

Segmentacja obrazów kolorowych wzorcami nieregularnymi

Witold ŻORSKI, Paweł SAMSEL

Zakład Automatyki, Instytut Teleinformatyki i Automatyki WAT,
ul. Kaliskiego 2, 00-908 Warszawa

STRESZCZENIE: Artykuł porusza zagadnienie rozpoznawania obiektów na obrazach kolorowych, z pominięciem etapu ekstrakcji cech. Na etapie przetwarzania wstępnego, jedynym poprzedzającym segmentację, dokonywana jest kwantyzacja kolorów, mająca na celu uniknięcie problemów związanych ze stosowaniem pełnej przestrzeni barw RGB. Proces segmentacji wykorzystuje uogólnioną transformatę Hougha w formie narzędzia rozpoznawania obiektów nieregularnych. Zagadnieniu kwantyzacji kolorów poświęcono szczególną uwagę, gdyż ma ono decydujące znaczenie dla jakości rozpoznawania. Wybór konkretnej techniki kwantyzacji powinien być podyktowany charakterem analizowanych, w danym systemie widzenia komputerowego, obrazów wejściowych.

SŁOWA KLUCZOWE: widzenie komputerowe, rozpoznawanie obrazów, segmentacja, transformata Hougha, grupowanie, kwantyzacja kolorów.

1. Wprowadzenie

Systemy widzenia komputerowego są dość złożonymi strukturami sprzętu i oprogramowania, które w celu realizacji postawionego im zadania dokonują przetwarzania obrazu wejściowego na kilku etapach [10]. Centralnym etapem jest segmentacja, zazwyczaj poprzedzona przez przetwarzanie wstępne i ekstrakcję cech. Ten etap należy do analizy danych i daje podstawę do wnioskowania, którego rezultatem jest klasyfikacja i opis. Analiza danych i wnioski należą do podsystemu analizy obrazu, który uzupełniany jest przez podsystem rozpoznawania obrazu, dokonujący reprezentacji symbolicznej. Efektem końcowym działania systemu widzenia komputerowego jest zazwyczaj interpretacja i opis.

W niniejszym opracowaniu kluczowe będą dwa etapy: przetwarzanie wstępne i segmentacja. Ze względu na przyjęte założenie, iż procesowi segmentacji będą poddawane bezpośrednio obrazy kolorowe, zaistniała konieczność przeprowadzenia odpowiednio zaawansowanej procedury przetwarzania wstępnego, która koncentruje się na problemie kwantyzacji kolorów. Narzędziem do przeprowadzenia procesu segmentacji będzie uogólniona transformata Hougha.

Jest w tym miejscu niemal rzeczą nieuniknioną napisać, że transformata Hougha została opatentowana w 1962 roku, jako technika wykrywania wzorców na obrazach binarnych. Podstawowym założeniem była możliwość wyznaczenia zbioru parametrów opisujących szukane wzorce. W roku 1981 spostrzeżono [3], że opracowana przez Hougha metoda jest szczególnym przypadkiem transformaty Radona [16] znanej od 1917 roku. Pierwotnie rozważane były problemy segmentacji krzywymi analitycznymi (proste, okręgi, elipsy), ale w dość naturalny sposób rozszerzono ten zakres na obiekty nieregularne, dokonując tym samym uogólnienia transformaty [1]. Nastąpiło jednocześnie przedefiniowanie przestrzeni parametrów, gdyż określa ona najczęściej położenie i orientację obiektów, a nie same wzorce, które to w przypadku uogólnionym stanowią zbiór pikseli. Pierwotnie, ze względu na branie pod uwagę zagadnienia i złożoność obliczeniową, rozpatrywane były tylko obrazy binarne. Nie stoi jednak nic na przeszkodzie, aby rozpatrywać również obrazy w poziomach szarości, co jest szczególnie istotne ze względu na niedopuszczalną obecnie utratę informacji w procesie binaryzacji. Należy też dodać, że opracowano interesujące warianty transformaty Hougha: szybka [13], hierarchiczna [14], randomizowana [17], probabilistyczna [11], rozmyta [5]; mające przede wszystkim na celu obniżenie złożoności obliczeniowej. Dzięki niezwyklejmu rozwojowi techniki komputerowej realne stały się implementacje transformaty Hougha o dużej złożoności obliczeniowej i znacznym zapotrzebowaniu na pamięć dla przestrzeni parametrów [19], [22].

W dalszej części przedstawiona jest propozycja segmentacji prowadzonej bezpośrednio na obrazach kolorowych, będącej naturalnym rozwinięciem prac prowadzonych w zakresie uogólnionej transformaty Hougha [20], [23].

2. Uogólniona transformata Hougha

Jak już zostało wspomniane, pierwotnie rozważane były problemy segmentacji obrazów krzywymi analitycznymi, ale w roku 1980 dokonano uogólnienia [1] transformaty Hougha na wzorce nieregularne. Uogólnienie polegało na potraktowaniu wzorca jako zbioru (listy) pikseli $\{(x_i, y_i)\}, i = 1, \dots, N$; natomiast zbiór parametrów został określony poprzez

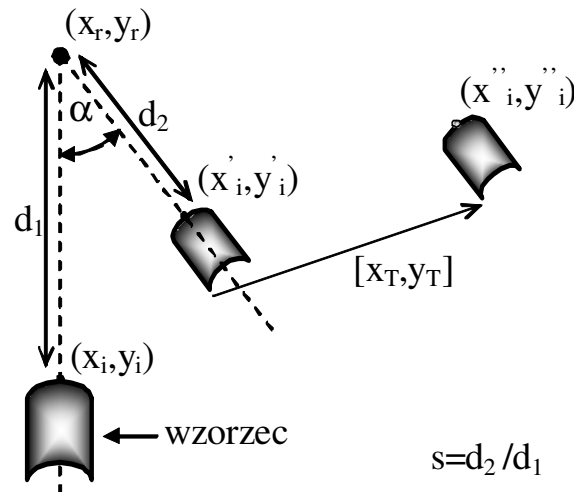
operacje: obrotu, przesunięcia i skalowania. Dokładniejsze wprowadzenie do zagadnienia można znaleźć w [9], a w tym miejscu zostanie przytoczona jedynie docelowa postać transformaty Hougha dla obrazów w poziomach szarości zaproponowana wcześniej w [18].

