użyciu *dyskretnej transformaty falkowej*.
- (5) Oblicz cechy charakterystyczne dyskretnej transformaty falkowej takie jak: *kontrast, klaster odcieni, i klaster wypukłości, z macierzy współwystąpień poziomów szarości*.
- (6) Powtarzaj kroki 2 do 5 dla wszystkich bloków podobrazu.
- (7) Posortuj sumę wartości cech charakterystycznych wszystkich okien w porządku rosnącym, następnie wybierz okno, mające maksymalne połączone wartości cech charakterystycznych (K_{wyp}) jako okno ziarna.
- (8) Uzyskaj próg, czyli, średnią sumę cech charakterystycznych pierwszych okien $n\%$
- (9) Nałóż algorytm rosnącego obszaru używając metody średniej odległości poprzez złączenie okien oparty na progu i sąsiedztwie.
- (10) Zakreśl obiekt zainteresowania prostokątem.

Rozważmy obrazy wejściowe o rozmiarze $N \times N$. Wykrywanie twarzy jest realizowane poprzez analizowanie nienachodzących na siebie podobrazów (np. bloków) różnych rozmiarów, zależnych od obrazów docelowych. Każdy odrębny blok podobrazu, wzięty z górnego lewego narożnika obrazu oryginalnego, jest zdekomponowany przy użyciu jedno- lub dwu-poziomowej dyskretnej transformaty falkowej i macierzy falkowych współwystąpień poziomów szarości (C) które są wydzielone dla $\theta = 135^\circ$ i $d = (1, 1)$ (np. jeden piksel poniżej i jeden piksel na lewo) dla szczegółowych podpasm (np., LH1, HL1, HH1, LH2, HL2, oraz HH2).

Należy zauważyć, iż bardzo ważne jest aby wymagany poziom dekompozycji dyskretnej transformaty falkowej zależał od użytego rozmiaru okna, czyli, dla większych rozmiarów okien, obraz może zostać dekomponowany na więcej poziomów dyskretnej transformaty falkowej, podczas gdy przy mniejszym rozmiarze okna, należy użyć mniejszej liczby poziomów dekompozycji dyskretnej transformaty falkowej. Oznacza to, że rozmiar okna analizy zależy od rozmiaru obiektu i obrazu.

Następnie, z matrycy współwystąpień poziomów szarości (C), obliczane są cechy charakterystyczne dyskretnej transformaty falkowej, takie jak kontrast, klaster odcieni, i klaster wypukłości przy użyciu wzorów podanych w relacji (13) do (15)

$$\text{Kontrast} = \sum_{i,j=1}^N (i-j)^2 \cdot C(i,j) \quad (13)$$

$$\text{Klaster odcieni} = \sum_{i,j=1}^N (i - M_x + j - M_y)^3 \cdot C(i,j) \quad (14)$$

$$\text{Klaster wypuk.} = \sum_{i,j=1}^N (i - M_x + j - M_y)^4 \cdot C(i,j) \quad (15)$$

gdzie:

$$M_x = \sum_{i,j=1}^N i \cdot C(i,j), \quad M_y = \sum_{i,j=1}^N j \cdot C(i,j) \quad (16)$$

Te wartości cech charakterystycznych są poddane liniowej lub logarytmicznej normalizacji, w zależności od ich zasięgów dynamicznych. Cechy kontrastu mają umiarkowane wartości i w związku z tym są poddawane normalizacji liniowej, podczas gdy klaster odcieni i klaster wypukłości są poddawane normalizacji logarytmicznej, ponieważ mają one bardzo duże wartości zasięgów dynamicznych. Wybierając blok ziarna często można bazować na istocie problemu. Kiedy wcześniejsza informacja nie jest dostępna, liczymy w każdym pikselu lub subregionie ten sam zestaw właściwości, które ostatecznie zostaną użyte do wybrania ziarna oraz do procesu rozrostu. W proponowanej implementacji, blok podobrazu, z maksymalnymi połączonymi znormalizowanymi cechami charakterystycznymi (takimi jak: wartość kontrastu, klaster odcieni i klaster wypukłości (K_{wyp}) jest identyfikowany jako blok ziarna lub okno ziarna. Koncepcja cech falkowych i współwystąpień poziomów szarości wskazuje, że wartości tych cech są wysokie dla okna, które jest najprawdopodobniej fragmentem obiektu. Obszar wzrostu jest procesem segmentacji

opartym na obszarze, w którym podobszary rosną w większe obszary w oparciu o predefiniowane kryteria, takie jak próg i sąsiedztwo.

