

Zbigniew Bubliński*, Zbigniew Mikrut*

Efektywność wybranych metod etykietowania obiektów**

1. Wprowadzenie

Etykietowanie (często też nazywane indeksacją) obiektów jest jedną z częściej wykonywanych operacji w analizie obrazów. Jej zadaniem jest nadanie wszystkim obiektom unikalnych etykiet: wszystkie punkty pierwszego obiektu uzyskują identyczną wartość – np. 1, drugiego np. 2, itd. Taki sposób postępowania pozwala na proste i szybkie wydzielenie dowolnego obiektu do szczegółowych badań, zapewniając niejako przy okazji precyzyjną informację o liczbie obiektów występujących na obrazie. W praktyce, znakomita większość systemów analizy obrazów wykorzystuje etykietowanie jako jeden z etapów przetwarzania – między innymi systemy automatycznej analizy obrazów ruchu drogowego [1–3]. W takich właśnie systemach szczególnego znaczenia nabiera czas analizy – optymalnym rozwiązaniem byłaby sytuacja, gdy kompletna analiza obrazu mogła być dokonywana w czasie rzeczywistym. Oznacza to konieczność wybrania takiej metody etykietowania, która zapewni minimalny czas przetwarzania, jeśli nie we wszystkich, to przynajmniej w większości typowych zadań.

2. Metody etykietowania obiektów

Do najczęściej spotykanych w literaturze [7–9] i wykorzystywanych w praktycznych zastosowaniach należą dwie metody. Pierwsza z nich to indeksacja oparta na rekurencyjnym przeglądaniu punktów obiektu popularnie nazywana metodą „pożaru prerii” (w języku angielskim ten algorytm nosi nazwę *floodfill*). Jej idea sprowadza się do znalezienia dowolnego punktu jeszcze niezaetykietowanego obiektu i „podpalenia” go. Punkt ten staje się źródłem „pożaru” – punkty z nim sąsiadujące „płoną” przenosząc „ogień” z kolei na swoich jeszcze nie „zapalonych” sąsiadów, itd. – stąd nazwa metody. Jeśli „wypalone” punkty będą oznaczane identycznymi wartościami, a zmiana wartości będzie dokonywana przy okazji „podpalenia” kolejnego obiektu, to uzyskamy w pełni funkcjonalną metodę etykietowania. Klasyczna jej realizacja oparta jest na rekurencyjnym przeglądaniu punktów, które mogą być kandydatami do „zapalenia”. Ze względu na bardzo duże wymagania pamięciowe (za-

* Katedra Automatyki, Akademia Górniczo-Hutnicza w Krakowie

** Praca zrealizowana w ramach badań własnych nr 10.10.120.39

jętość stosu), wersja taka nie nadaje się do przetwarzania obrazów o rozmiarach spotykanych w typowych systemach wizyjnych. Możliwe jest zastąpienie wersji rekurencyjnej przez wersję iteracyjną – aby tego dokonać, konieczne jest zapamiętywanie punktów będących potencjalnymi kandydatami do „spalenia” w pewnej jawnej strukturze danych (zamiast – jak w rekurencji – niejawnie na stosie procesora). Wybranie jako takiej struktury stosu nie poprawia sytuacji – może się okazać, że sam stos zajmie kilka razy więcej miejsca niż sam etykietowany obraz. Jedynym sensownym sposobem rozwiązania istniejących problemów jest wybranie kolejki jako formy organizacji danych – punkty do „zapalenia” pobierane są z początku kolejki, a na jej koniec dopisywane są punkty będące sąsiadami już „zapalonych” jako potencjalni kandydaci do stania się źródłem „ognia” w kolejnym kroku. Taka implementacja pozwala dokonać etykietowania obrazu o w zasadzie dowolnych rozmiarach – realnym problemem może być tylko mała efektywność metody wynikająca z tego, że niektóre punkty obiektu mogą być wielokrotnie przeglądane (co wynika z tego, że dany punkt jest sąsiadem dziewięciu innych punktów). Jako że efektywność takiej nie można oszacować teoretycznie, gdyż zależy ona od postaci samego obrazu, konieczne jest wyznaczenie jej na drodze eksperymentalnej. Metoda ta rodzi też istotne problemy przy próbach stworzenia jej sprzętowej implementacji.