Transformata Hougha $H(x_T, y_T, \alpha, s)$, dla obrazu danego w poziomach szarości $I_G(x, y)$, w procesie identyfikacji wzorca M_P , zostaje zdefiniowana w następującej postaci:

$$H(x_T, y_T, \alpha, s) = \sum_{(x_i, y_i) \in M_P} h(x_i, y_i, x_T, y_T, \alpha, s), \quad (1)$$

$$h(x_i, y_i, x_T, y_T, \alpha, s) = |I_G(x''_i, y''_i) - M_P(x_i, y_i)|, \quad (2)$$

$$\begin{cases} x''_i = x_r + s(x_i - x_r)\cos(\alpha) - s(y_i - y_r)\sin(\alpha) + x_T \\ y''_i = y_r + s(x_i - x_r)\sin(\alpha) + s(y_i - y_r)\cos(\alpha) + y_T \end{cases} \quad (3)$$

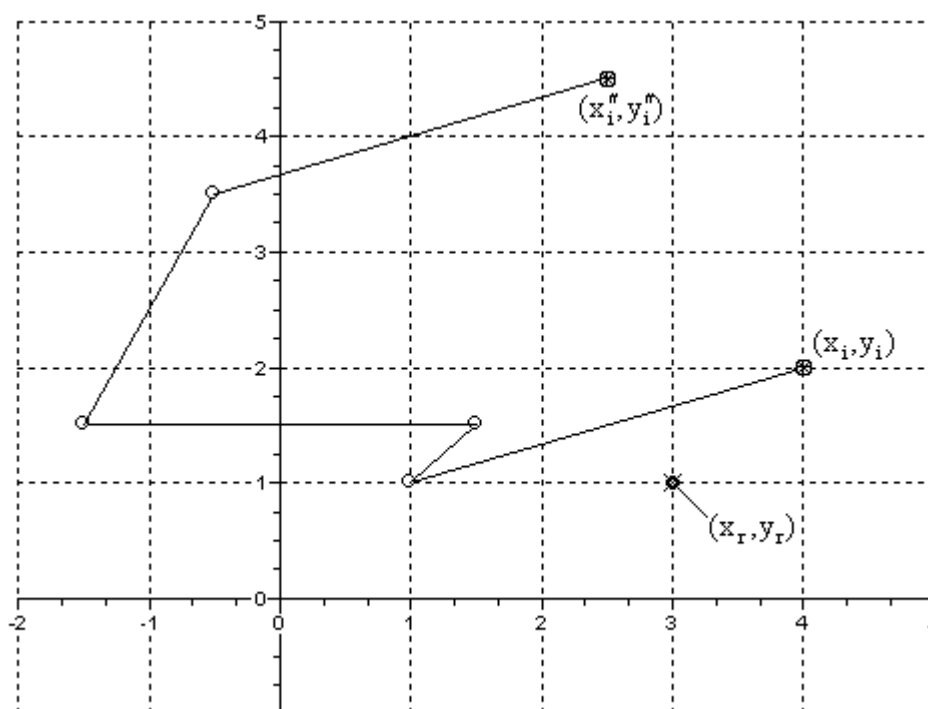


Rys. 1. Obrót, skalowanie i translacja wzorca względem arbitralnego punktu (x_r, y_r)

Wzór (3) można też zapisać w równoważnej postaci macierzowej:

$$\begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix} = \begin{bmatrix} s \cos(\alpha) & -s \sin(\alpha) & -sx_r \cos(\alpha) + sy_r \sin(\alpha) + x_T + x_r \\ s \sin(\alpha) & s \cos(\alpha) & -sx_r \sin(\alpha) - sy_r \cos(\alpha) + y_T + y_r \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}, \quad (4)$$

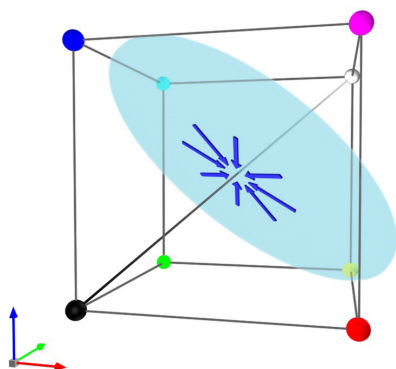
w której jawnie ukazany jest operator odpowiedzialny za złożenie operacji obrotu, skalowania i translacji względem arbitralnego punktu (x_r, y_r) . Należy pamiętać, że tylko w wyjątkowym przypadku, gdy $s_x = s_y = s$, operacje obrotu i skalowania mogą być wykonane w dowolnej kolejności. Znaczenie czynników we wzorze (4) zostało zobrazowane na rys. 2, gdzie zilustrowano złożenie omawianych operacji (dla: $s = 1.5$, $\alpha = \pi/2$, $x_T = 1$, $y_T = 2$).



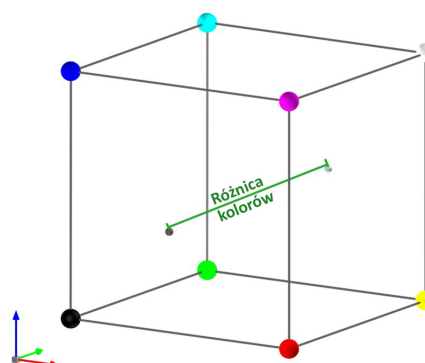
Rys. 2. Skalowanie, obrót i translacja dla punktu (x_i, y_i) względem arbitralnego punktu (x_r, y_r)

3. Problem obrazów kolorowych

Najprostszym sposobem uporania się z obrazami kolorowymi jest ich wstępna konwersja do obrazów w poziomach szarości, ale takie postępowanie, choć w większości przypadków wystarczające i skuteczne, może niekiedy przysporzyć poważnych problemów. Należy sobie zdawać sprawę z tego, że konwersja dokonywana jest na drodze rzutowania danego punktu sześcianu RGB na jego przekątną (rys. 3a), co może prowadzić do sytuacji utożsamienia pikseli o tej samej jasności, ale różnych składowych.



Rys. 3a. Konsekwencje konwersji do poziomów szarości



Rys. 3b. Odległość pikseli obrazu kolorowego

Wykorzystanie pełnej informacji o kolorach wymaga modyfikacji wzoru (2) do postaci uwzględniającej składowe pikseli obrazu kolorowego. Przykładowo, wykorzystując metrykę euklidesową (rys. 3b) oraz przyjmując: $I_C(x_i, y_i) = (r_i, g_i, b_i)$, $M_P(x_i, y_i) = (r_p, g_p, b_p)$, rozważany wzór można przepisać do postaci:

$$h(x_i, y_i, x_T, y_T, \alpha, s) = \sqrt{(r_i - r_p)^2 + (g_i - g_p)^2 + (b_i - b_p)^2} \quad (5)$$

Modyfikacja wzoru (2) jest poprawna, lecz konieczność wykonania potęgowania i pierwiastkowania będzie przyczyną nieakceptowalnego wzrostu czasu obliczeń. Powstaje zatem problem odszukania takiego rozwiązania, które zniesie powyższą niedogodność. Można tutaj pokusić się o zmianę metryki z euklidesowej na „miejską”:

$$h(x_i, y_i, x_T, y_T, \alpha, s) = |r_i - r_p| + |g_i - g_p| + |b_i - b_p|, \quad (6)$$

co jednak nie poprawia dostatecznie sytuacji, gdyż pozyskanie składowych kolorów pikseli obrazu i wzorca jest czasowo kosztowne, a całość nadal nieakceptowalna względem prostej operacji (znalezienia „odległości” poziomów szarości) reprezentowanej wzorem (2).

Ponieważ liczebność zbioru kolorów dla obrazów rastrowych jest skończona, więc można zaproponować wstępne tablicowanie odległości dla wszystkich możliwych par pikseli. Najczęściej mamy do czynienia z obrazami TrueColor, których głębina kolorów jest 24-bitowa, co niestety prowadzi do tablicy o zawrotnej liczbie 281474976710656 komórek.