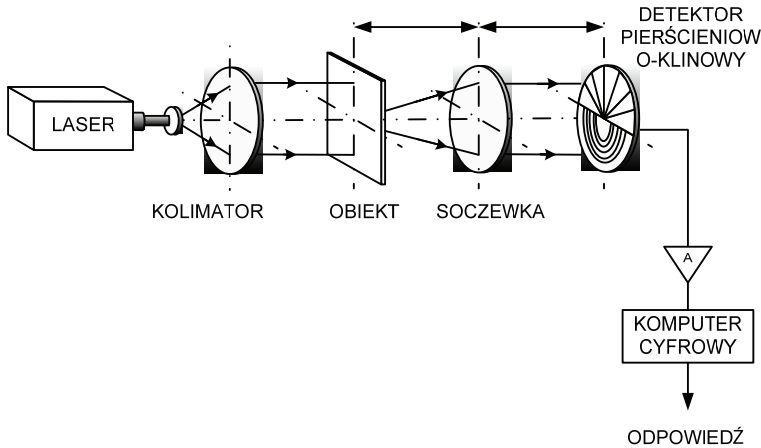
W proponowanej implementacji, algorytm obszaru wzrostu jest oparty na metodzie średniej odległości. Pierwszym krokiem w tej metodzie, jest sortowanie wartości cech charakterystycznych wszystkich okien, czyli bloków podobrazu w porządku rosnącym, tak aby okno którego wartość jest największa było oknem ziarna. Próg jest określony poprzez znalezienie wartości średniej (A) pierwszych n okien, które są wybrane adaptacyjnie w zależności od obrazu docelowego. Następnie, wartości cech charakterystycznych wszystkich 8 sąsiadujących bloków są porównywane z wartością średniej A , i wartością cech charakterystycznych K_{wyp} . Okno, którego wartość jest bliższa do wartości progowej K_{wyp} zostanie połączone z oknem ziarna. Proces ten jest powtarzany dla wszystkich 8 przyległych bloków. Jeżeli żaden z 8 przylegających bloków nie zostanie połączony, algorytm kończy działanie. Jeżeli przynajmniej jedno okno spośród 8 przyległych bloków zostanie połączone, powyższa procedura będzie powtarzana z 16 przyległymi i tak dalej. Na koniec jest rysowany prostokąt otaczający wszystkie łączone okna, aby zakreślić obszar wykrytego obiektu zainteresowania. Algorytm wykrywania obiektu jest podany w ramce Algorytm wykrywania obiektów.

Rozpoznawanie obrazów dyfrakcyjnych twarzy

W przeprowadzonym eksperymencie testowano system optoelektroniczny oparty na próbkowaniu obrazu dyfrakcyjnego przez cyfrowy detektor pierścieniowo-klinowy do rozpoznawania obrazów twarzy (rys. 8) przez sztuczną sieć neuronową.

Zastosowanie próbkowania obrazu dyfrakcyjnego detektorem pierścieniowo-klinowym (DPK) oraz sztucznych sieci neuronowych (SSN) w roli klasyfikatorów prowadzi do opracowania łatwych i szybkich algorytmów rozpoznawania obrazów twarzy. Algorytmy te potrzebują mniej czasu na napisanie oprogramowania.

Celem eksperymentu było oszacowanie optymalnych wymiarów cyfrowego detektora pierścieniowo-klinowego do próbkowania transformaty Fouriera (kwadratu modułu transformaty) przez matrycę detektorów. Przyjęto założenie, że obrazy twarzy wykorzystane w eksperymencie mają zróżnicowany rozmiar i zróżnicowane nachylenie głowy. W naszym podejściu realizowane było próbkowanie kwadratu modułu funkcji intensywności transformaty Fouriera obrazów twarzy przez detektor pierścieniowo-klinowy oraz trening sztucznych sieci neuronowych z tymi właśnie obrazami. Twarze użyte podczas treningu przez sieć neuronową różniły się wielkością i nachyleniem głowy.