Zupełnie innym podejściem do etykietowania jest druga z metod – tzw. liniowe przeglądanie obrazu połączone ze sklejaniem uprzednio wstępnie zaetykietowanych obiektów. W metodzie tej wykorzystywane są proste operacje sprowadzające się do przeglądnięcia linii po linii, kolumnie po kolumnie wszystkich punktów obrazu i sprawdzenie dla każdego punktu kilku prostych warunków logicznych, których rezultat warunkuje nadanie danemu punktowi określonej etykiety. Ze względu na to, że w pewnych sytuacjach może okazać się, iż obiekty dotychczas uważane za rozłączne są faktycznie ze sobą połączone (np. obiekt w kształcie litery U), konieczny jest dodatkowy etap przetwarzania. Jego zadaniem jest dokonanie końcowej reindeksacji obrazu tak, aby cały obiekt uzyskał tę samą etykietę. Na etapie tym wykorzystywane są informacje o ewentualnych połączeniach obiektów o różnych etykietach zapamiętane w tzw. tablicy sklejeń. Ze względu na małe skomplikowanie algorytmu i proste zależności czasowe w trakcie przetwarzania, metoda ta szczególnie dobrze nadaje się do realizacji sprzętowej – opis przykładowej implementacji można znaleźć w [5, 6].

3. Implementacja algorytmów i metodologia porównania

W celu porównania obu metod pod kątem ich efektywności czasowej, zaimplementowano je w postaci filtru działającego w środowisku VirtualDub [4]. Pozwoliło to na elastyczne testowanie algorytmów na różnych sekwencjach obrazów – wygenerowanych i zarejestrowanych w rzeczywistych warunkach. W stworzonym filtrze zawarto możliwość wybrania metody etykietowania przed rozpoczęciem przetwarzania. Umieszczenie kodu obu metod w jednym filtrze pozwoliło na zapewnienie identycznych warunków w trakcie badań porównawczych – wszystkie fragmenty niezbędne do działania (konwersje, przygotowanie danych, itp.) i wspólne dla obu metod mają tę samą postać, a co za tym idzie, czas ich wykonania jest taki sam.

Czasy wykonania procedur etykietowania mierzono za pomocą tzw. High Performance Timer dostępnego w środowisku Windows. Timer ten wykorzystuje zmienne o wielkości 64 bitów, co zapewnia rozdzielczość pomiaru czasu na poziomie pojedynczych mikrosekund i trochę mniejszą dokładność. Nie jest to jednak istotnym problemem, gdyż czasy etykietowania w zależności od wybranej metody i zawartości samego obrazu mieszczą się w przedziale od kilku do kilkudziesięciu milisekund. Aby zwiększyć dokładność pomiaru czasu uzyskane wyniki skorygowano, uwzględniając czas trwania samego pomiaru czasu. Należy jednak pamiętać, że w większości środowisk wielozadaniowych (a do takich należy MS Windows) pomiar czasu dowolnej operacji nie może być zawsze precyzyjny ze względu na równoległe działanie innych procesów. Minimalizacja powstałych błędów jest jednak możliwa i prosta do wprowadzenia – wystarczy po prostu kilkanaście lub kilkadziesiąt razy powtórzyć daną operację i dokonać uśrednienia otrzymanych wyników.

Jedynym kryterium stosowanym przy porównaniu był czas trwania etykietowania obrazu uzyskane dla wybranej metody. Ponieważ jednak, jak już wcześniej wspomniano, czasy te zależą w znacznym stopniu od treści obrazu (jego zawartości), w celu uzyskania szerszej wiedzy o przydatności poszczególnych metod indeksacji do różnych klas obrazów, zrealizowano szereg eksperymentów na kilku znacznie różniących się od siebie sekwencjach obrazów.

