

Hafed ZGHIDI, Paweł PRUSKO  
Politechnika Śląska, Instytut Informatyki

## DETEKCJA TWARZY, WYBRANE TECHNIKI I METODY

**Streszczenie.** Celem niniejszej pracy jest przedstawienie metody rozwiązania problemu detekcji twarzy w obrazach barwnych opartej na informacji o kolorze odgrywającym najważniejszą rolę zarówno w czasie segmentacji pikseli koloru skóry, a także wyodrębnienia z twarzy prawdopodobnych obszarów, z których wyłonione zostaną oczy. Otrzymany w ten sposób zbiór wykrytych elementów cech twarzy poddaje się dalszej analizie, zgodnie z geometrią twarzy.

**Słowa kluczowe:** detekcja twarzy, wizja komputerowa, krawędzie, kolor, skala szarości, segmentacja

## FACE DETECTION, CHOSEN TECHNICS AND METHODS

**Summary.** The paper presents a method for solving face detection problem in colour images. The presented method is based on color information which plays a fundamental role during skin pixel segmentation, as well as during the detection of regions in which the eyes probably are in. The detected face elements are next analyzed in accordance with face geometry.

**Keywords:** face detection, computer vision, edges, color, grey-scale, segmentation,

### 1. Wstęp

Wizja komputerowa to obecnie prężnie rozwijająca się dziedzina, która odnotowuje coraz większy i dalej idący postęp w zakresie naśladowania ludzkiego wrodzonego systemu wzrokowego. Dąży się dziś do stawiania przed nią ambitnych celów o coraz bardziej ogólnym charakterze, do których należą detekcja twarzy, rozpoznawanie twarzy czy kodowanie sygnału wideo. Staje się nieoceniona w identyfikacji biometrycznej, inteligentnych interfejsach pomiędzy człowiekiem i komputerem, a także w dziedzinie wideokonferencji oraz systemach

do przeszukiwania na podstawie zawartości. Sięga się po nią także jako narzędzie do wstępnego przetworzenia obrazu celem wyodrębnienia i zlokalizowania obszaru pokrywającego twarz, by następnie przekazać tę wiedzę jako parametr wejściowy do systemu zajmującego się jej rozpoznawaniem. W detekcji twarzy za cel przyjmujemy zlokalizowanie obszarów obrazu, w których umiejscowione są ludzkie twarze, zakładając, że na takim obrazie w ogóle jakaś twarz się znajduje. Cały proces przetwarzania obejmuje kilka następujących po sobie etapów, takich jak segmentacja, wyodrębnienie, weryfikacja, lecz może również zostać rozszerzony i uwiarygodniony przez zastosowanie analizy cech twarzy. Chcielibyśmy również, aby strategia podejścia, którą obraliśmy, miała na uwadze oraz skutecznie radziła sobie z takimi czynnikami, jak zmienne warunki oświetlenia, pozycje twarzy, a także odległość od urządzenia rejestrującego, np. aparatu.

## 2. Ewolucja technik detekcji twarzy

Na początku lat 70. XX wieku, gdy w dziedzinie detekcji twarzy stawiano pierwsze kroki, jedynymi stosowanymi technikami była heurystyka i wykorzystanie wiedzy na temat antropometriki. Były one jednak tak bardzo rygorystyczne w swoich założeniach z powodu ograniczeń dotyczących frontalnych pozycji twarzy i gładkiego tła, że nawet drobna zmiana warunków otoczenia, w jakich zrobiono zdjęcie, niemal zawsze oznaczała konieczność jego reorganizacji lub przeprojektowania. Rozwój w dziedzinie detekcji i rozpoznawania twarzy, a co za tym idzie również i zainteresowanie w kręgach naukowych, utkwily więc w martwym punkcie aż do początku lat 90., kiedy to pojawiły się pierwsze, służące do tego celu systemy o praktycznym zastosowaniu. Efekty postępu ostatnich dwóch dekad widzimy dziś począwszy od wykorzystywania schematów segmentacji, w których niejednokrotnie uwzględnia się również ruch i kolor aż po sięgnięcie do statystyki czy sieci neuronowych, które z powodzeniem znalazły swoje zastosowanie na polu detekcji twarzy w obrazach ze złożonym, niejednorodnym tłem.

Wiedzę o budowie twarzy wykorzystywano w różny sposób, co dało początek dywersyfikacji na dwie obszerne kategorie technik ich detekcji. Pierwsza z nich składa się z technik działających na podstawie cech charakterystycznych twarzy, stąd też ich nazwa (ang. *feature-based*). Obejmują one na pierwszym miejscu etap analizy opartej na jawnej wiedzy o twarzy oraz postępują wedle tradycyjnej metodologii wyszukiwania na obrazie właściwości niskiego poziomu, wyodrębniając charakterystyczne jej cechy, jak oczy, usta czy nos. Są one następnie wzajemnie ze sobą zestawiane w przeróżnych konfiguracjach, do konstrukcji których używa się pomiaru odległości między nimi, kątów nachylenia czy też zajmowanych przez nie powierzchni. W przeciwieństwie do pierwszej kategorii, druga zakłada udział wiedzy o ele-

mentach twarzy w sposób niejawnny przez schematy treningowe oraz klasyfikacje za pośrednictwem algorytmów je wykorzystujących. Nie jesteśmy tu zainteresowani cechami twarzy, lecz pracujemy wyłącznie na dwuwymiarowej reprezentacji obrazu w postaci tablic wartości nasycenia barw (image-based).

### 3. Rodzaje technik detekcji twarzy

W klasycznym zadaniu detekcji twarzy, pierwszy krok najczęściej polega na analizie niskiego poziomu, co oznacza wykorzystanie własności obrazu, na przykład już na poziomie wartości nasycenia barw. Wizualne artefakty podlegają następnie analizie, w czasie której traktuje się je jako cechy twarzy, zgodnie z jej geometrią eliminując niejednoznaczności, jakie niesie przetwarzanie niskiego poziomu.

#### 3.1. Krawędzie

Krawędzie, uważane za jedną z najprostszych i najbardziej powszechnych cech analizowanych w dziedzinie wizji komputerowej, w odniesieniu do detekcji twarzy wykorzystywane są najczęściej jako narzędzie do wykrywania okularów. Dysponujemy dziś szeroką gamą operatorów, dzięki którym możliwe jest przystąpienie do znalezienia reprezentacji krawędziowej, z których na pewno warto wymienić operator Sobla oraz Marr-Hildreth. Istnieje jednak wiele metod, które opierają swoje działanie na pierwszej i drugiej pochodnej rozkładu Gaussa. Kiedy wykrywanie krawędzi jest zastosowane do detekcji twarzy, muszą one być indeksowane oraz dopasowane do modelu twarzy w celu weryfikacji wynikowego efektu detekcji. Przykładem takiej strategii jest podejście zastosowane przez Govindaraju [1], który oznaczał krawędzie jako leżące poza granicą włosów oraz po lewej lub prawej stronie widoku frontального twarzy. Następnie sprawdzał ich dopasowanie do modelu twarzy, używając zasady złotego podziału:

$$\frac{height}{width} = \frac{1 + \sqrt{5}}{2} \quad (1)$$

Oznaczone elementy krawędzi są łączone, tworząc możliwe konfiguracje umiejscowień charakterystycznych cech twarzy na podstawie funkcji kosztu. Algorytm ten poddano testom na 60 obrazach, zawierających 90 twarzy i uzyskano 76% skuteczności poprawnego wykrycia ze średnią dwóch fałszywych trafień na obraz.