Rys. 8. Schemat ideowy próbkowania obrazu dyfrakcyjnego twarzy^{37, 38}

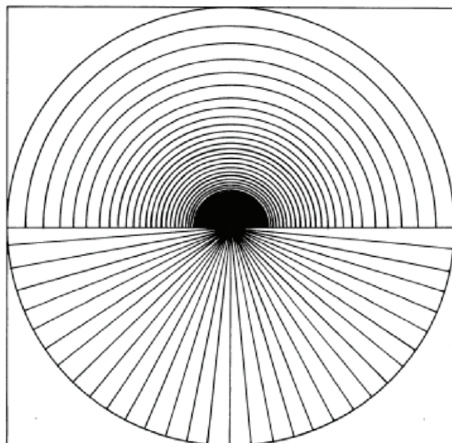
Obraz w płaszczyźnie dyfrakcyjnej zawiera wszystkie informacje związane z rozpoznawanym obiektem (twarzą). Rozkład energii światłej w obrazie dyfrakcyjnym może być rejestrowany w funkcji kąta i promienia przez próbkowanie przestrzeni Fouriera w reprezentacji biegunowej. Proces próbkowania w płaszczyźnie dyfrakcyjnej pola świetlnego padającego na wybrany element (obszar) detektora pierścieniowo-klinowego (DPK) jest obliczaniem kwadratu modułu funkcji intensywności. DPK to właściwie okrągła matryca detektorów w postaci półpierścieni i klinów. Detektor pierścieniowo-klinowy (rys. 9) posiada sensory fotoelektryczne w postaci półokrągłych pierścieni oraz klinów rozchodzących się ze środka matrycy detektorów. Transformata Fouriera jest symetryczna tzn. w każdej półpłaszczyźnie DPK znajduje się pełna informacja o obrazie wejściowym.

Próbkowanie promieniowe obrazów dyfrakcyjnych przez półpierścienie zapewnia niezależność detekowanej informacji od orientacji przestrzennej częstości w obrazie. Natomiast próbkowanie kątowe obrazów dyfrakcyjnych przez kliny zapewnia detekcję informacji niezależnej od skali oraz orientacji przestrzennych częstości. Wzrost obszarów półpierścieni od środka matrycy pomaga zrekompensować spadek intensywności, gdy wzrasta kąt rozproszenia.

³⁷ T. Niedziela, *Automatic target recognition from diffraction patterns*, "Prace Naukowe Politechniki Warszawskiej. Transport", z. 95/2013, s. 383–390.

³⁸ T. Niedziela, *Automatic target recognition of complex internal structure*, "Archives of Transport", 32/2014, 4, s. 61–72.

Rys. 9. Schemat ideowy 64-elementowego detektora pierścieniowo-klinowy (32 półpierścienie i 32 kliny)



Fotodetektor pierścieniowo-klinowy opracowany i opatentowany przez George'a i Wanga⁴² posiadał 32 półpierścienie i 32 kliny z wyraźnym spadkiem strumienia mocy (około 34,5 dB) na wyższych częstotliwościach przestrzennych, gdzie duży obszar wysokich częstotliwości jest zaletą tego detektora.

Do cyfrowych symulacji detektora pierścieniowo-klinowego popularne są dwie techniki: sumowanie binarne i sumowanie maskowe.

W *sumowaniu binarnym* każdy z obszarów częstotliwości jest traktowany jako osobny punkt i dane pierścieniowo-klinowe są obliczane przez sumowanie wartości wszystkich punktów próbnych w obrębie każdego indywidualnego obszaru próbkującego tzn. sumowane są wartości szarości dla binarnej reprezentacji sektora, w którym piksel jest zlokalizowany.

W *sumowaniu maskowym* istnieje szczególna maska (w formie macierzy zero jedynkowej). Te maski są mnożone przez transformatę Fouriera (kwadrat modułu) i następnie wszystkie wartości są sumowane. Sumowanie maskowe jest intensywne obliczeniowo, ale działa lepiej dla mniejszych obszarów obrazów. Symulacje cyfrową detektora pierścieniowo-klinowego z 32 pierścieniami i 32 klinami przeprowadzono jako szczególną macierz masek, w której każdemu pikselowi została przypisana liczba korespondująca z regionem, do którego należy. Każdy element macierzy zachowuje się jak indeks elementu wektora wyjściowego.

Wartości elementów wektora wyjściowego wzrastają z wartością szarości w macierzy intensywności transformaty Fouriera poprzez zgodność z „wykładnikami” wektora wyjściowego oraz „wartości” macierzy masek.