Rozwiązaniem jest zastosowanie procesu kwantyzacji kolorów. Można łatwo wykazać, że kwantyzacja kolorów nawet dla obrazów złożonych kolorystycznie, przy zachowaniu rozsądnej liczby reprezentantów, nie powoduje znacznej (obserwowalnej) utraty jakości obrazu. Zauważmy, że redukcja palety kolorów, przykładowo do 1000, stwarzałaaby możliwość wyznaczenia tablicy odległości dla wszystkich kombinacji pikseli, gdyż miałyby ona milion elementów, co jest całkowicie dopuszczalne dla współczesnych komputerów.

Można odnieść wrażenie, że problem jest rozwiązany, ale tak niestety zdecydowanie nie jest. Szkopuł tkwi w tym, że system widzenia komputerowego nie przetwarza tylko jednego obrazu i powstaje problem ustalenia wspólnej palety kolorów, dla której można by wyliczyć tablicę ich odległości. Problem ten daje się jednak ominąć na dwa sposoby. Pierwszym się na podział jednorodny przestrzeni RGB i wyznaczenie dla niej stałego zbioru reprezentantów. Drugie podejście zakłada, że obrazy przetwarzane w danym systemie widzenia komputerowego mają podobne histogramy i wówczas dla odpowiednio dużej populacji takich obrazów można dokonać kwantyzacji kolorów, która nie musi być jednorodnym podziałem przestrzeni RGB.

Zakładając dalej, że możliwe jest przeprowadzenie kwantyzacji kolorów obrazów wejściowych, możliwe staje się też wyznaczenie tablicy D odległości kolorów-reprezentantów. Można zaproponować następujący odpowiednik funkcji (2), dla przypadku obrazów kolorowych:

$$h(x_i, y_i, x_T, y_T, \alpha, s) = D[\tilde{I}_C(x_i'', y_i''), \tilde{M}_P(x_i, y_i)], \quad (7)$$

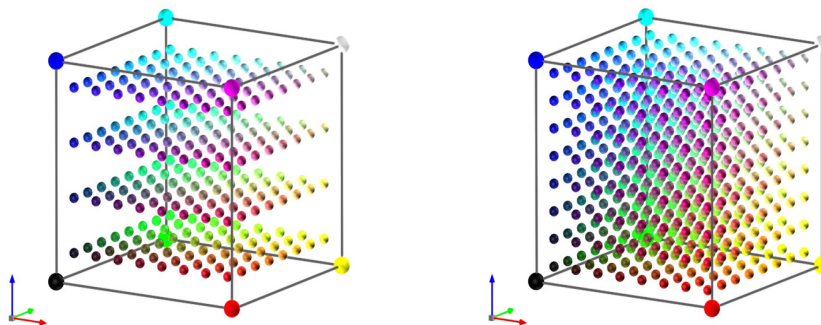
gdzie \tilde{I}_C , \tilde{M}_P oznaczają odpowiednio obraz I_C i wzorzec M_P po procesie kwantyzacji. Wartości $\tilde{I}_C(x_i'', y_i'')$, $\tilde{M}_P(x_i, y_i)$ pełnią rolę indeksów do tablicy D . Zatem proces obliczeń sprowadza się do odczytu wskazanej tymi indeksami komórki tablicy D .

4. Kwantyzacja obrazów

W części tej przedstawione są cztery podstawowe techniki grupowania, które mogą być wykorzystane na potrzeby kwantyzacji kolorów.

4.1. Metoda jednorodnego podziału przestrzeni RGB

Jest to jedna z najprostszych technik, z zaprezentowanych w tym zestawieniu i jako jedyna nie wykorzystuje rozkładu kolorów obrazu wejściowego, więc jest od niego niezależna. Raz wygenerowana paleta barw może być wykorzystywana dla różnych danych wejściowych. Proces generowania palety polega na podziale sześcianu RGB na jednakowe prostopadłościany reprezentujące wybrane kolory. Każda z trzech krawędzi sześcianu RGB może być traktowana niezależnie i dzielona na dowolną liczbę jednakowych odcinków. Dla utworzonych w ten sposób prostopadłościanów wyznacza się reprezentujące je kolory. Przy często spotykanej w technice palety 256 kolorów stosuje się podział $8 \times 8 \times 4$, zgodnie z którym składowe R i G dzielimy na 8 części, a B na 4. W rezultacie, przy założeniu, że operujemy w 24-bitowej przestrzeni kolorów, otrzymujemy 256 prostopadłościanów o wymiarach $32 \times 32 \times 64$. W środku każdego z nich zlokalizowany jest punkt o współrzędnych odpowiadającym składowym reprezentującego go koloru. Przy podziale $8 \times 8 \times 4$ otrzymamy paletę 512 kolorów. Porównanie omawianych palet zobrazowano na rys. 4.

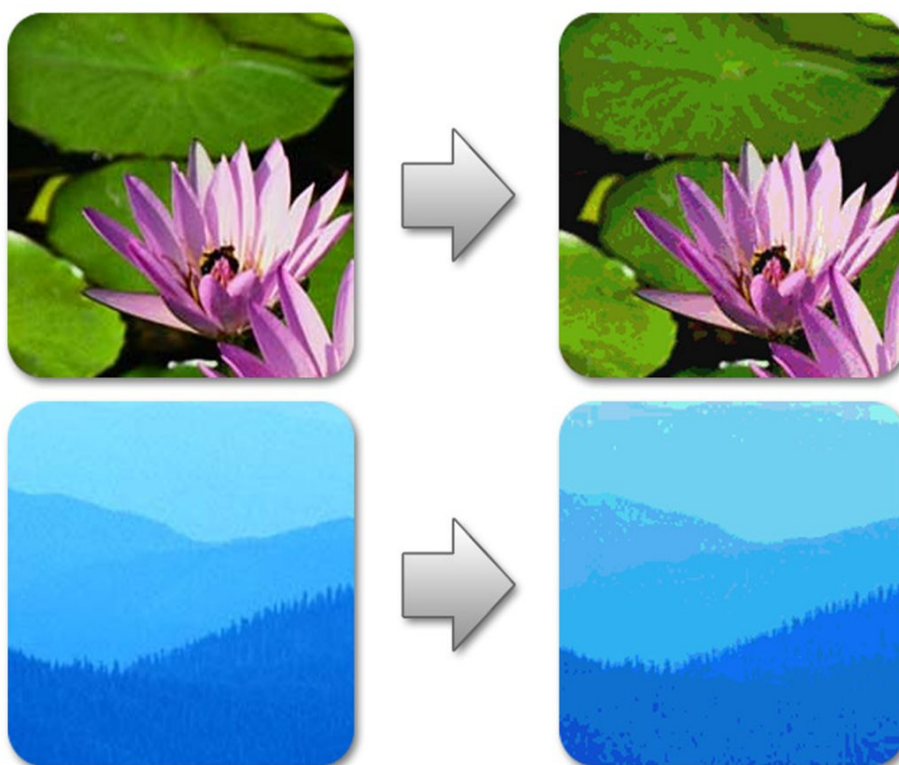


Rys. 4. Porównanie palet jednorodnych z 256 i 512 kolorami

Wadą tego rozwiązania jest brak analizy kolorów danego obrazu, przez co w wyznaczonej palety często występują kolory, które nie będą wykorzystane przy kwantyzacji. Wpływa to negatywnie na rezultat odwzorowania barw.