### 3.2. Skala szarości

Pewne obserwacje obszarów pokrytych przez cechy ludzkich twarzy, jak brwi, źrenice czy usta, pozwoliły na odróżnienie części twarzy za sprawą ciemniejszej ich barwy niż obszarów je otaczających. W ostatnich latach powstały algorytmy, które potrafią taką informację wykorzystać w celu znalezienia lokalnego minimum oraz lokalnego maksimum w obrębie segmentowanego obrazu. Aby poprawić jakość występujących na nim ciemniejszych obszarów i umożliwić później zastosowanie procesu detekcji twarzy, obraz poddaje się najpierw operacjom morfologicznym w skali szarości oraz stosuje się inne metody, jak rozciągnięcie kontrastu (ang. *contrast-stretching*). Ciemniejsze obszary są następnie wyodrębniane przez progowanie. Przykładem niech będzie tu algorytm zaprezentowany przez Hoogenbooma oraz Lewa i in.[2], którzy definiują maksimum lokalne jako jasny piksel w otoczeniu 8 ciemniejszych, co może zostać zinterpretowane jako czubek nosa. Znalezione w ten sposób punkty są porównywane z szablonami cech twarzy dla dokonania korelacji pomiarów.

### 3.3. Kolor

W porównaniu do skali szarości kolor dzięki swojej wielowymiarowości rzuca zupełnie inne światło na wygląd rozpoznawanego obiektu na obrazie. Może się teraz okazać, że dwa kształty obiektu o tej samej wartości nasycenia w skali szarości zajmują całkowicie inną część przestrzeni kolorów. Do zastosowań typowych dla detekcji twarzy używa się różnych modeli kolorów. Może to być najbardziej powszechny RGB czy HSI, YIQ, HSV, YES, YUV, a także YCbCr. Co ciekawe, odkryto bardzo osobliwą właściwość koloru ludzkiej skóry, która umiejscawia go w spójnym i ograniczonym klastrze przestrzeni kolorów. Dzięki prowadzonym badaniom, obserwacjom oraz testom wydaje się być to prawdą bez względu na przynależność do jakiegokolwiek grupy etnicznej. Wnioskujemy z tego zatem, że kolor naszej skóry nie różni się aż tak bardzo w jego składowej koloru. W literaturze na temat ludzkiego systemu wzrokowego podejmowany jest temat zjawiska, wedle którego ludzkie oczy reagują w większym stopniu na zmiany w nasyceniu światła niż koloru. Rozważmy w tym miejscu przykład Y.Dai wraz z Y.Nakano [3] oraz G.Wei wraz z I.K. Sethi [4], którzy w swojej pracy nad detekcją twarzy wykorzystali przestrzeń kolorów YIQ. Zauważyli oni, że po konwersji z RGB do YIQ składnik I nadawał się do dużo lepszego przedstawienia wyglądu koloru skóry u osób reprezentujących społeczność Azji. Warto zwrócić uwagę, że jest to komponent, który odpowiada za kolory z zakresu pomiędzy pomarańczowym a morskim. W efekcie zastosowania konwersji między wspomnianymi przestrzeniami kolorów inne barwy obecne w tle obrazu są skutecznie blokowane.

Ponieważ poleganie jedynie na metodach niskiego poziomu może zdominować wyniki przez niejednoznaczność ich interpretacji, w literaturze znajdujemy dwa główne podejścia do

rozwiązania problemu weryfikacji wyodrębnionych elementów ludzkiej twarzy. Pierwsze z nich określa się mianem „wyszukiwania cech” i odnosi się do schematu, wedle którego pierwszorzędne znaczenie ma odnalezienie cechy charakterystycznej, odznaczającej się w sposób szczególny w porównaniu do innych. Takiej charakterystyce i opisowi odpowiada na przykład para oczu lub owal twarzy. To właśnie na jej podstawie opiera się poszukiwanie następnych, do czego służą pomiary antropometryczne, mając na uwadze strukturę geometryczną twarzy. Dobrze rozwiązanie w tej kategorii zaproponował De Silva i in.[7], który rozpoczyna swój algorytm od hipotetycznego wyznaczenia górnej granicy twarzy. Posuwając się w dół, spodziewa się napotkać płaszczyznę oczu, co sygnalizować będzie zagęszczona obecność krawędzi, a jednocześnie jej odległość od założonej wcześniej górnej linii twarzy stanie się odległością odniesienia. To ona daje początek utworzeniu elastycznego szablonu nałożonego na obraz wejściowy, a pokrywającego oczy oraz usta. Początkowo jego kształt otrzymuje się przez korelację pomiędzy odległością odniesienia a odległością antropometryczną. Jego końcowa postać to efekt dostrojenia przez funkcję kosztu bazującą na krawędziach. Algorytm ten osiąga skuteczność 82% dla twarzy w widoku frontalnym oraz  $\pm 30^\circ$  na gładkim tle. Sprawdzono, że radzi sobie dobrze dla przedstawicieli różnych ras, gdyż nie używa właściwości koloru, lecz może zawieść w przypadku osób noszących okulary lub mających czoło zasłonięte przez włosy.

### 3.4. Modele aktywne kształtu

Aktywne modele kształtu dostosowują swój, początkowo naniesiony kształt powoli, odwzorowując prawdziwy obiekt na obrazie przez oddziaływanie z jego lokalnymi właściwościami, jak krawędzie czy jasność pikseli.

Wyróżniamy trzy rodzaje aktywnych modeli: węże, szablony deformowalne oraz inteligentne węże. W kontekście detekcji twarzy, pierwszy z nich, czyli węże bądź aktywne kontury, przydaje się do wykrywania zarysu i brzegów głowy. Bardzo istotną rolę po zdecydowaniu się na tę metodę jest właściwe zainicjalizowanie węża w ścisłym sąsiedztwie owalu głowy. W kolejnych etapach będzie on naśladował kształt obiektu, coraz bardziej go odwierciedlając. Idea przyświecająca tej metodzie jest związana z równaniem przedstawiającym energię węża, a ściślej mówiąc z jej minimalizacją.

$$E_{Snake} = E_{Internal} + E_{External} \quad (2)$$

Energia wewnętrzna, której rolę pełnić może na przykład energia sprężystości, definiuje zachowanie, takie jak kurczenie się czy rozszerzanie. Z kolei energia zewnętrzna, która zależy od typu właściwości obrazu, pozwala na założenie pewnego kształtu obiektu w stanie równowagi przez dopuszczenie odstępstw od jego naturalnej ewolucji, co tak naprawdę sprowadza się do przeciwdziałania energii wewnętrznej. Może ona na przykład reagować na gra-

dient, przyciągając kontur do krawędzi. Mimo iż uważa się ją za metodę bardzo skuteczną, aktywne kontury mają też swoje wady. Pierwszą z nich jest blokowanie na niepożądanych obiektach obrazu, drugą jest znana nieumiejętność poradzenia sobie z obiektami o niewypukłej strukturze, gdyż z definicji dążą do utrzymania minimalnej krzywizny.

Warto choć w kilku słowach wspomnieć o metodzie PDM (ang. *point distribution models*). Opracował ją i po raz pierwszy zastosował Lanitis i in. [10] jako model elastyczny. Przeprowadzając badania na zbiorze 160 obrazów, otrzymał go przez ręczne rozmieszczenie 152 punktów kontrolnych. Aby dopasować PDM do twarzy, model twarzy średniej musi najpierw zostać ulokowany w pobliżu twarzy. Następnie, stosując lokalne wyszukiwanie w skali szarości, każdy z punktów przysuwany jest do odpowiadającego mu punktowi granicznemu. Dzięki przetworzeniu zbioru treningowego model ten nabywa umiejętności konsekwentnego, elastycznego dopasowywania. Do głównych zalet PDM należy jej sparametryzowany opis, dobre wyniki w przypadku częściowo zakrytych obiektów dzięki znajdowaniu globalnego optymalnego rozwiązania, a włączając współpracę z algorytmami genetycznymi nadaje się także do rozstrzygnięcia problemu wielu kandydatów twarzy.