Przeprowadzono badania optymalizacji wymiarów cyfrowego detektora pierścieniowo-klinowego dla próbkowania transformaty Fouriera. Efekty cyfrowych błędów (artefaktów) zostały przeanalizowane w symulacji detektora pierścieniowo-klinowego o różnych powierzchniach. Wyjście detektora pierścieniowo-klinowego było analizowane w funkcji cyfrowych błędów w związku z pikselizacją detektora. Stwierdzono, że detektor pierścieniowo-klinowy o elementarnej powierzchni rzędu 300x300 pikseli jest wystarczający do otrzymywania niskich błędów cyfrowych w związku z pikselizacją elementu detekcyjnego.

Dla powyższych rezultatów przeprowadzono następujące analizy. Losowo elementy macierzy o różnych wymiarach były transformowane Fourierem. Wymiary macierzy po transformacji fourierowskiej (kwadrat modułu) były przechowane tak samo jak wymiary losowych macierzy. Detektor pierścieniowo-klinowy symulowany w postaci macierzy, miał zestaw pikseli zorganizowanych w kształcie pierścieni i klinów. Ustalono, że błędy cyfrowe mają znaczący wpływ na sygnał wyjściowy. Większa powierzchnia detektora zmniejsza błędy cyfrowe. Próbkowanie dwuwymiarowego kwadratu modułu transformaty Fouriera daje liniowe zależności dla klinowego histogramu.

Symulowano detektor pierścieniowo-klinowy o różnych powierzchniach elementów aktywnych (od 100x100 do 600x600 pikseli). Zaobserwowano, że *znormalizowane sygnały wyjście DPK* są gładziej dla dużych powierzchni elementów detekcyjnych w porównaniu z małymi powierzchniami detekcyjnymi. *Błąd odchylenia standardowego* sygnałów na wyjściu DPK w funkcji rozmiaru powierzchni spada gwałtownie, gdy wymiary elementu aktywnego DPK wzrastają (100x100 do 300x300 pikseli). Jednak, spadek błędu odchylenia standardowego nie przekracza granicy 300x300 pikseli. Zatem cyfrowy i praktycznie najlepszy DPK posiada elementarną powierzchnie detekcyjną w pobliżu 300 pikseli. W związku z powyższy cyfrowy detektor pierścieniowo-klinowy o wymiarach 300x300 pikseli był stosowany w eksperymencie.

Przeprowadzone symulacje cyfrowego DPK wykazały, że wykorzystanie wewnętrznych półpierścieni jako wejścia dla SSN nie jest przydatne jak również wykorzystanie wyższych częstotliwości przestrzennych również nie jest przydatne. Zatem detektory klinowe nie wykorzystywały danych z wnętrza elementu detekcyjnego

Błąd odchylenia standardowego sygnału wyjścia detektora pierścieniowo-klinowego malał w funkcji powierzchni elementarnej detektora.

W przeprowadzonym eksperymencie klasyfikator został zasymulowany w postaci SSN z wsteczna propagacja błędu. Była to trójwarstwowa SSN z 64 neuronami wejściowymi, 20 neuronami ukrytymi oraz 5 neuronami wyjściowymi. Funkcja transferu użyta od warstwy wejściowej do warstwy ukrytej

była sigmoidalna, natomiast funkcja transferu użyta od warstwy ukrytej do warstwy wyjściowej była liniowa. Obie warstwy, wejściowa i wyjściowa sieci były jednobiegunowe. Założono, że obrazy twarzy na wejściu sieci należą do jednych z pięciu osób, które mają zostać rozpoznane. Element współistniejący z pięciu elementów wektora wyjściowego powinien być albo „1” albo „0”. Przyjęcie tych jednobiegunowych wektorów wejściowych i wyjściowych powoduje, że system staje się bardziej realny do zaimplementowania optycznego.

Czterdzieści obrazów twarzy było użyte do badań, 20 do treningu sieci oraz 20 do testowania (rys. 10).

Rys. 10. Obrazy twarzy wykorzystane do treningu sieci neuronowych. Obrazy te mają różne pochylenia głowy oraz rozmiar³⁹



Zestaw treningowy



Zestaw testowy

Zestaw *treningowy* sieci neuronowych składał się z 20 obrazów twarzy (4 obrazów 5 osób). Obrazy 128x120 pikseli uzyskane zostały z bazy danych⁴⁰ i były o odwróconym kontraście (obrazy negatywne). Obrazy różniły się rozmiarem i nachyleniem głowy. Do testu *uogólnienia osoby* zostały użyte te same osoby, ale wielkość i nachylenie głowy różniły się od tych użytych do treningu. Przykładowo osoby z niewielkim nachyleniem głowy w lewo było wykorzystane do *treningu*, ale normalna twarz (bez przechylenia) była użyta do *testowania uogólnienia pierwszej osoby*.