Rys. 5 prezentuje wynik procesu kwantyzacji z wykorzystaniem jednorodnej palety 512 kolorów. Nawet bez prezentowania obrazów wejściowych można dostrzec charakterystyczne niedostatki tej metody (regiony wypełnione jednolitym kolorem, co jest szczególnie widoczne w przypadku obszarów o łagodnych przejściach jednej składowej w obrazie wejściowym).

Zalety omawianej techniki to: prosta i szybka w działaniu implementacja oraz dobre rezultaty dla obrazów o zróżnicowanych kolorach. Niestety rzadko się zdarza, aby wszystkie kolory z palety zostały wykorzystane, co powoduje wyraźne (percepcyjne) obniżenie jakości obrazu, szczególnie w przypadku obrazów o mało zróżnicowanej liczbie kolorów.

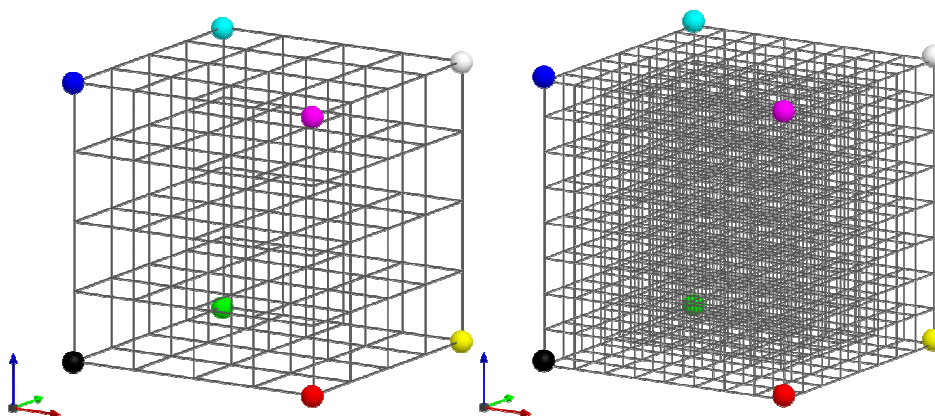


Rys. 5. Wyniki działania procesu kwantyzacji metodą podziału jednorodnego, dla przypadku 512 kolorów¹

¹ Artykuł w wersji kolorowej, wydrukowany do formatu PDF, dostępny jest na: www.ita.wat.edu.pl

4.2. Algorytm popularności (ang. *popularity algorithm*)

Technika ta wykorzystuje informację o rozkładzie kolorów w obrazie, co pozwala na lepsze, w porównaniu do metody podziału jednorodnego, oddanie barw na bitmapie wynikowej. Idea sprowadza się do wyboru palety najpopularniejszych kolorów w obrazie wejściowym. Reprezentantów można wyszukać w maksimach lokalnych histogramu RGB wygenerowanego na podstawie obrazu. Niestety podejście to ma poważną wadę, wynikającą z dużego skupienia maksymalnych wartości w przestrzeni barw (bardzo częsty przypadek w obrazach rzeczywistych). Wynikiem byłaby paleta posiadająca bardzo zbliżone kolory, mało efektywnie oddająca przestrzeń barw z obrazu wejściowego. Z tego powodu, w opisanym przez Heckberta² metodzie generowania palety, wykorzystano podział przestrzeni RGB na N^3 prostopadłościanów (rys. 6), jednak w liczbie w zdecydowanie większej, niż w przypadku metody jednorodnej (przykładowe wartości N podziału to: 16, 32 lub 64).

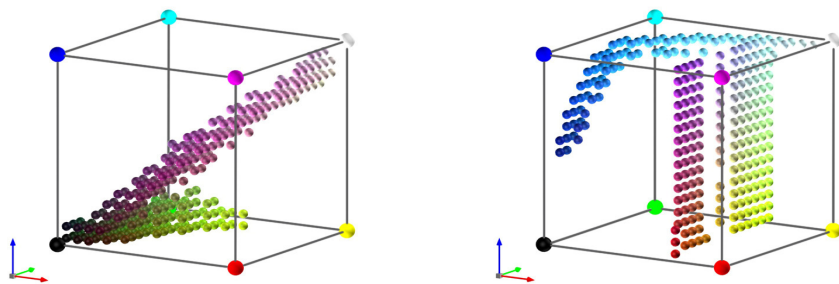


Rys. 6. Przestrzeń RGB podzielona na 64 i 512 sześcianów, odpowiednio dla $N = 4$ i $N = 8$

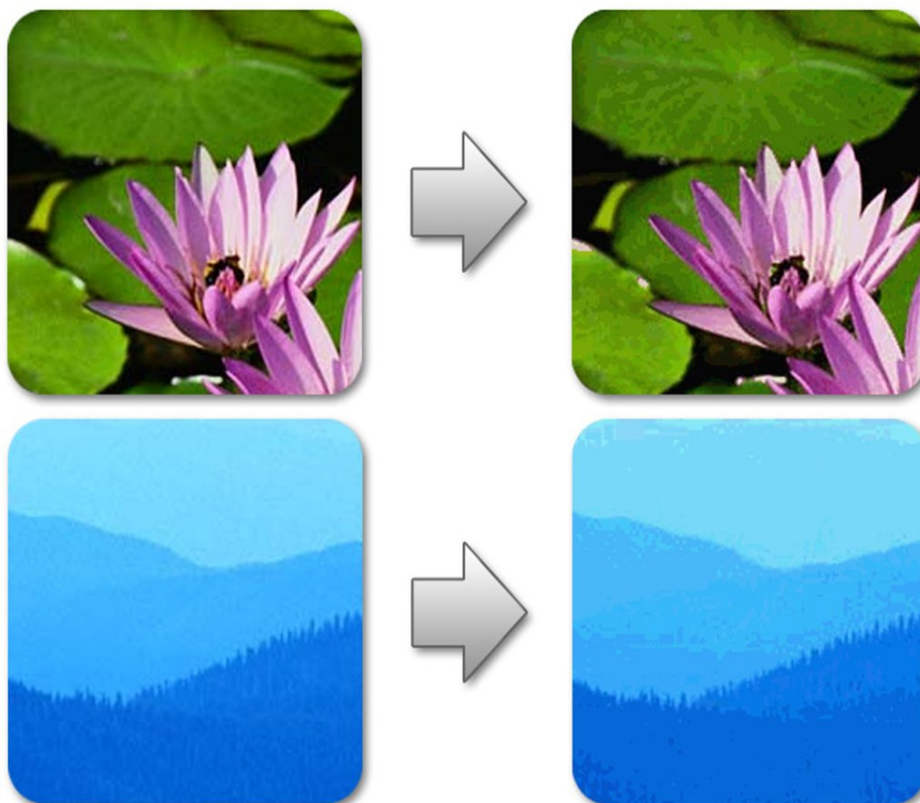
Dla każdego z utworzonych prostopadłościanów³ obliczana jest wartość określająca liczbę przynależnych pikseli obrazu wejściowego. Paletę kolorów określa się na drodze wyboru prostopadłościanów o największej liczbie głosów (rys. 7).

² Heckbert, Paul S.: *Color image quantization for frame buffer display*, Thesis (B.S.) - MIT, Dept. of Mathematics, 1980.

³ W literaturze stosuje się czasem porównanie prostopadłościanów do pojemników bądź wiader (ang. *buckets*), do których wpadają piksele z obrazu wejściowego.