Jedna z nowszych metod, opracowana przez Moghaddam i Pentland [11], bazuje na technice PCA, którą udoskonaili w porównaniu do jej pierwszych wersji. Mogła ona teraz wziąć pod uwagę przestrzeń twarzy oraz jej ortogonalne uzupełnienie, a nie jak poprzednio jedynie pierwsze z nich. Gdyby nie ta zmiana, przestrzeń twarzy musiałaby zakładać jednolitą gęstość. W przypadku detekcji lewego oka na zbiorze 7000 obrazów przedstawiono wynik w 95% pozytywny. Metodę tę zaopatrzoną w analizę koloru oraz ruchu zaimplementowano również w systemach do śledzenia ludzkich twarzy.

By uzupełnić metody, które w historii badań nad detekcją twarzy miały istotne znaczenie, należy zaznaczyć wpływ algorytmów bazujących w swym działaniu na sieciach neuronowych. Nowatorska architektura do uczenia w systemach detekcji twarzy opracowana została przez Rotha i in.[12]. W skrócie nazywana jest SNoW (ang. *Sparse network of windows*) i stanowi ją sieć neuronowa wyposażona w dwie liniowe jednostki progowe (LTU). Jednostki te, które w zamierzeniu mają reprezentować wspomniane klasy twarzy i nietwarzy, operują na przestrzeni wejściowej elementów logicznych. System charakteryzujący się najlepszymi osiąganiami może wykrywać cechy w następujący sposób {pozycja  $x$ , średnie nasycenia  $\bar{x}$ , wariancja nasycenia} dla podokien  $1 \times 1$ ,  $2 \times 2$ ,  $4 \times 4$ ,  $10 \times 10$ . Proces trenowania sprowadzony został do prostego uczenia, w którym odpowiednio przyporządkowane wagi są promowane bądź degradowane, w zależności od tego czy wynik wykrycia jest poprawny czy nie.

### 3.5. Metody statystyczne

Do detekcji twarzy wypróbowano również SVM (ang. *support vector machine*). Swoje rozwiązania zaprezentowali tu między innymi Osuna [13] oraz Kumar i Poggio[14], którzy

posłużyli się tym rozwiązaniem przy analizie, a także śledzeniu twarzy w czasie rzeczywistym.

#### 4. Kolor jako istotny czynnik w detekcji twarzy

Przedstawiona w niniejszym artykule metoda detekcji twarzy zainspirowana została wynikami badań nad właściwościami koloru ludzkiej skóry oraz pracą ludzi nauki, takich jak Rein-Lien Hsu, Mohamed Abdel-Mottaleb i Anil K. Jain [1]. Proponowane rozwiązanie może być stosowane do detekcji twarzy w wielu okolicznościach i warunkach charakteryzujących się zróżnicowanym oświetleniem, a także dużym dopuszczalnym zakresem pozycji twarzy. Kluczowym zagadnieniem stanowiącym niemałe wyzwanie jest wyodrębnienie obszarów skóry modelując jej kolor za pomocą elipsy parametrycznej w płaszczyźnie 2D przestrzeni kolorów YCbCr. Z uwagi na pomysł wykorzystania właściwości koloru, efektywność tej metody jest z góry ściśle uzależniona od warunków oświetlenia, które w przypadku ich zmienności mogą stanowić trudną do pokonania czy też ominięcia przeszkodę. Ponieważ nigdy nie można liczyć na obecność jednolitego oświetlenia, prawie zawsze dostrzec można efekt śladów poświaty żółtawej lub też innego koloru użytego źródła światła. Są one w stanie zmodyfikować oryginalną barwę pikseli koloru skóry do tego stopnia, że przestają ją przypominać, a zatem jeśli większą część twarzy stanowią piksele zmienione z powodu wrogich efektów iluminacji, może to prowadzić do poważnych problemów najpierw ze znalezieniem obszarów-kandydatów do miana regionów twarzy, jak i później ze znalezieniem samej twarzy. Staje się to rozważaniem szczególnie ważnym, gdy w grę wchodzi wartość przedstawiana przez kolor, jak ma to miejsce w niniejszej pracy. Fakt ten czyni kompensację oświetlenia jednym z najbardziej intrygujących problemów fazy wstępnej algorytmu tuż przed detekcją obszarów twarzy. Zainteresowanie nakreślonym tematem zaowocowało rozwojem sposobów radzenia sobie z przedstawionym niechcianym i wrogim w swych konsekwencjach zjawiskiem. Idąc w ślad za opinią zaczerpniętą z pracy autorstwa Terillona i in. dotyczącej segmentacji obrazu na podstawie koloru, najkorzystniej jest przeprowadzić ją w znormalizowanej przestrzeni kolorów. Ma to również zapewnić lepsze wyniki niż w przypadku dokonania tego na innych, nieznormalizowanych przestrzeniach. W niniejszej pracy wykorzystano znormalizowaną przestrzeń kolorów RGB. Z kolei, jako narzędzie do rozwiązania problemu zagadnienia kompensacji oświetlenia wykorzystano wariację znanego algorytmu Grey World, którą zaproponował Phil Chen. Zwyczajowo algorytm GW przedstawia się w formie:

$$S_c = \frac{C_{std}}{C_{avg}} \quad (3)$$

gdzie  $S_c$  rozumiany jest jako współczynnik skali dla poszczególnego kanału koloru oznaczonego indeksem  $c$ .

Wartości poprawione przez algorytm GW otrzymuje się mnożąc znormalizowane wartości pikseli przez współczynnik skali dla odpowiedniego kanału. Zakłada się, że  $C_{std}$  stanowi zwykle 50% idealnej szarości, co jednak w obecności dużego obszaru tła pokrytego ciemnym kolorem prowadzi do przesylenia jasności obiektów na pierwszym planie. Zauważył to i wziął pod uwagę Phil Chen, który zmodyfikował wzór na optymalną wielkość  $C_{std}$ , czego efekt prezentuje się następująco:

$$C_{std} = \frac{\sum_1^m [\max(N_R, N_G, N_B) + \min(N_R + N_G + N_B)]}{2n} \quad (4)$$

$$n = m - \sum_1^m (N_R = N_B = N_G = 0) \quad (5)$$

gdzie  $m$  to liczba pikseli na obrazie,  $n$  - liczba pikseli koloru innego niż czarny,  $C_{avg}$  to średnia pikseli koloru innego niż czarny obecnych w poszczególnych kanałach.

#### 4.1. Modelowanie koloru skóry

Chcąc wyodrębnić interesujące nas obszary z obrazu, potrzebna jest przestrzeń kolorów, która da nam możliwość zauważenia, jak różnią się one w swojej reprezentacji od innych na nim obecnych. Śladami autorów, których praca posłużyła tu za inspirację, decydujemy się na wybór przestrzeni YCbCr, która w swej reprezentacji rozdziela kanały na luminancję i składowe kolorów. Definiuje dzięki temu również model podobny do tego, w jaki sposób człowiek postrzega kompozycję koloru oraz jasności. Niemalą rolę odgrywa przy tym wybo-rze fakt, że można za jego pomocą zaobserwować spójny klaster koloru ludzkiej skóry. Za pomocą próbek kolorów pochodzących z 9 tematów, a przeanalizowanych w Instytucie Heinricha-Hertz (HHI) zaobserwowano, że wartości odpowiadające odcieniom ludzkiej skóry mogą być z powodzeniem opisane przy użyciu modelu eliptycznego, który służy jako granica obszaru obejmującego ich największe zagęszczenie.