³⁹ D. Ganotra, J. Joseph, K. Singh, *Neural network based face recognition by using diffraction pattern sampling with a digital ring-wedge detector*, "Optics Communications" 202, 61–68, 2002.

⁴⁰ N. George, S. Wang, *Neural Networks Applied to Diffraction-Pattern Sampling*. Appl. Opt. 33, 3127–3134, 1994.

Wyjście sieci neuronowej było wektorem z pięcioelementowym, każdy element należał do jednej osoby. Im więcej elementów było bliżej pierwszej osoby, tym lepiej sieć wykryła obraz wejściowy należący do danej osoby.

Parametry architektury sieci, przykładowo takie jak: stały moment, wskaźnik uczenia, liczba ukrytych neuronów, wartości macierzy wag, mają istotny wpływ na wydajność SSN. Optymalizacja tych parametrów jest niezbędna do prawidłowego działania sieci neuronowych. Jeżeli powyższe parametry zostaną dobrane niewłaściwie mogą prowadzić do negatywnych skutków.

Obrazy numerycznie transformowano do przestrzeni Fouriera oraz obliczano kwadrat modułu transformaty i konwertowano do obszaru o powierzchni 256x256 pikseli. Powyższy zestaw intensywności obrazów był próbkowany przez detektor pierścieniowo-klinowy. Każdy obraz intensywności generował 64-elementowy wektor po próbkowaniu przez detektor pierścieniowo klinowy. Tak otrzymywano zestaw 20 wektorów o 64 elementach każdy. Wartości pojedynczych elementów wektora różniły się co do wielkości. Podawanie tych danych jako wejście do sieci neuronowych sprawia, że rozpoznawanie staje się kłopotliwe.

W przeprowadzonym eksperymencie przyjęto wielkość średnią z wektorów i otrzymano 64 elementowy wektor średni. Każdy sygnał otrzymany z DPK był dzielony przez wartości średniego wektora. Przebadano sieć w funkcji liczby iteracji oraz wskaźników uczenia się. Zauważono, że efekty zmiany wskaźników uczenia się oraz liczby iteracji mogą być badane bardziej szczegółowo tylko, gdy ten sam zestaw losowych liczb zostanie wybrany dla początkowych wag. Eksperyment potwierdził łatwość użycia danych z próbkowania detektorem pierścieniowo-klinowego (DPK) do aplikacji sieci neuronowej dla celów rozpoznawania wielu twarzy.

Podsumowanie

Rozmiar elementu detekcyjnego, który może generować gładkie sygnały wyjściowe podczas próbkowania transformaty Fouriera (kwadratu modułu transformaty Fouriera) obrazów twarzy jest około 300x300 pikseli i jest optymalnym wymiarem do utrzymania niskich błędów cyfrowych.

Klasyfikacja intensywności obrazów twarzy w płaszczyźnie transformaty Fouriera może zostać z sukcesem przeprowadzona przez trzywarstwową sieć neuronową ze wsteczną propagacją błędów. Widmowe intensywności Fouriera można łatwo otrzymać przez symulacje komputerową a trenowanie i uogólnienia sztucznej sieci neuronowej oraz badana funkcji wskaźnika uczenia się i liczby iteracji pomagają w ocenie skuteczności optoelektronicznego systemu rozpoznawania twarzy.

Biorąc pod uwagę rolę techniki we współczesnych zaawansowanych systemach rozpoznania, automatyzacja wykrywania obiektów jest bardzo ważna. Metryczne falkowe cechy współwystąpień poziomów szarości wykorzystane w proponowanej implementacji mogą być odpowiednie do tego zadania. Zaproponowany algorytm może być skuteczny dla zaszumionych obrazów twarzy.

Dokonano oszacowania optymalnych wymiarów cyfrowego ekstraktora w postaci detektora pierścieniowo-klinowego dla próbkowania transformaty Fouriera przez macierzę fotodetektorów. Kwadrat modułu transformaty Fouriera funkcji obrazowej twarzy były próbkowane przez strukturę pierścieniowo-klinową i zostały użyte do trenowania i generalizacji sztucznej sieci neuronowej dla celów rozpoznawania wielu twarzy.