Rys. 7. Reprezentacja dwóch palet 256 kolorów, wygenerowanych przy $N = 16$



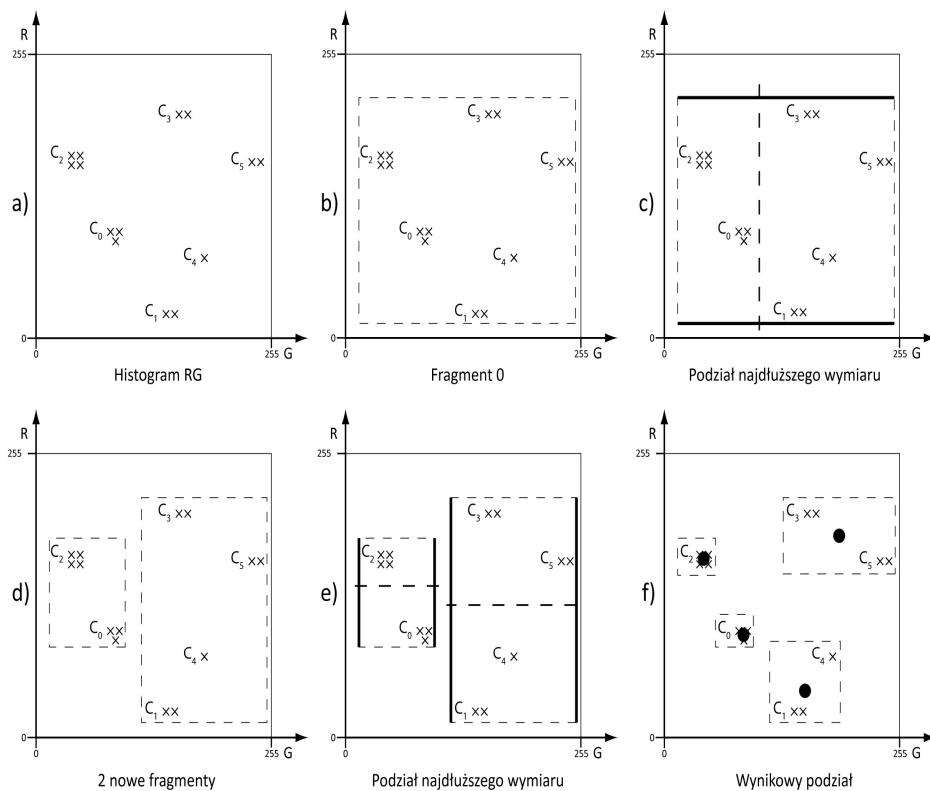
Rys. 8. Wyniki kwantyzacji dla 256 kolorów, przy $N = 16$

Jak można zauważyć, obserwując szczegóły i porównując je do wyników działania metody podziału jednorodnego, zmniejszyły się jednolite regiony. Niestety, szczególnie w przypadku obrazu drugiego, rezultat jest nadal mierny. Jakość obrazu wynikowego można próbować poprawić, zwiększając rozdzielczość podziału przestrzeni RGB na mniejsze sześciiany, przykładowo o wymiarach $4 \times 4 \times 4$ ($N = 64$). Niestety takie postępowanie może prowadzić do bardzo niepożądanego efektu zanikania mało reprezentatywnych barw w obrazie, co raczej jest niedopuszczalne.

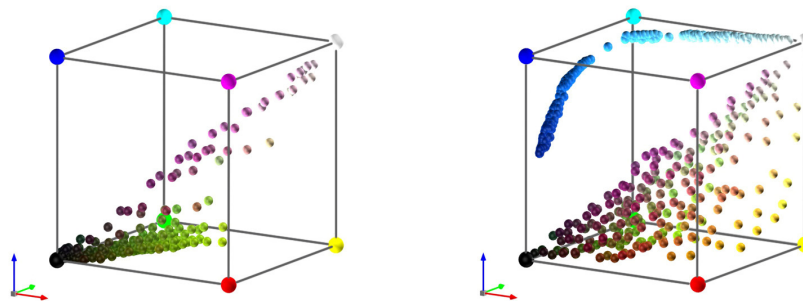
Podstawową zaletą algorytmu jest jego prosta implementacja i niewielka złożoność. Trudności pojawiają się jednak przy ocenie odpowiedniej wartości N . Występuje też zauważalny efekt pogorszenia jakości obrazów oraz może wystąpić zanikanie barw, których reprezentacja w obrazie nie jest liczna.

4.3. Metoda Median-Cut

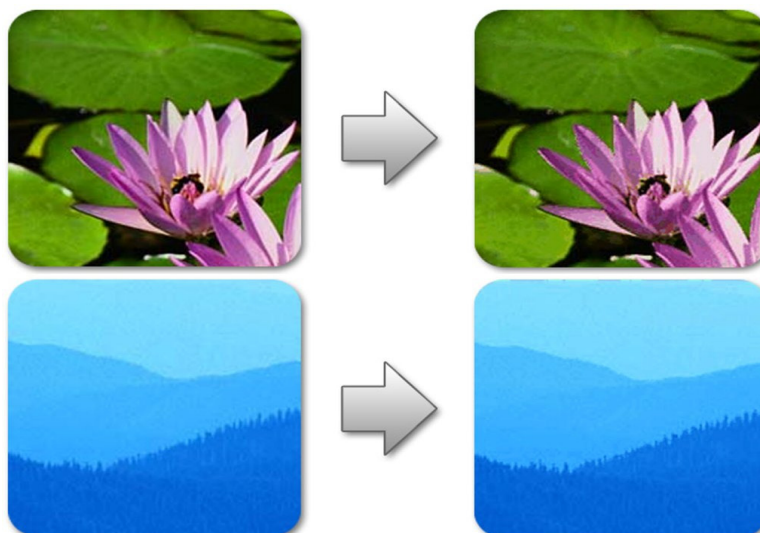
Opracowana przez Heckberta [7] technika bazuje na specyficznej metodzie podziału przestrzeni RGB na coraz mniejsze fragmenty, tak aby każdy z nich zawierał mniej więcej jednakową liczbę pikseli obrazu wejściowego. Pierwszym etapem procesu jest wyznaczenie obszaru w RGB obejmującego wszystkie piksele bitmapy. Wyznaczony obszar jest następnie dzielony na dwie części, poprzez podział najdłuższego z wymiarów tak, aby powstałe w ten sposób obszary zawierały tyle samo pikseli. Do następnego podziału wybierany jest ten, spośród wszystkich dotychczas powstałych obszarów, który obejmuje największą liczbę pikseli. Procedura jest powtarzana do uzyskania pożądanej liczby fragmentów (rys. 9). Na koniec dla każdego z fragmentów obliczany jest uśredniony kolor reprezentujący wszystkie zawarte w nim piksele. Należy zaznaczyć, że przedstawiony proces może zakończyć się przed uzyskaniem zadanej liczby fragmentów ze względu na brak możliwości dalszego podziału.



Rys. 9. Zobrazowanie działania metody Median-Cut dla czteroelementowej palety kolorów



Rys. 10. Reprezentacja dwóch palet 256 kolorów, wygenerowanych techniką Median-Cut



Rys. 11. Wyniki kwantyzacji przeprowadzonej metodą Median-Cut, dla 256 kolorów

Metoda umożliwia wygenerowanie palety kolorów dobrze oddającej barwy z obrazu wejściowego, co można zaobserwować na rys. 11. Przy porównaniu wynikowych bitmap z ich pierwowzorami trudno doszukać się różnic. Technika ta, spośród opisanych tutaj, daje najlepsze rezultaty przy percepcyjnym porównywaniu obrazów, niemal bez względu na ich charakter.