Model eliptyczny dla odcieni koloru ludzkiej skóry zdefiniowany jest następująco:

$$\frac{(x - ec_x)^2}{a^2} + \frac{(y - ec_y)^2}{b^2} = 1 \quad (6)$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} C_b - c_x \\ C_r - c_y \end{bmatrix} \quad (7)$$

Rozważany sposób modelowania należy rozumieć jako proces klasyfikacji, który powinien pozwolić na wyodrębnienie obszaru pokrytego kolorem ludzkiej skóry. Jednak koloru takiego mogą również być inne objekty, znajdujące się na analizowanym obszarze. Obraz z tak wyodrębnionymi łatanami poddaje się następnie binaryzacji, po której obszary ludzkiej

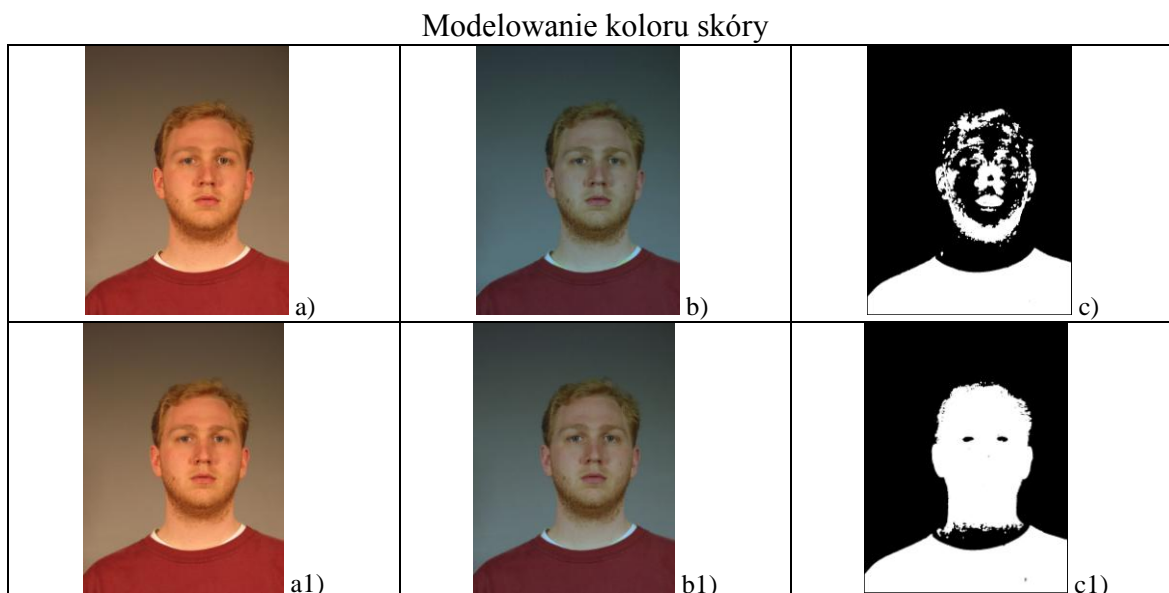


skóry oznacza się na białą, a tło blokuje się wypełniając je kolorem czarnym. W odniesieniu do prezentowanego modelu eliptycznego rozumieć możemy to jako promowanie pikseli znajdujących się wewnątrz elipsy oraz degradowanie wszystkich poza nią, jako posiadających kolor inny niż spodziewany odcień skóry. Aby dodatkowo usprawnić proces przekształcania kolorów, transformację z jednej jego przestrzeni w drugą reprezentację w postaci znormalizowanej możemy go dodatkowo zmodyfikować, stosując przekształcenie nieliniowe. W zamierzeniu ma ono poprawić procent liczby pikseli właściwie sklasyfikowanych jako reprezentujących kolor ludzkiej skóry. Jego postać oraz stała zawarta we wzorze wyznaczone zostały eksperymentalnie i przetestowane na zdjęciach przedstawiających twarze.

$$Nc = \frac{C}{C + a} \tag{8}$$

gdzie stałej C wyznaczona została wartość 410.

Tabela 1



W tabeli 1 kolejne obrazy oznaczają: a, a1) obrazy oryginalne; b, b1) obrazy po kompensacji oświetlenia bez stosowania przekształcenia nieliniowego; c, c1) efekty modelowania koloru skóry.

Na zamieszczonych obrazach widać wyraźną różnicę w interpretacji koloru skóry, według której wiele pikseli zostało niewłaściwie sklasyfikowanych. Gdyby wspomniane przekształcenie nie zostało użyte, część z nich, mimo iż przedstawia twarze, zostałyby w konsekwencji tego efektu pominięta w dalszej analizie.

#### 4.2. Segmentacja obszarów twarzy

Znalezione w poprzedniej fazie algorytmu łaty traktowane są dalej jako hipotetyczne obszary zawierające twarz. Aby móc w sposób unikalny odróżnić je od siebie, przeprowadza się

segmentację obrazu z wykorzystaniem indeksowania elementów połączonych z liczbą połączeń równą 8 (ang. *connected component labelling with 8-connectivity*). Tylko te wyodrębnione obszary staną się przedmiotem następnych etapów, kiedy poddane zostaną analizie w celu stwierdzenia obecności cech charakterystycznych twarzy. Dwuwymiarowa tablica pikseli skanowana jest dwukrotnie w kierunku od strony lewej do prawej i z góry do dołu. W pierwszym przebiegu w celu zidentyfikowania połączonych z sobą pikseli analizowane jest sąsiedztwo każdego z nich z osobna, jeden po drugim w celu odszukania nowych segmentów bądź przypisania ich do istniejących oraz oznaczenia części punktów biorących udział w analizie jako równoważne, w przypadku gdy dany piksel sąsiaduje z więcej niż jednym obszarem istniejącym. To właśnie ta zależność będzie przedmiotem drugiego przebiegu, w którym etykiety zostaną pogrupowane w klasy, z których każda otrzyma jedną unikalną, co zapewni wszystkim równoważnym indeksom tę samą wartość.

### 4.3. Konstruowanie map cech twarzy

Konstruowana mapa cech twarzy przyjmuje parę oczu za swój kluczowy atrybut, co z uwagi na jednoznaczne przez nich zlokalizowanie twarzy jest powodem częstego ich wyboru do tego celu. Rozumowanie Hsu, Mottaleb oraz Jain powinno pozwolić na bezpośrednie odnalezienie umiejscowienia oczu, opierając się na ich charakterystycznej i osobliwej kombinacji składników koloru oraz luminancji. W tym celu konstruowane są dwie niezależne mapy oczu, jedna bazująca na informacji zawartej w składowej koloru i druga, opierająca swe założenia na znaczeniu natężenia oświetlenia.

### 4.4. Konstruowanie mapy oczu na podstawie składowej koloru

Dzięki obserwacji wykazującej wokół oczu większe wartości składnika  $C_b$  oraz mniejsze wartości  $C_r$  zrodził się pomysł na stworzenie pierwszej ze wspomnianych map, którą opisać można przy użyciu następującego wzoru:

$$MapaOczuChroma = \frac{1}{3} \left[ (C_b^2)_+ + (\bar{C}_r^2)_+ + (C_b / C_r) \right] \quad (9)$$

gdzie:  $C_b^2$ ,  $C_r^2$  i  $C_b/C_r$  przedstawiają wartości znormalizowane do zakresu (0, 255), a  $\bar{C}_r$  jest rozumiane jako uzupełnienie  $C_r$  (tj.  $\bar{C}_r = 255 - C_r$ ).  $\bar{C}_r$  do stworzenia mapy oczu ze składowej luminancji potrzebne będzie zrozumienie działania morfologicznych operatorów: dylatacji i erozji.