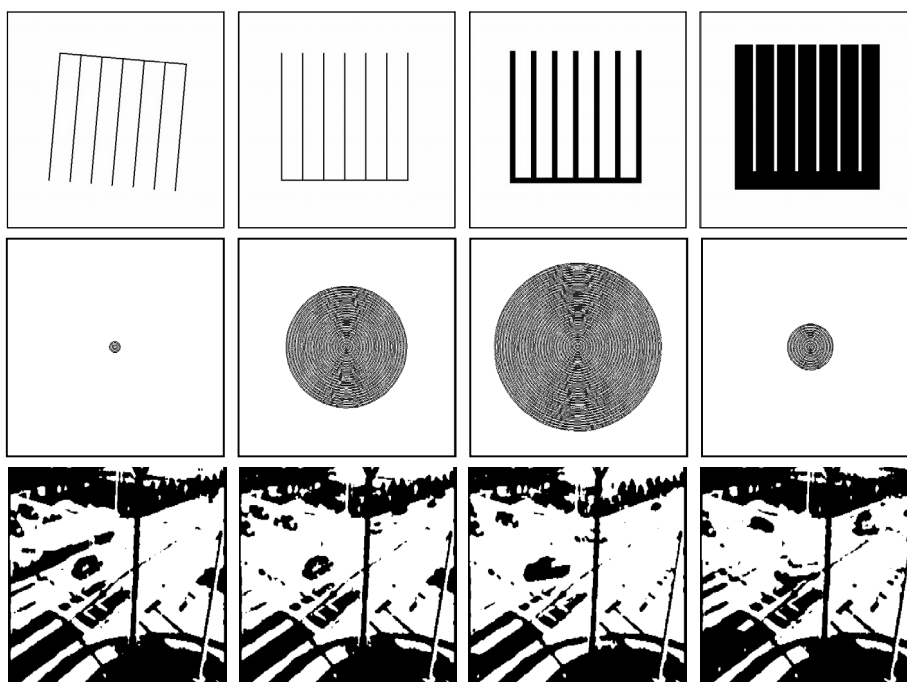
4. Wyniki testów

Dokonując interpretacji pokazanych dalej wyników, należy pamiętać o dwóch kwestiach. Pierwsza to problem z dokładnym pomiarem czasu wykonania fragmentu kodu w systemie wielozadaniowym. Wykonywany kod może zostać w dowolnym momencie przerwany przez inne, bardziej istotne dla systemu operacyjnego, zadania. W takiej sytuacji zmierzony czas będzie większy niż faktyczny. W praktyce należy więc liczyć się z pojawianiem się w pomiarach co jakiś czas znaczących odchyłek w górę. Druga kwestia to silne zróżnicowanie czasu trwania etykietowania w zależności od postaci obrazu. Samo zróżnicowanie wyniku wprost z algorytmu – problem polega na tym, iż trudno z góry przewidzieć, jaki obraz będzie „łatwy” do indeksacji, a jaki „trudny”. Oznacza to, że czasy etykietowania obrazów podobnych do siebie (np. jeden jest odwróconą do góry nogami wersją drugiego) mogą się znacząco różnić. Można to zauważyć, analizując wyniki uzyskane dla filmu „Grzebień”.

Zamieszczone dalej wyniki uzyskano dla trzech filmów – we wszystkich przypadkach rozmiary klatki obrazu wynosiły 512×512 pikseli. Tablica sklejeń w metodzie przeglądania liniowego miała rozmiar 4096, co oznacza możliwość zaetykietowania 4094 obiektów (wartość 0 jest zarezerwowana dla tła, a wartość 4095 dla obiektów, którym jeszcze nie nadano żadnej etykiety).

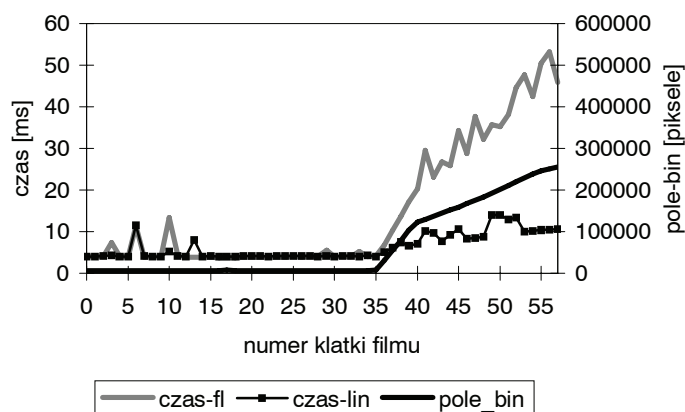
Do testów użyto trzech filmów, które umownie nazwano: „Grzebień”, „Okręgi” oraz „Ruch”, oraz – w celu dokładniejszego zbadania zależności od wielkości etykietowanych obiektów – ich zanegowanych wersji. Przykładowe klatki klatek z ww. filmów przedstawiono na rysunku 1. Film „Grzebień” to obrazy obiektu o kształcie grzebień – obiekt najpierw jest obracany wokół środka klatki, a następnie ulega dylatacji aż do połączenia się wszystkich zębów. Obiektem użytym w filmie „Okręgi” są koncentryczne okręgi – do poło-