Podstawową zaletą metody jest bardzo dobry dobór kolorów palety, co przekłada się na dobre rezultaty dla dużej klasy obrazów. Metoda nie jest jednak łatwa w implementacji i dość czasochłonna w działaniu.

4.4. Metoda k-means [6]

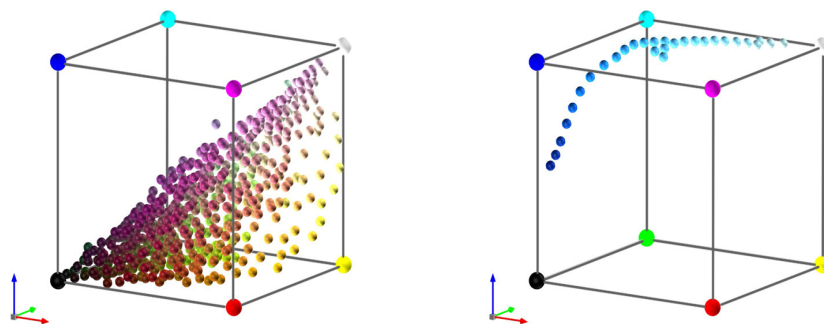
Algorytm centroidów jest jedną z najwcześniejszych [15] i najprostszych technik grupowania. Zasada działania opiera się na ustaleniu k centroidów, wstępnie reprezentujących skupiska, a następnie przyporządkowywaniu do nich najbliższych im punktów z danych wejściowych (kolorów w przypadku kwantyzacji). W każdej nowopowstałej grupie wyznaczany jest, na podstawie środka ciężkości, nowy centroid. Dla tak określonych środków nowych grup ponownie przyporządkowywane są punkty. Proces jest powtarzany do chwili, w której centroidy nie zmieniają już swojej pozycji w kolejnej iteracji.

W kolejnych krokach algorytmu minimalizowana jest wartość funkcji oceny jakości wyznaczonych grup (przykładowo reprezentowana przez funkcję kwadratu błędu):

$$J = \sum_{j=1}^k \sum_{i=1}^n |x_i^{(j)} - c_j|^2, \quad (8)$$

gdzie $|x_i^{(j)} - c_j|^2$ jest miarą odległości pomiędzy punktami ze zbioru danych $x_i^{(j)}$ a centroidem grupy c_j (j – numer grupy, n – liczba punktów budujących daną grupę).

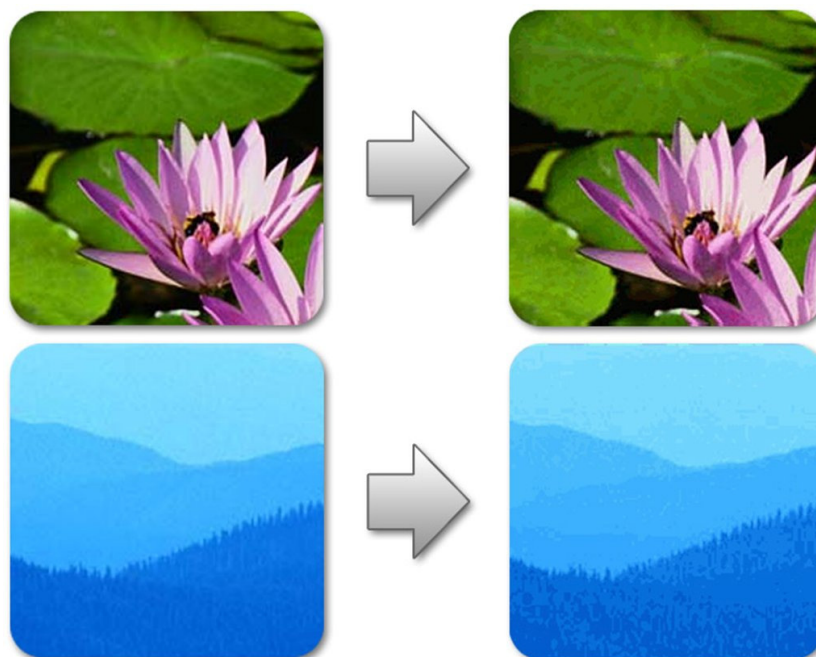
Nie ma jednoznacznego przepisu na początkowe ustalenie lokalizacji k środków. Często stosowaną metodą jest ich losowy wybór. Jednak jakość procesu grupowania (określona wartością J) jest ściśle uzależniona od początkowego rozlokowania środków grup, przez co w wyniku ich przypadkowego rozmieszczenia daje różne rezultaty.



Rys. 12. Palety kolorów wygenerowanych techniką k-means, przy ustalonych progach skupienia: 553 kolory dla obrazu pierwszego oraz 30 kolorów dla drugiego

Alternatywnym sposobem wykorzystania powyższej techniki, w przypadku, gdy nie jesteśmy w stanie określić liczby k , jest wykorzystanie współczynnika skupienia określonego jako:

$$s(j) = \frac{1}{n} \sum_{i=1}^n |x_i^{(n)} - c_j|^2 \quad (9)$$



Rys. 13. Przykładowy wyniki kwantyzacji kolorów metodą k-means

Pomysł polega na cyklicznym podziale danych na coraz mniejsze dwie grupy przy założeniu, że środkami nowych skupisk zostają dwa skrajne punkty oraz każdorazowo powtarzana jest metoda k-means dla ustalenia optymalnego położenia centroidów. Proces trwa do chwili, kiedy współczynniki skupienia wszystkich grup osiągną wartości poniżej ustalonego a priori progu.

Dla pierwszego obrazu, w przykładzie prezentowanym na rys. 13, metoda k-means dała aż 553 kolory palety, co zaowocowało dobrą jakościowo bitmapą wynikową. W drugim przypadku ta sama przyjęta wartość progu skupienia pozwoliła na wygenerowanie jedynie 30 kolorów, co niestety spowodowało powstanie dużych regionów wypełnionych jednolicie. Wadę tą można wyeliminować poprzez przyjęcie bardziej rygorystycznej wartości progu skupienia.

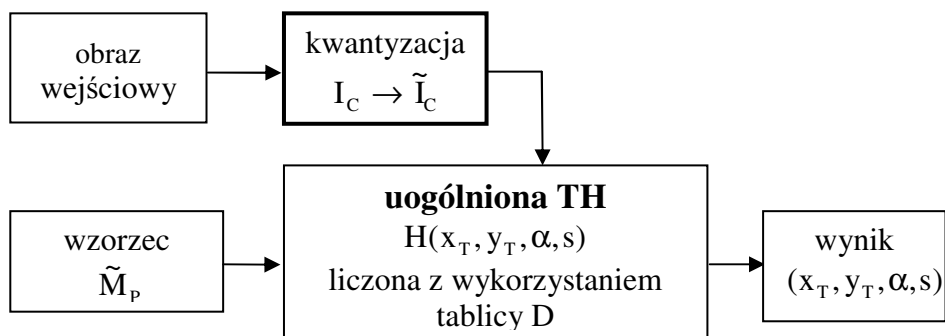
Dość dobry dobór barw palety wynikowej i wykorzystanie naturalnych skupień w przestrzeni RGB, to niewątpliwe zalety metody. Do wad natomiast należy zaliczyć dość trudną implementację i czas obliczeń wprost proporcjonalny do liczby elementów a odwrotnie do przyjętego progu skupienia.