### 4.5. Konstruowanie mapy na podstawie luminancji

Wykorzystanie wspomnianych operatorów pozwoliło na realizację pomysłu stworzenia mapy, używając bezpośrednio składowych luminancji danego obrazu, a jego zwarty zapis matematyczny przedstawia się następująco:





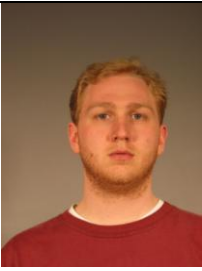

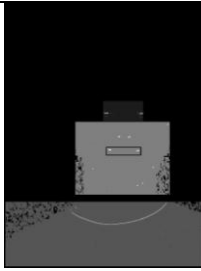

$$MapaOczuLuma = \frac{I(x, y) \oplus g_{\sigma}(x, y)}{I(x, y) \ominus g_{\sigma}(x, y) + 1} \quad (10)$$

gdzie:  $\oplus$   $\ominus$  – oznaczają odpowiednio operacje dylatacji i erozji na obrazie w skali szarości jako funkcje  $f : F \subset R^2 \rightarrow R$  z użyciem elementu strukturalnego w postaci funkcji  $g : G \subset R^2 \rightarrow R$ , których definicje znaleźć można w [15].

Nadeszła pora, aby połączyć obie mapy w jedną, mnożąc wartości jednej przez wartości drugiej (AND), co w efekcie pozwoli zauważyć wyraźną wskazówkę co do lokalizacji oczu. Zanim jednak do tego dojdzie, mapa powstała z komponentu koloru zostaje poddana procesowi wyrównania histogramu (ang. *histogram equalization*). Obraz, będący ostatecznym rezultatem tego etapu, jest następnie normalizowany, progowany oraz ponownie poddany operacjom morfologicznym, aby w jak największym stopniu podkreślić obecność pożądaných efektów oraz zablokować niechciane.

Tabela 2

Mapy oczu

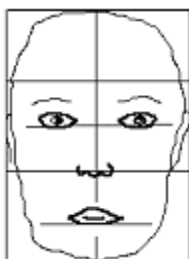
 a)	 b)	 c)	 d)
 a1)	 b1)	 c1)	 d1)

W tabeli 2 kolejne obrazy oznaczają: a,a1) obraz oryginalny, b,b1) mapa oczu z komponentu kolorów, c,c1) mapa oczu z komponentu luminancji, d,d1) efekt końcowy po połączeniu i operacjach uzupełniających.

#### 4.6. Odrzucenie nieprawdziwych cech

Białe obiekty, które pozostały na obrazie po poprzednim etapie, to tak naprawdę to, czego oczekiwaliśmy, z wyjątkiem jednego zauważalnego efektu. Metoda ta jest tak pomocna i użyteczna w przypadku oczu, ponieważ wykorzystuje rozkład kolorów zaobserwowanych wokół nich. Sęk w tym, iż nie tylko oczy charakteryzują się taką dystrybucją kolorów w swoim sąsiedztwie, dlatego używając jej zawsze należy wziąć pod uwagę ewentualność znalezienia jeszcze innych obiektów. Z tymi niechcianymi, mogącymi reprezentować każdy element rzeczywistego wyglądu, nie tylko w obrębie twarzy, ale również innych części ciała oraz ubrania, jak na przykład guziki na marynarce, trzeba sobie zatem jakoś poradzić. Najlepiej gdy obiekty te leżą w narożnikach analizowanego obrazu, wtedy można je od razu wykluczyć, bo to eliminuje je ze zbioru potencjalnych kandydatów do sklasyfikowania jako oczu. W trudniejszych przypadkach weryfikacja ich autentyczności jako oczu wymagać będzie bardziej dogłębnej analizy. Czasem zdarza się, że proporcje, które normalnie prowadzą algorytm do wskazania prawdziwych oczu, sprowadzają go do bliskiego sąsiedztwa nosa. Nie jest to sytuacja, z którą poradzenie sobie można zaliczyć do rzeczy łatwych, niemniej jednak zagadnienie to jest jednym z podejmowanych dalej rozważań, w których tematem jest odrzucenie błędnie znalezionych obiektów cech twarzy, a zaakceptowanie właściwych jako znalezionych poprawnie i odpowiadających prawdziwym oczom.

Zakładając, iż miejsce położenia twarzy zostało znalezione i otoczone prostokątem ograniczającym, istnieje duża szansa na poprawne wskazanie rejonu, w którym można spodziewać się umiejscowienia oczu. W niniejszym artykule zastosowano podejście, zgodnie z którym stopniowo odrzuca się obiekty typowane jako niebędące poszukiwanymi oczami. Proces rozpoczyna się od obrazu zawierającego obszary prawdopodobnych twarzy otoczonych prostokątami wraz ze znalezionymi w poprzedniej fazie całego algorytmu, białymi łatami oznaczającymi hipotetyczne obiekty oczu. Biorąc pod uwagę skuteczność metody, za pomocą której modelowany jest kolor skóry oraz dokładność wyznaczenia rejonów twarzy, zakłada się, że wyznaczając punkt przecięcia przekątnych takiego rejonu, będzie on leżał w samym sercu obszaru obejmującego charakterystyczne cechy ludzkiej twarzy. Biorąc pod uwagę jej rzeczywiste proporcje, możliwe jest zawężenie obszaru poszukiwań. Aby zachować tę właściwą, ludzką twarz, zwykle dzieli się na sześć równych kwadratów przylegających do siebie i tworzących prostokąt o wymiarach 2x3 takich właśnie kwadratów. Górny bok głównego prostokąta jest styczny do wierzchołka głowy, podczas gdy dolny bok przylega do podbródka. Oczekuje się, że oczy znajdować się będą w środkowej części kwadratów rozmieszczonych horyzontalnie. Jest to oczywiście, jedynie model twarzy w jej frontalnej pozycji, a zatem należy wziąć pod uwagę zmiany w położeniu oczu w przypadku jej odchylenia bądź obrótu. Założonym celem algorytmu jest umiejętność poprawnej detekcji twarzy w pozycji frontalnej oraz półprofilu, a także przechylonej lub odchylonej w bok o kąt do 35°.