wy filmu ich liczba i średnice ulegają zwiększeniu począwszy od pustego obrazu aż do jego całkowitego wypełnienia, zaś w drugiej połowie sytuacja jest odwrotna – okręgi maleją i znikają. Ze względu na pewne niedokładności w trakcie generowania okręgów niektóre z nich są ze sobą połączone, co powoduje, że liczba obiektów zmienia się w sposób nieregularny. Ostatni z filmów („Ruch”) to zbinaryzowany obraz ruchu drogowego zarejestrowany na skrzyżowaniu.



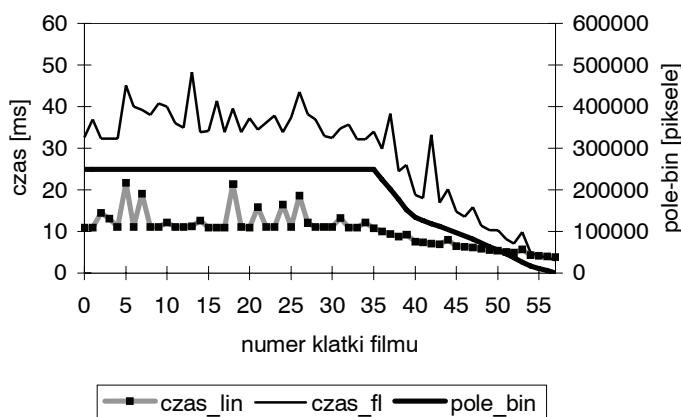
Rys. 1. Przykłady klatek z filmów użytych do testów

Wykresy czasów wykonania obu metod indeksacji dla kolejnych klatek filmu „Grzebień” przedstawiono na rysunku 2. Dodano też do niego przebieg zmian liczby pikseli składających się na etykietowane obiekty. Na wykresie wyraźnie widocznych jest kilka znacznych odchyłeń – jako że miejsca te występują nieregularnie i trudno je skorelować z zawartością poszczególnych klatek obrazu, dlatego najprawdopodobniej przyczyną tych zaburzeń jest system Windows, który w tym czasie wykonuje własne procesy o wyższym priorytecie niż samo zadanie indeksacji. W przypadku gdy obiekt poddany etykietowaniu jest relatywnie niewielki, trudno zauważyć jakiegokolwiek różnice w czasach wykonania pomiędzy testowanymi metodami. Rozbieżności w wynikach pojawiają się dopiero wtedy, gdy obiekt składa się ze znacznie większej liczby punktów – wtedy metoda przeglądania liniowego i sklepania wykazuje swoją przewagę.



Rys. 2. Film „Grzebień”: czasy etykietowania dla obu metod oraz rozmiar obiektu (krzywa *pole-bin*)

Jeszcze wyraźniej widać to na rysunku 3, na którym przedstawiono wyniki uzyskane na zanegowanej wersji filmu „Okregi”. Przez większą część filmu etykietowany obiekt zawiera wiele punktów, co od razu przekłada się na czasy indeksacji. Różnice zacierają się na końcowych klatkach filmu, kiedy obiekt jest już niewielki.