5. Segmentacja obrazów kolorowych

Pierwsza opisana metoda kwantyzacji jest uniwersalna i umożliwia przygotowanie systemu do pracy zanim zostaną wprowadzone obrazy wejściowe i wzorce. Odpowiednio gęste podzielenie przestrzeni RGB powinno zagwarantować dobre wyniki w większości przypadków. Należy jednak pamiętać, że może się zdarzyć, iż system będzie analizował obrazy, dla których rozkład kolorów w sześcianie RGB tworzy nieliczne skupiska i wówczas jednorodny podział przestrzeni RGB będzie wysoce nieodpowiedni. W takim niekorzystnym przypadku zwiększanie liczby reprezentantów kolorów będzie poprawiać sytuację tylko nieznacznie, a ceną będzie szybki rozrost tablicy odległości. Z przeprowadzonych badań wynika, że podział sześcianu RGB na $8^3 = 512$ reprezentantów jest wystarczający, a tablica D zawiera wówczas $512^2 = 262144$ komórek. Można też pokusić się o gęstszy podział na $16^3 = 4096$ części, co spowoduje wzrost liczebności elementów tablicy D do $4096^2 = 16777216$. Przy założeniu, że każda komórka tablicy D jest reprezentowana w pamięci komputera na dwóch bajtach (Integer bez znaku 0-65536) otrzymujemy odpowiednio 512KB i 32MB zapotrzebowania na pamięć, odpowiednio dla każdego z przypadków, co jest współcześnie akceptowalne.

W przypadku innej metody kwantyzacji system widzenia komputerowego będzie musiał wprawdzie dokonać analizy potencjalnych obrazów i wzorców, dopasować odpowiednią metodę kwantyzacji i wyliczyć tablicę odległości. Wadą takiego podejścia do problemu jest wrażliwość na zmianę charakterystyki obrazów wejściowych, co tylko częściowo można skompensować na etapie tworzenia tablicy odległości poprzez nacisk na charakter (kolorystykę) wzorców. Wzorce są tutaj decydujące, gdyż to one mają być dopasowane do obiektów na obrazie wejściowym, ale może zaistnieć sytuacja, gdy kwantyzacja „nowego” obrazu wejściowego doprowadzi do niepożądanego wyeksponowania niewłaściwych obiektów i obniżenia niezawodności systemu.

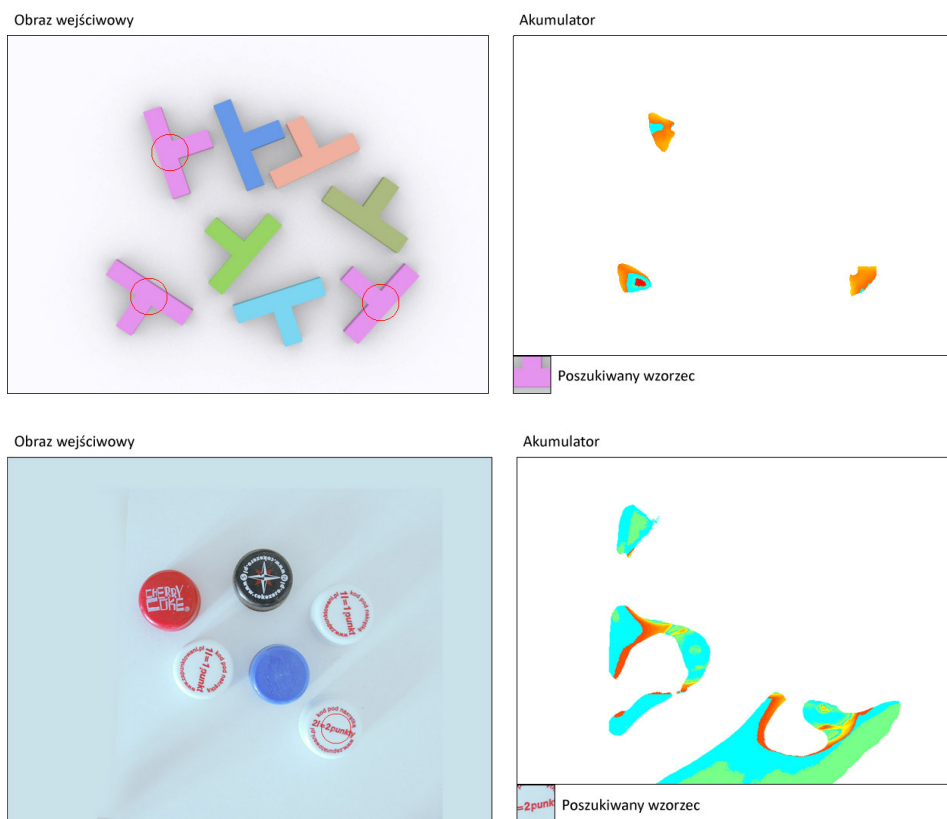
Bez względu na wybór ścieżki postępowania (na etapie kwantyzacji) otrzymamy w wyniku tablicę odległości kolorów D . Zastosowanie tej tablicy przy wyznaczaniu wartości funkcji $h(x_i, y_i, x_T, y_T, \alpha, s)$ przedstawia wzór (7), który w przypadku obrazów kolorowych zastępuje wzór (2). Można zatem powiedzieć, że wzory (1), (7) i (3) definiują transformatę Hougha dla przypadku segmentacji obrazów kolorowych wzorcami nieregularnymi. Proces identyfikacji obiektów na obrazach kolorowych został przedstawiony na rys. 14.



Rys. 14. Proces dopasowywania wzorca do obiektu na obrazach kolorowych

Wyznaczenie tablicy D jest realizowane przed uruchomieniem systemu, a w rezultacie otrzymujemy mapę kolorów, która będzie wykorzystywana przy kwantyzacji kolejnych obrazów, a wcześniej wzorców. Kwantyzacja obrazu wejściowego I_C będzie skutkowałą zastąpieniem pikseli obrazu przez indeksy \tilde{I}_C tablicy D . Wyznaczenie transformaty Hougha sprowadzi się do akumulowania odczytów z tablicy $D[\tilde{I}_C(x_i'', y_i''), \tilde{M}_p(x_i, y_i)]$.

Na rys. 15 przedstawiono przykładowe rezultaty otrzymane przy zastosowaniu uogólnionej transformaty Hougha. Przetwarzane były obrazy kolorowe pozyskane kamerą i aparatem cyfrowym, które poddano procesowi kwantyzacji metodą Median-Cut 256. Wybór tej metody kwantyzacji podyktowany był percepcyjnie najlepszym odwzorowaniem kolorów. Rozdzielczość obrazów wejściowych wynosiła ok. 670x500 pikseli, a czas obliczeń był poniżej 5s na komputerze z procesorem Pentium 4 2,8GHz. Na skrócenie czasu obliczeń duży wpływ miała analiza histogramów. Analiza histogramów nie jest związana z procesem kwantyzacji. Zadaniem tej analizy jest wykluczenie z obliczeń tych elementów (lokalizacji), dla których występują zbyt duże różnice w histogramach danego (przez wartości: x_T, y_T, α, s) obszaru obrazu wejściowego i wzorca. Technika ta została przedstawiona dla obrazów w poziomach szarości w [18] oraz [21], a jej zastosowanie przynosi redukcję wyznaczanych elementów akumulatora nawet o 90% (rys. 15).