Próbkowanie obrazu dyfrakcyjnego detektorem pierścieniowym – klinowym oraz użycie sztucznej sieci neuronowych (SSN) prowadzi do tworzenia łatwych i szybkich algorytmów rozpoznawania obrazów twarzy.

Bibliografia

- Anila S., Devarajan N., *Global Journal of Computer Science and Technology Graphics & Vision Pre-processing Technique for Face Recognition Applications under Varying Illumination Conditions*, 2012.
- Boccignone G., Chianese A., Picariello A., *Using Renyi's information and wavelets for target detection: an application to mammograms*, *Pattern Analysis and Applications*, vol. 3, no. 4, pp. 303–313, 2000.
- Bureau of Justice Statistics*, U. S. Department of Justice, April 1990, pp. 43-66; Search Group, *Legal and Policy Issues Relating to Biometric Identification Technologies*.
- Caarls W., Jonker P., Corporaal H., *Smartcam: Devices for embedded intelligent cameras*, in *PROGRESS 2002, 3rd seminar on embedded systems*, Proceedings (Utrecht, The Netherlands).
- Cai D., He X., Han J., Zhang H.-J., *Orthogonal Laplacianfaces for Face Recognition*, "IEEE Transactions On Image Processing", 2006.
- Castellano G., Boyce J., Sandler M., *Moving target detection in infrared imagery using a regularized CDWT optical flow*, in *Proc. IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications (CVBVS '99)*, pp. 13–22, Fort Collins, Colo, USA, June 1999.
- Chang T., Kuo C.-C.J., *Texture analysis and classification with tree-structured wavelet transform*, "IEEE Trans. Image Processing", vol. 2, no. 4, pp. 429–441, 1993.
- Chen P.C., Pavlidis T., *Segmentation by texture using correlation*, "IEEE Trans. on Pattern Analysis and Machine Intelligence", vol. 5, no. 1, pp. 64–69, 1983.
- Cole M., *Algorithmic skeletons: structured management of parallel computations*, Glasgow 1989.
- Espinal F., Huntsberger T.L., Jawerth B.D., Kubota T., *Wavelet-based fractal signature analysis for automatic target recognition*, "Optical Engineering", vol. 37, no. 1, pp. 166–174, 1998.

- Fatemi H., Malek H.E., Kleihorst R., Corporaal H., Jonker P., *Real-time face recognition on a mixed SIMD VLIW architecture*, in *PROGRESS 2003, 4th seminar on embedded systems, Proceedings (Nieuwegein, The Netherlands)*, 22 October 2008.
- Ganotra D., Joseph J., Singh K., *Neural network based face recognition by using diffraction pattern sampling with a digital ring-wedge detector*, "Optics Communications" 202, 61–68, 2002.
- George N., Wang S., *Neural Networks Applied to Diffraction-Pattern Sampling*, "Appl. Opt." 33, 3127–3134, 1994.
- Haddadnia J., Faez K., Moallem P., *Human face recognition with moment invariants based on shape information*, in *Proceedings of the International Conference on Information Systems, Analysis and Synthesis, vol. 20 (Orlando, Florida USA)*, International Institute of Informatics and Systemics (ISAS), 2001.
- Haley G.M., Manjunath B.S., *Rotation-invariant texture classification using modified Gabor filters*, in Proc. IEEE International Conference on Image Processing (ICIP '95), pp. 262–265, Washington, DC, USA, October 1995.
- Haralick R.M., Shanmugam K., Dinstein I., *Textural features for image classification*, "IEEE Trans. Systems, Man, and Cybernetics", vol. 3, no. 6, pp. 610–621, 1973.
- He D.-C., Wang L., *Texture unit, texture spectrum, and texture analysis*, "IEEE Transactions on Geoscience and Remote Sensing", vol. 28, no. 4, pp. 509–512, 1990.
- Howard A., Padgett C., Brown K., *Real time intelligent target detection and analysis with machine vision*, in Proc. 3rd International Symposium on Intelligent Automation and Control, World Automation Congress (ISIAC-WAC '00), Maui, Hawaii, June 2000.
- Huntsberger T.L., Jawerth B.D., *Wavelet based algorithms for acoustic and non-acoustic anti-submarine warfare*, "Annual Tech. Rep., University Research Initiative Program for Combat Readiness", University of South Carolina, Columbia, SC, USA, 1998.
- Huntsberger T.L., Jawerth B.D., *Wavelet based automatic target detection and recognition*, Annual Tech. Rep., University Research Initiative Program for Combat Readiness, University of South Carolina, Columbia, SC, USA, 1998.
- Itti L., Gold C., Koch C., *Visual attention and target detection in cluttered natural scenes*, "Optical Engineering", vol. 40/2001, no. 9, pp. 1784–1793.
- Jain A., Moulin P., Ramchandran K., *Chernoff bounds for automatic target recognition from compressed data*, in Proc. IEEE Information Theory Workshop on Detection, Estimation, Classification and Imaging (DECI '99), p. 57, Santa Fe, NM, USA, February 1999. 2674 EURASIP Journal on Applied Signal Processing
- Kubota T., *Adaptive pixel-based data fusion for boundary detection*, in Proc. Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR '99), Lecture Notes in Computer Science, pp. 173–188, Springer-Verlag, York, UK, July 1999.
- Kubota T., Huntsberger T.L., *Adaptive pattern recognition system for scene segmentation*, "Optical Engineering", vol. 37, no. 3, pp. 829–835, 1998, Special section on advances in pattern recognition.
- Kubota T., Huntsberger T.L., Alford C.O., *A vision system with real-time feature extractor and relaxation network*, "International Journal of Pattern Recognition and Artificial Intelligence", vol. 12, no. 3, pp. 335–354, 1998.
- Kubota T., Huntsberger T.L., *Edge dipole and edge field for boundary detection*, "Hybrid Image and Signal Processing VI", vol. 3389 of Proceedings of SPIE, pp. 179–189, Orlando, Fla, USA, April 1998.