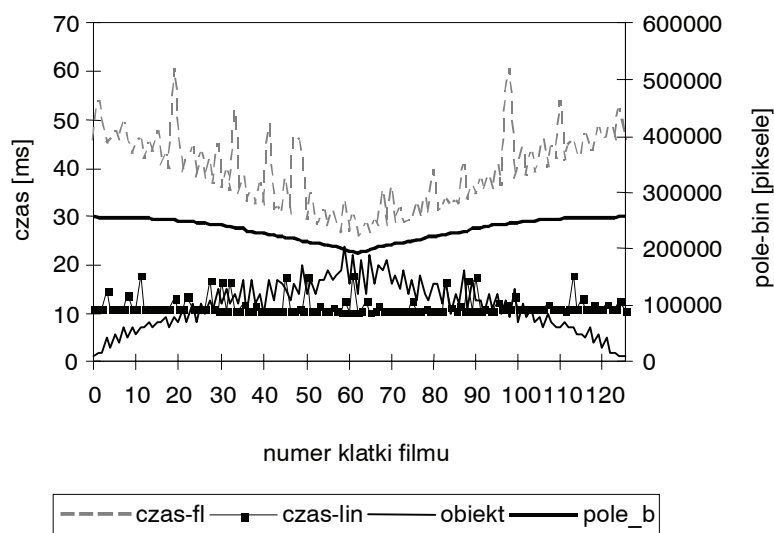
Rys. 3. Film „Negatyw grzebień”: czasy etykietowania dla obu metod oraz rozmiar obiektu

Zbiorcze wyniki takie, jak średni czas etykietowania, czasy minimalne i maksymalne, wyrażone zarówno w wartościach bezwzględnych, jak i względnych, uzyskane dla filmu „Grzebień” i jego zanegowanej wersji, przedstawiono w tabeli 1. Widać z niej wyraźnie, że metoda przeglądania liniowego i sklejanie zapewnia generalnie szybsze działanie, chociaż dla obu filmów minimalny czas indeksacji uzyskano w metodzie bazującej na rekurencji.

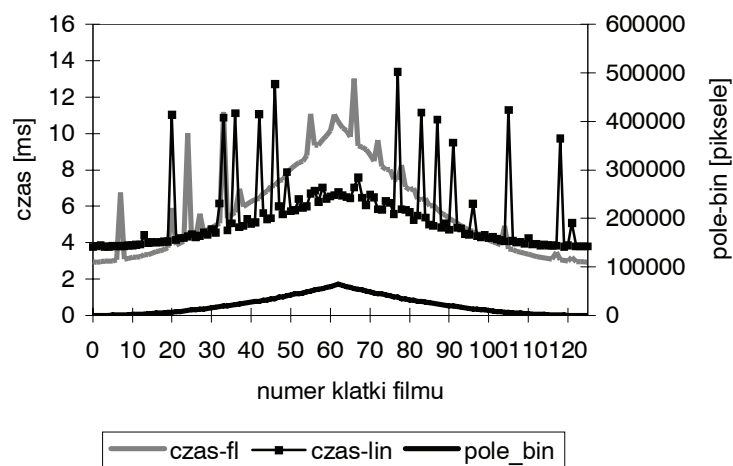
Tabela 1
Filmy „Grzebień” i „Negatyw grzebień”: czasy etykietowania

	„Grzebień”		„Negatyw grzebień”	
	<i>Floodfill</i>	<i>Linear</i>	<i>Floodfill</i>	<i>Linear</i>
Średnia [ms]	14,96	6,44	28,56	10,26
Maksimum [ms]	53,31	14,08	48,31	21,74
Minimum [ms]	3,82	4,01	2,96	3,81
Maks. [%]	256,49	118,72	69,15	111,80
Min [%]	74,47	37,73	89,65	62,89

Podobne rezultaty uzyskano dla filmu „Okregi” i jego zanegowanej wersji, chociaż tutaj charakter zmian wygląda zupełnie inaczej. Można to zaobserwować na rysunkach 4 oraz 5. Ponownie można zauważyć częste skokowe odchylenia w czasach przetwarzania wywołane działaniem systemu operacyjnego. Zmierzone i uśrednione czasy etykietowania dla obu filmów zawarto w tabeli 2.