Rys. 15. Przykładowe wyniki zastosowania uogólnionej transformaty Hougha na obrazach kolorowych poddanych procesowi kwantyzacji; rozpoznane obiekty zaznaczono okręgami na obrazie wejściowym

6. Podsumowanie

Na podstawie wielu przeprowadzonych doświadczeń można wyciągnąć wniosek, że zastosowanie kwantyzacji kolorów i uogólnionej transformaty Hougha przynosi bardzo dobre rezultaty w procesie segmentacji obrazów kolorowych wzorcami nieregularnymi. Problemy z rozpoznawaniem obiektów pojawiają się przy zmianie warunków oświetlenia sceny; natomiast przy zachowaniu stałych warunków oświetlenia zastosowana technika jest niemal niezawodna, a otrzymany rezultat niewiele zależy od wyboru metody kwantyzacji obrazów kolorowych. Obserwowane skrócenie czasu obliczeń, w stosunku do metod bez kwantyzacji kolorów, jest ponad dziesięciokrotne.

Czas wykonywania obliczeń w stosunku do obrazów w poziomach szarości nie ulega zauważalnej zmianie, gdyż mimo poddawania obrazu wejściowego procesowi kwantyzacji kolorów pojawia się zysk na etapie wyliczania akumulatora. Podstawowym i chyba jedynym problemem jest niesatysfakcjonujący czas obliczeń wynoszący kilka sekund, który będzie nieakceptowalny w wielu systemach widzenia komputerowego.

Literatura

- [1] BALLARD D. H., *Generalizing the Hough Transform to Detect Arbitrary Shapes*, Readings in Computer Vision: Issues, Problems, Principles, and Paradigms, Los Altos, CA, 1987, pp. 714-725.
- [2] DAVIES E. R., *Machine Vision: Theory, Algorithms, Practicalities*, Academic Press Ltd, 24/28 Oval Road, London NW1 7DX, United Kingdom, 1990.
- [3] DEANS S. R., *Hough transform from the Radon transform*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 3, No. 2, 1981, pp. 185-188.
- [4] FU K. S., GONZALEZ R. C., LEE C. S. G., *ROBOTICS: Control, Sensing, Vision, and Intelligence*, McGraw-Hill, New York, 1987.
- [5] HAN J. H., KOCZY L. T., POSTON T., *Fuzzy Hough transform*, Pattern Recognition Letters, Vol. 15, No. 7, 1994, pp. 649-658.
- [6] HARTIGAN J. A., WONG M. A., *A K-Means Clustering Algorithm*, Applied Statistics, Vol. 28/1979, No. 1, pp. 100-108.
- [7] HECKBERT P., *Color image quantization for frame buffer display*, Proceedings of the 9th annual conference on Computer graphics and interactive techniques, Boston, 1982, pp. 297-307.
- [8] HOUGH P. V. C., *Method and means for recognizing complex patterns*, U.S. Patent 3,069,654, Dec. 18, 1962.
- [9] ILLINGWORTH J., KITTLER J., *A survey of the Hough Transform*, Computer Vision, Graphics and Image Processing 44, 1988, pp. 87-116.
- [10] JAIN A. K., *Fundamentals of Digital Image Processing*, Prentice-Hall, New Jersey, 1989.
- [11] KIRYATI N., ELDAR Y., BRUCKSTEIN A. M., *A probabilistic Hough transform*, Pattern Recognition, Vol. 24, No. 4, 1991, pp. 303-316.
- [12] LEAVERS V. F., *Shape Detection in Computer Vision Using the Hough Transform*, Springer, London, 1992.

- [13] LI H., LAVIN M. A., LEMASTER R. J., *Fast Hough transform*, Proceedings of the Third Workshop on Computer Vision: Representation and Control (Bellaire, MI, October 13-16, 1985), IEEE Publ. 85CH2248-3, pp. 75-83.
- [14] LI H., LAVIN M. A., LEMASTER R. J., *Fast Hough transform: a hierarchical approach*, Computer Vision, Graphics, and Image Processing, Vol. 36, 1986, pp. 139-161.
- [15] MACQUEEN, J. B., *Some Methods for classification and Analysis of Multivariate Observations*, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, University of California Press 1967, pp. 281-297.
- [16] RADON J., *Über die Bestimmung von Funktionen durch ihre Integralwerte langs gewisser Mannigfaltigkeiten*, Berichte Sachsische Akademie der Wissenschaften Leipzig, Math Phys Kl., 69, 1917, pp 262-267.
- [17] XU L., OJA E., KULTANEN P., *A new curve detection method: Randomized Hough Transform (RHT)*, Pattern Recognition Letters, Vol. 11, No. 5, 1990, pp. 331-338.
- [18] ŻORSKI W., FOXON B., BLACKLEDGE J., TURNER M., *Irregular Pattern Recognition Using the Hough Transform*, Machine Graphics & Vision, 9, 2000, pp. 609-632.
- [19] ŻORSKI W., ŻAK A., TURNER M., *Hardware Implementation of the Hough Technique for Irregular Pattern Recognition*, Proceedings of the 8th IEEE International Conference MMAR 2002, 2-5 September 2002, Vol.1, pp. 561-566.
- [20] ŻORSKI W., *Application of the Hough Technique for Irregular Pattern Recognition to a Robot Monitoring System*, Proceedings of the 11th IEEE International Conference MMAR 2005, pp. 725-730.
- [21] ŻORSKI W., *The Hough Transform Application Including Its Hardware Implementation*, Advanced Concepts for Intelligent Vision Systems: Proceedings of the 7th International Conference, Lecture Notes in Computer Science, Springer-Verlag Vol. 3708/2005, pp. 460-467.
<http://www.springerlink.com/content/50yk3q0fw71x1qlld>.
- [22] ŻORSKI W., *Fast Hough Transform Based on 3D Image Space Division*, Advanced Concepts for Intelligent Vision Systems: Proceedings of the 8th International Conference, Lecture Notes in Computer Science, Springer-Verlag Vol. 4179/2006, pp. 1071-1079. <http://www.springerlink.com/content/6216256332x1p166>.
- [23] ŻORSKI W., *Unknown scale objects recognition*, Biuletyn WAT (652), 4/2008, pp. 197-207.

Colour image segmentation with irregular patterns

ABSTRACT: This paper considers the problem of object recognition in colour images, excepting the feature extraction process. On the pre-processing stage, the only preceding segmentation, a colour quantisation technique is applied to avoid the use of a whole RGB colour space. The generalized Hough transform is used as a tool of segmentation to identify irregular objects. The colour quantisation process is essential for the recognition reliability and received special attention. In a given computer vision system, the choice of a quantisation technique should be dictated by the nature of analyzed input images.

KEYWORDS: computer vision, image recognition, segmentation, Hough transform, clustering, colour quantisation

Praca wpłynęła do redakcji: 12.10.2009