- Kubota T., Huntsberger T.L., Martin J.T., *Edge based probabilistic relaxation for sub-pixel contour extraction*, in Proc. 3rd International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR '01), pp. 328–343, INRIA, Sophia Antipolis, France, September 2001.
- Nautiyal J., Gahlot S., Kumar Mishra P., *An Automated Technique for Criminal Face Identification Using Biometric Approach*, Conference on Advances in Communication and Control Systems 2013 (CAC2S 2013).
- Nicolescu C., Jonker P., *A data and task parallel image processing environment*, "Lecture Notes in Computer Science", vol. 2131, pp. 393–408, 2001.
- Niedziela T., *Automatic target recognition from diffraction patterns*, "Prace Naukowe Politechniki Warszawskiej. Transport", z. 95/2013, s. 383–390.
- Niedziela T., *Automatic target recognition of complex internal structure*, "Archives of Transport", 32, 4/2014, s. 61–72.
- Niedziela T., *Metody falkowe analizy sygnałów*, „Logistyka” 6/2014, s. 7788–7796.
- Raghu P.P., Yegnanarayana B., *Segmentation of Gabor-filtered textures using deterministic relaxation*, "IEEE Trans. Image Processing", vol. 5, no. 12, pp. 1625–1636, 1996.
- Serot D., Derutin J., *Skipper: A skeleton-based programming environment for image processing applications*, in *Proceeding of the Fifth International Conference on Parallel Computing Technologies*, 1999.
- Sklansky J., *Image segmentation and feature extraction*, "IEEE Trans. Systems, Man, and Cybernetics", vol. 8, no. 4, pp. 237–247, 1978. Automatic Target Detection Using Wavelet Transform 2673.
- Tian Y., Qi H., Wang X., *Target detection and classification using seismic signal processing in unattended ground sensor systems*, in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP '02), vol. 4, p. 4172, Orlando, Fla, USA, May 2002.
- Unser M., Eden M., *Multiresolution feature extraction and selection for texture segmentation*, "IEEE Trans. on Pattern Analysis and Machine Intelligence", vol. 11, no. 7, pp. 717–728, 1989.
- Unser M., *Local linear transforms for texture measurements*, "Signal Processing", vol. 11, no. 1, pp. 61–79, 1986.
- Unser M., *Texture classification and segmentation using wavelet frames*, "IEEE Trans. Image Processing", vol. 4, no. 11, pp. 1549–1560, 1995.
- Van de Wouwer G., Scheunders P., Van Dyck D., *Statistical texture characterization from discrete wavelet representations*, "IEEE Trans. Image Processing", vol. 8, no. 4, pp. 592–598, 1999.