Rys. 1. Model proporcji twarzy

Fig. 1. Face proportions model

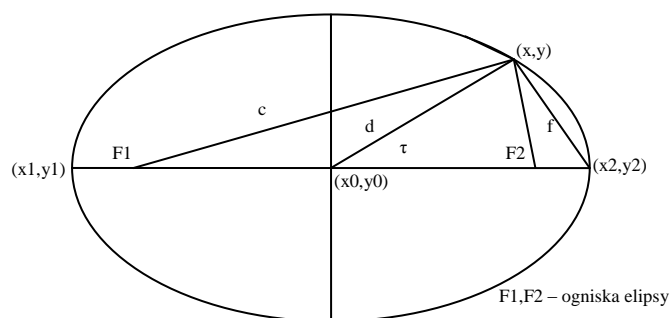
Proporcje zauważone na modelu twarzy posłużyły tu do wytłumaczenia i wizualizacji, dlatego postanowiono wykorzystać punkt przecięcia się przekątnych jako punkt odniesienia. Przypuszcza się, że wyznaczając współrzędne tego punktu, jego bliskie sąsiedztwo obejmuje też rzeczywiste oczy. W tym miejscu należy uświadomić sobie i założyć prawdopodobną obecność także innych obiektów wykrytych na etapie konstrukcji mapy oczu. Można do nich zaliczyć znalezione obiekty, takie jak znaki szczególne, znamiona na twarzy, dziurki w nosie lub też elementy zupełnie niezwiązane z twarzą i niebędące jej wrodzoną, integralną częścią. Tę jakże delikatną i istotną kwestię rozwiązuje się począwszy od wyznaczenia odległości każdego ze znalezionych obiektów do określonego punktu odniesienia. Bez względu na to, jak wiele zostało ich wykrytych, do dalszej analizy kwalifikuje się co najwyżej cztery położone najbliżej punktu odniesienia. Oczywiście, przypadek skrajny występowania zaledwie jednego obiektu przedstawiającego oko jest od razu odrzucany, co natychmiast staje się również warunkiem koniecznym dla regionu twarzy, aby mogła w dalszej części algorytmu zostać uznana za twarz. Dalsza analiza zależy od tego, czy znaleziono 2, 3, 4 lub więcej obiektów/kandydatów do zostania sklasyfikowanymi jako oczy. W przypadku gdy twarz obejmuje 4 lub więcej obiektów ubiegających się o to miano, zakłada się górną granicę w liczbie 4, co jest najbardziej skomplikowanym przypadkiem w porównaniu do liczby 2 czy 3 obiekty, ponieważ te ostatnie dwa wymagają przeanalizowania mniejszej ilości możliwości i ograniczają tym samym liczbę kombinacji. Decyzję dobrania górnej granicy w liczbie 4 obiektów tłumaczy się obserwacją, zgodnie z którą w większości przypadków jest to wystarczająca ilość w sąsiedztwie punktu odniesienia, aby odnaleźć wśród nich rzeczywiste oczy. Główną motywacją przy dokonywaniu tego wyboru była ewentualność wzięcia pod uwagę przypadkowo zakwalifikowanych dziurek od nosa, które zabrałyby dwa miejsca z czterech. W następnym kroku wyznacza się kąty nachylenia linii przechodzących przez każdą możliwą parę oczu. W ten sposób wyłonione zostają wstępnie dwie pary oczu o najmniejszym spośród wszystkich nachyleń, z których na końcu jedna zostanie wybrana jako przedstawiająca parę prawdziwych oczu, a druga odrzucona. Warunki sformułowane dla oceny tych dwóch par pod względem ich autentyczności i wytypowania jednej jako pary prawdziwych oczu zostały wyznaczone na podstawie proporcji zaprezentowanego modelu twarzy, po przeanalizowaniu kilku zbiorów zawierających ich zdjęcia. Wśród nich najważniejsze to tangens kąta ich nachylenia

względem siebie nie większy niż 0,7, co zapewnić ma możliwość przechylenia głowy o maksymalnie  $35^\circ$ . Oprócz tego odległość dzieląca oczy została dostosowana do średniego rozmiaru obszaru twarzy znalezionej po analizie całej bazy zdjęć. Wyniki w postaci współrzędnych oczu zostały następnie zestawione z wyznaczonymi niezależnie wartościami koordynatów, które dołączone są do zbioru zdjęć twarzy. Zgodnie z dostarczonymi danymi, na tym etapie wynik prawidłowej detekcji plasował się w granicach 80% w przypadku bazy zdjęć zatytułowanej „Notre Dame”. Jednakże po analizie otrzymanych wyników wraz z odpowiadającymi im mapami oczu stało się jasne, iż część twarzy nie została wykryta, gdyż etap detekcji oczu zawiódł z powodu obiektów leżących poza owalem twarzy, do których proces na tym etapie był zbieżny. Obiekty te były wzięte pod uwagę, ponieważ ich odległości od punktu odniesienia były na tyle niewielkie, aby kwalifikowały się jako 4 najbliższe. Lekarstwem na problem tej natury, zapobiegającym tym nieoczekiwanym i wielce niepożądanym efektem, jest algorytm do detekcji elips, który ograniczy obszar, wewnątrz którego obiekty oczu mają prawo występować, a poza którym są odrzucane. Jest to tym razem dużo bardziej restrykcyjne założenie niż prostokątny obszar ograniczający cały rejon twarzy.

#### 4.7. Wykrywanie elips

Za jeden z najbardziej znanych i efektywnych sposobów detekcji elips w obrazach cyfrowych uważa się Uogólnioną Transformatę Hougha. Jednakże z uwagi na jej duże wymagania pamięciowe, jak i obliczeniowe, w artykule tym opisano inną metodę według pomysłu Yonghong Xie i Quiang Ji [16]. Opracowali oni nową metodę detekcji elips, zgodnie z którą wykorzystuje się dłuższą oś elipsy, aby możliwie jak najszybciej zasugerować jej parametry. Nie wymaga przeprowadzania skomplikowanych operacji matematycznych i potrzebuje zaledwie jednowymiarowego akumulatora, w którym za pomocą mechanizmu „głosowania” gromadzi się informacje o długości krótszej osi.

Aby opisać elipsę, potrzebnych jest pięć parametrów: współrzędne środka elipsy  $(x_0, y_0)$ , długości dłuższej i krótszej osi  $a$  i  $b$ , a także kąt nachylenia.



Rys. 2. Elipsa i jej charakterystyczne parametry

Fig. 2. Characteristic ellipse parameters

Do zastosowania tego algorytmu należy przygotować przetworzony obraz przedstawiający wykryte na nim krawędzie. Można do tego celu posłużyć się operatorem detekcji krawędzi Canny. Ponieważ złożoność algorytmu znacznie wydłuża czas jego działania w przypadku dużej ilości wykrytych pikseli krawędzi, zmniejszenie rozmiarów obrazu jest jednym ze sposobów skrócenia go, nie tracąc przy tym ważnych informacji zawartych na obrazie.

W trakcie działania algorytm stosuje założenie, że współrzędne każdej pary pikseli  $(x_1, y_1)$  oraz  $(x_2, y_2)$  należących do wykrytych wcześniej krawędzi (jeśli ich odległość jest większa lub mniejsza od przyjętej wartości progowej), reprezentuje współrzędne położenia dłuższej z osi elipsy. To wstępne założenie pozwala na obliczenie wartości czterech z pięciu potrzebnych parametrów:

$$x_0 = \frac{(x_1 + x_2)}{2} \quad (11)$$

$$y_0 = \frac{(y_1 + y_2)}{2} \quad (12)$$

$$a = \frac{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}}{2} \quad (13)$$

$$\alpha = a \tan\left(\frac{(y_2 - y_1)}{(x_2 - x_1)}\right) \quad (14)$$

gdzie  $(x_0, y_0)$  to współrzędne środka elipsy, a rozumiane jest jako połowa długości dłuższej osi, a  $\alpha$  to kąt nachylenia/inklinacji. Piąty parametr zostanie wybrany w następnej fazie algorytmu. Zachowując znalezione dotychczas wstępne wartości parametrów, następuje drugi przebieg, w którym dla każdego punktu należącego do krawędzi obliczana jest jego odległość od środka elipsy jako długości krótszej z osi. Posłużą do tego następujące wzory:

$$b^2 = \frac{(a^2 d^2 \sin^2 \tau)}{(a^2 - d^2 \cos^2 \tau)} \quad (15)$$

$$\cos \tau = \frac{(a^2 + d^2 - f^2)}{(2ad)} \quad (16)$$

gdzie  $d$  stanowi odległość pomiędzy środkiem elipsy a jednym z przetwarzanych w danej chwili punktów krawędzi.