Rys. 4. Film „Okregi”: czasy etykietowania dla obu metod oraz liczba i łączne rozmiary obiektów

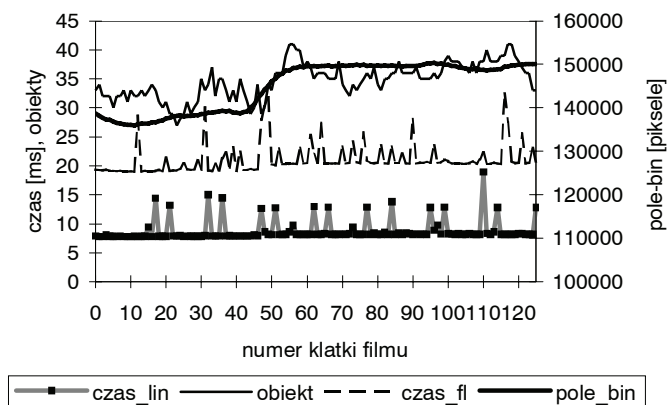


Rys. 5. Film „Negatyw okręgów”: czasy etykietowania dla obu metod oraz rozmiar obiektu

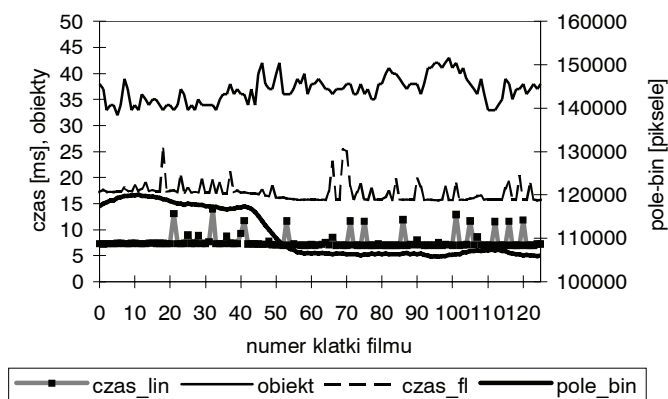
Tabela 2
Filmy „Okregi” i „Negatyw okręgów”: czasy etykietowania

	„Okregi”		„Negatyw okręgów”	
	<i>Floodfill</i>	<i>Linear</i>	<i>Floodfill</i>	<i>Linear</i>
Średnia [ms]	38,94	11,65	5,75	5,48
Maksimum [ms]	60,85	18,06	13,02	13,40
Minimum [ms]	26,15	10,34	2,93	3,77
Maks. [%]	56,28	55,03	126,40	144,29
Min [%]	32,84	11,26	49,05	31,27

Ostatnim z testowanych filmów był „Ruch” oraz jego negatywowa wersja. W tych filmach występuje znaczne zróżnicowanie liczby obiektów poddawanych etykietowaniu na poszczególnych klatkach. Aby wykryć ewentualne korelacje pomiędzy liczbą obiektów a czasami indeksacji, zdecydowano się na dodanie do wykresów jeszcze jednej krzywej pokazującej liczbę etykietowanych obiektów na danej klatce filmu. Wyniki dla filmu „Ruch” pokazano na rysunku 6, zaś rysunek 7 ilustruje rezultaty uzyskane dla filmu „Negatyw ruchu”. Warto zauważyć, że jakkolwiek wielkości obiektów na obu filmach znacznie się różnią, to liczby obiektów są w zasadzie prawie że identyczne. Porównując wykresy, łatwo dostrzec wyższość metody przeglądania liniowego – zapewnia ona średnio ponad dwukrotnie krótszy czas etykietowania niż druga z badanych metod. Również w skrajnych przypadkach (wartości minimalne i maksymalne) tendencja ta jest zachowana co wyraźnie widać w tabeli 3.



Rys. 6. Film „Ruch”: czasy etykietowania dla obu metod oraz liczba i łączne rozmiary obiektów



Rys. 7. Film „Negatyw ruchu”: czasy etykietowania dla obu metod oraz liczba i łączne rozmiary obiektów

Tabela 3
Filmy „Ruch” i „Negatyw ruchu”:
czasy etykietowania

	„Ruch”		„Negatyw ruchu”	
	<i>Floodfill</i>	<i>Linear</i>	<i>Floodfill</i>	<i>Linear</i>
Średnia [ms]	9,04	11,65	16,49	7,76
Maksimum [ms]	23,39	36,39	28,66	17,68
Minimum [ms]	7,76	18,71	14,33	6,63
Maks. [%]	158,70	67,80	73,81	127,78
Min [%]	14,18	13,74	13,08	14,58