W wyniku tych operacji otrzymujemy pewną wartość, dla której zwiększany jest licznik w akumulatorze. Na końcu największa wartość przechowywana w akumulatorze sugeruje prawdopodobną długość krótszej osi dla elipsy o założonych wcześniej parametrach, pod warunkiem że jest większa od założonej wcześniej wartości progowej. Drugi przebieg powtarzany jest dla każdej założonej długości dłuższej osi elipsy wyznaczonej w pierwszej części algorytmu. Opisane podejście daje wymierne rezultaty, trzeba jednak zauważyć obecność wykrywanych tą metodą błędnych wykryć. Potrzebny jest zatem sposób weryfikacji i sprawdzenia, czy za otrzymanymi parametrami kryje się elipsa odzwierciedlająca autentyczny obiekt na obrazie, który można by scharakteryzować także owalnym kształtem. Polega on na

nałożeniu elipsy o wyznaczonych parametrach na obraz zawierający obszary prawdopodobnych twarzy w postaci binaryzowanej. Obecność prawdziwej elipsy uważa się za potwierdzoną, jeśli co najmniej 70% obejmowanych przez nią pikseli należy do obszaru wyznaczonego we wcześniejszych etapach całego algorytmu jako regionu twarzy. W przypadku gdy elipsa o tak sformułowanych warunkach nie może zostać wykryta, region otaczający owal twarzy zostaje rozszerzony z bardziej dokładnej elipsy do mniej dokładnego prostokąta, obejmującego cały obszar twarzy.

## 5. Testy

Dwa różne zbiory zdjęć zostały przetestowane według algorytmu opisanego w tym artykule wraz z postawionymi w nim założeniami. Pierwsza z nich, Notre Dame, liczy sobie 215 zdjęć twarzy w rozdzielczości 852x1136 przedstawiających twarzy osób należących do różnych etnicznie grup. Druga, AR, zawierała 620 obrazów o rozdzielczości 768x576 z osobami prezentującymi różne wyrazy twarzy, a w pewnych przypadkach również noszącymi szalik częściowy zakrywający twarz bądź okulary. Jak wspomniano wcześniej, obie bazy zdjęć wyposażono w zbiór plików w formacie xml, zawierających dane referencyjne wskazujące prawdziwe położenie oczu, a które wyznaczone zostały ręcznie. Dzięki temu możliwe stało się obliczenie względnego błędu między wynikiem otrzymanym po zaimplementowaniu prezentowanego algorytmu a wartościami referencyjnymi. Do tego celu stosuje się następujący wzór:

$$err = \frac{|\max[d(lEye_1, lEye_2), d(rEye_1, rEye_2)]|}{dist(lEye_2, rEye_2)} \quad (17)$$

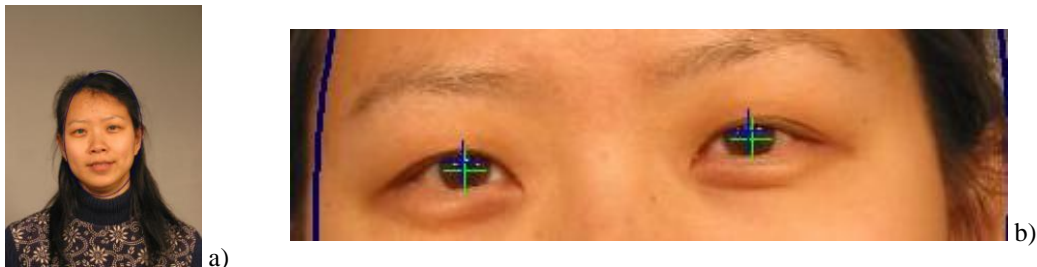
gdzie:  $lEye_1$ ,  $rEye_1$  oznaczają lokalizację lewego i prawego oka wyznaczone przez program podczas gdy  $lEye_2$ ,  $rEye_2$  analogicznie oznaczają wartości odczytane z plików danych referencyjnych.

Twarz uważa się za wykrytą, jeśli obliczony błąd nie jest na tyle duży, aby programowo wyznaczona lokalizacja znajdowała się poza źrenicą zarówno lewego, jak i prawego oka. Zgodnie z postawionymi założeniami algorytm charakteryzuje się skutecznością 89,3% na testowanym zbiorze 215 obrazów i 73,7% w przypadku 620 zdjęć drugiego zbioru testowego. Na poniższych przykładowo wybranych zdjęciach testowych lokalizację wykrytych oczu uwidoczniono przez oznaczenie ich niebieskim krzyżykiem, podczas gdy dane referencyjne wskazywane są krzyżykiem zielonym. Elipsa wokół owalu twarzy jest efektem wynikowym opisanego wcześniej algorytmu, służącego do jej wykrycia.



## 6. Przykłady detekcji

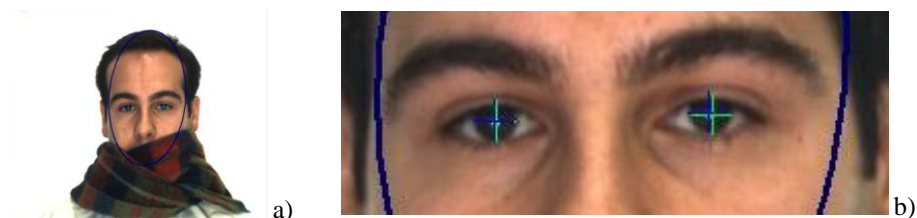
Przykłady prawidłowej detekcji przedstawiono na rysunkach:



Rys. 3. Przykład detekcji: a) oryginał, b) wynik detekcji  
 Fig. 3. Detection example: a) original image, b) detection result



Rys. 4. Przykład detekcji: a) oryginał, b) wynik detekcji  
 Fig. 4. Detection example: a) original image b) detection result



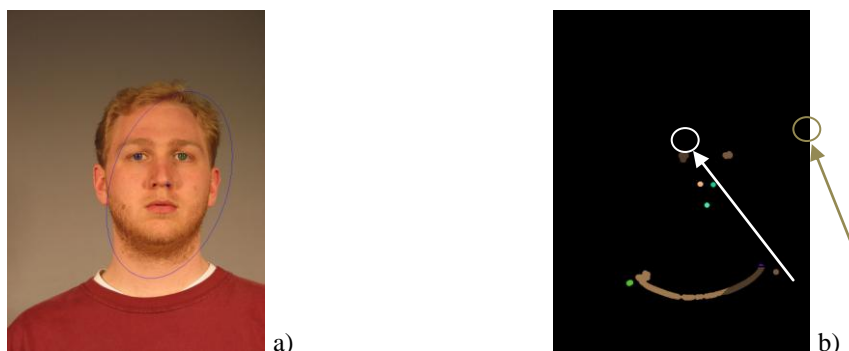
Rys. 5. Przykład detekcji: a) oryginał, b) wynik detekcji  
 Fig. 5. Detection example: a) original image b) detection result

Przykłady błędnej detekcji przedstawiono na rysunkach:



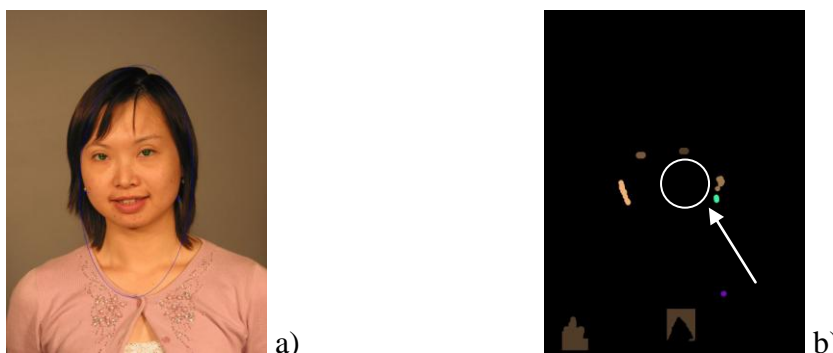
Rys. 6. Przykład błędnej detekcji: a) oryginał, b) wynik detekcji  
 Fig. 6. Erroneous detection example: a) original image b) detection result

Rysunek 6b przedstawia mapę oczu ilustrującą małżowinę uszną jako jedną z przyczyn, które mogą sprawić, że algorytm zawiódł.



Rys. 7. Przykład błędnej detekcji: a) oryginał, b) wynik detekcji  
 Fig. 7. Erroneous detection example: a) original image, b) detection result

Rysunek 7b przedstawia przyczynę niepowodzenia zidentyfikowania jako dziurki nosa.



Rys. 8. Przykład błędnej detekcji: a) oryginał, b) wynik detekcji  
 Fig. 8. Erroneous detection example: a) original image b) detection result

Na rys 8b pokazano prawy kolczyk jako przykład przeszkody na drodze do poprawnego wyznaczenia prawdziwej pary oczu.

Oryginały obrazów pobrano z bazy zdjęć Notre Dame. Wykorzystane zdjęcia to: 04202s134.jpg (rys. 3), 04211d155.jpg (rys. 4), 04202d154.jpg(rys. 6), 04212d199.jpg (rys. 7), 04217d203.jpg(rys. 8)

Baza zdjęć AR: M\_001\_1.jpg (rys. 5).

Najczęstsze przyczyny niepowodzenia algorytmu zostały przedstawione, gdy usiłowano zdiagnozować przyczyny takich zachowań. Do zaobserwowanych w trakcie testów należy głównie obecność kolczyków w uszach, a także specyficzne przypadki okolic małżowiny usznej i dziurek od nosa. Rzadziej spotkać się można z niepożądanym efektem wywołanym nieregularnym kształtem części twarzy, w miejscu gdzie przylega do niej ucho.

## 7. Podsumowanie oraz sugerowany kierunek dalszego rozwoju

W artykule zademonstrowano sposób rozwiązania problemu detekcji twarzy w obrazach barwnych. Jego idea oparta jest na informacji o kolorze odgrywającym najważniejszą rolę

zarówno w czasie segmentacji pikseli koloru skóry, a także wyodrębnienia z twarzy prawdopodobnych obszarów, z których wyłonione zostaną oczy. Otrzymany w ten sposób zbiór wykrytych elementów cech twarzy poddaje się dalszej analizie, zgodnie z geometrią twarzy, co w efekcie pozwala na otrzymanie obiecujących rezultatów. Należy jednak pamiętać o ograniczeniach narzuconych na same obrazy, warunkujących jego poprawną analizę oraz detekcję. W obecnej wersji użytkownik może oczekiwać obrazu wynikowego zawierającego co najwyżej jedną twarz wykrytą. Dużo uwagi poświęcono efektywnemu radzeniu sobie z niepożądanymi efektami oświetlenia, a także umożliwieniu poprawnej detekcji dla szerokiego zakresu pozycji głowy, a zatem i twarzy. Granice te ustalono jako rozciągające się od pozycji całkowicie pionowej do odchylonej o  $35^\circ$ . Zakłada się również, że oczy osób znajdujących się na zdjęciach są widoczne.

Dalszy rozwój i udoskonalenie algorytmu skłania się w kierunku wyodrębnienia twarzy z bardziej złożonego tła, a także podejmowania decyzji o liczbie znalezionych twarzy w trakcie działania programu.

## BIBLIOGRAFIA

1. Yang J., Weibel A.: A real-time face tracker. IEEE Proc. of the 3rd Workshop on Applications of Computer Vision, Florida, 2006.
2. Chen P.: A Fast Skin Region Detector. ESC Division Research, 2005.
3. Jack K.: Video Demystified: A Handbook for the Digital Engineer, 4th Edition. Butterworth-Heinemann, August 2004.
4. Fisher R., Perkins S., Walker A., Wolfart E.: Hypermedia Image Processing – internet sources. 2003.
5. Gonzalez R. C., Woods R. E.: Digital Image Processing – Second Edition. Prentice Hall, Upper Saddle River, New Jersey 2002.
6. Govindaraju V.: Locating human faces in photographs. Int. J. Comput. Vision, 1996.
7. Hoogenboom R., Lew M.: Face detection using local maxima. IEEE Proc. of 2nd Int. Conf. on Automatic Face and Gesture Recognition, 2000.
8. Dai Y., Nakano Y.: Face-texture model based on sgltd and its application. Pattern Recognition, 1996, s. 1007÷1017.
9. Wei G., Sethi I. K.: Face detection for image annotation. Pattern Recognition Lett, 1999, s. 1313÷1321.
10. Reisfeld D., Wolfson H., Yeshurun Y.: Context-free attentional operators: The generalized symmetry transform. Int. J. Computer Vision, 1995, s. 119÷130.

11. Reisfeld D., Yeshurun Y.: Robust detection of facial features by generalised symmetry. Proc. of 11th Int. Conference on Pattern Recognition, The Hague, The Netherlands, August 1992.
12. De Silva L. C., Aizawa K., Hatori M.: Detection and tracking of facial features by using a facial feature model and deformable circular template. IEICE Trans. Inform. Systems E78-D(9), 1995, s. 1195÷1207.
13. Burl M. C., Leung T. K., Perona P.: Face localization via shape statistics. Int. Workshop on Automatic Face and Gesture Recognition, Zurich, Switzerland, June 1995.
14. Gunn S. R., Nixon M. S.: A dual active contour for head and boundary extraction. IEE Colloquium on Image Processing for Biometric Measurement, London, April 1994.
15. Lanitis A., Taylor C. J., Cootes T. F.: Automatic tracking, coding and reconstruction of human faces, using flexible appearance models. IEEE Electron Lett, 1994, s. 1578÷1579.
16. Moghaddam B., Pentland A.: Probabilistic visual learning for object representation. IEEE Trans. Pattern Anal. Mach. Intell, 1997.
17. Roth D., Yang M.-H., Ahuja N.: A SnoW-based face detector. Advances in Neural Information Processing Systems 12 (NIPS 12), MIT Press, Cambridge, MA, 2000.
18. Osuna E., Freund R., Girosi F.: Training Support Vector Machines: An application to face detection. IEEE Proc. of Int. Conf. on Computer Vision and Pattern Recognition, 6, 1997.
19. Kumar V., Poggio T.: Learning-based approach to real-time tracking and analysis of faces. Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 2000.
20. Jackway P. T., Deriche M.: Scale-Space Properties of the Multiscale Morphological Dilation-Erosion. IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 18, No. , January 1996, s. 38÷51.
21. Yonghong Xie and Qiang Ji. A new efficient ellipse detection method.
22. Hsu R. L., Abdel-Mottaleb M., Jain A. K.: "Face detection in colour images". IEEE Transactions on Pattern Analysis and Machine Intelligence, May 2002.

Recenzent: Dr inż. Krzysztof Fajarewicz

Wpłynęło do Redakcji 21 października 2010 r.

## Abstract

The paper represents another approach to detecting faces. Information of colour components play the first fiddle in segmentation pixels to human skin and non-skin colour classes

and extracting eye candidates from within probable face regions. They are then manipulated with according to rules governing human face geometry to finally give pretty decent and accurate results. This paper addresses the idea of colour usage to ephasise its essensial role to detecting faces. It then gets slightly developed by involving information of facial features geometry and combined with shape of human face to form autonomous toll for finding faces in colour images. The obtained results can be thought of as positive effect of the presented method.

### **Adresy**

Hafed ZGHIDI: Politechnika Śląska, Instytut Informatyki, ul. Akademicka 16,  
44-100 Gliwice, Polska, hafed.zghidi@polsl.pl.

Paweł PRUSKO: Politechnika Śląska, Instytut Informatyki, ul. Akademicka 16,  
44-100 Gliwice, Polska, pawel.prusko@gmail.com.