5. Podsumowanie

Zasadniczym celem opisywanych badań było sprawdzenie, która z dwóch najbardziej popularnych metod etykietowania obiektów jest szybsza. W tym celu zaimplementowano oba algorytmy w postaci filtru działającego w środowisku VirtualDub. Zrealizowano sześć eksperymentów w oparciu o trzy sekwencje klatek z obrazami do indeksacji oraz o ich negatywowe wersje. Platformą testową był komputer PC z procesorem Athlon 2400 działający pod kontrolą systemu operacyjnego Windows XP.

Rezultaty wszystkich eksperymentów przedstawiono w postaci wykresów, zaś wyniki zbiorcze w tabelach. Wynika z nich niezbicie, że metodą zapewniającą krótszy czas etykietowania jest algorytm liniowego przeglądania i sklejanie. Charakteryzuje się on dużą liniowością – czasy przetwarzania są prawie stałe (pomijając odchylenia spowodowane głównie poprzez ingerencje ze strony systemu operacyjnego) i nie zależą zbytnio od zawartości obrazu.

Metoda bazująca na podejściu rekurencyjnym cechuje się znaczną zależnością od wielkości etykietowanych obiektów – indeksacja obiektów o znacznych rozmiarach (porównywalnych z rozmiarem całego obrazu) wiąże się ze znacznym wydłużeniem czasu przetwarzania. W przypadku obiektów małych okazuje się jednak, że pozwala ona na uzyskanie lepszych wyników niż metoda przeglądania liniowego. Wynika to z jej „inteligencji” – analizowane są tylko punkty należące do obiektów, a nie wszystkie punkty obrazu, jak ma to miejsce w metodzie liniowej. Warto też zauważyć, że użycie szybszego komputera pozwoliłoby w większości przypadków na etykietowanie obiektów w czasie rzeczywistym pomimo swojej względnej „powolności”.

Podsumowując: obie metody umożliwiają indeksację obrazów w czasie rzeczywistym, jednak w świetle przeprowadzonych badań lepsze wyniki zapewnia podejście oparte o liniowe przeglądanie obrazu i sklejanie – jego dodatkową zaletą jest prosta implementacja sprzętowa, a realizacja etykietowania z wykorzystaniem specjalizowanego hardware’u mogłaby wiązać się ze znacznym skróceniem czasu przetwarzania i analizy obrazu.

Literatura

- [1] Adamski A.: *Inteligentne systemy transportowe: sterowanie, nadzór i zarządzanie*. Kraków, UWND AGH, 2003
- [2] Adamski A., Bubleński Z., Mikrut Z., Pawlik P.: *Image Analysis Based Automatic Safety Monitoring System For Traffic Intersections*. IV Konferencja „Telematyka i Bezpieczeństwo Ruchu Drogowego”, 19–20 XI 2004, Katowice-Ustroń, 2004
- [3] Hoose N.: *Computer Image Processing In Traffic Engineering*. RSP Ltd., Taunton, 1991
- [4] Jabłoński M.: *Od obrazu do matrycy znaku – Implementacja sprzętowa*. Seminarium Wyjazdowe nt.: Przetwarzanie i analiza sygnałów w systemach wizji i sterowania, Słok k/Belchatowa, 2002, 38–43
- [5] Jabłoński M., Gorgoń M.: *Handel-C implementation of classical component labelling algorithm*. Proceedings of the Euromicro Systems on Digital System Design, Rennes., 2004, 387–393

-
- [6] Lee A.: *VirtualDub project homepage*. <http://www.virtualdub.org>, Internet, 2004
 - [7] Parker J.R.: *Algorithms for Image Processing and Computer Vision*. Wiley Computer Publishing, New York, 1997
 - [8] Tadeusiewicz R., Korohoda P.: *Komputerowa analiza i przetwarzanie obrazów*. Wydawnictwo Fundacji Postępu Telekomunikacji, Kraków, 1997
 - [9] Umbaugh S.E.: *Computer Vision and Image Processing*. Prentice Hall, Upper Saddle River, 1998