

Zbigniew Mikrut*, Mariusz Duplaga**

Generowanie danych z obrazów bronchoskopowych w celu późniejszej klasyfikacji***

1. Wprowadzenie

Koncepcje i eksperymenty opisywane w artykule są częścią zakrojonego na szeroką skalę projektu, którego celem jest opracowanie zintegrowanego systemu wspomagającego diagnostykę bronchoskopową. Jednym z głównych zadań systemu ma być wspomaganie decyzji lekarza wykonującego zabiegi (będzie to moduł pracujący w czasie rzeczywistym). W trybie *off-line*, po archiwizacji zapisów badań, możliwa ma być analiza filmów polegająca na detekcji zmian patologicznych z wykorzystaniem biblioteki wzorców obrazów endoskopowych.

Nowoczesne bronchoskopy, cechujące się odwzorowaniem koloru w wysokiej jakości, oferujące opcje przekodowania kolorów w trakcie badania i możliwości rejestracji, nie są jeszcze szeroko rozpowszechnione. Prawdopodobnie dlatego brak jest doniesień literaturowych o próbach automatyzacji analizy tego typu obrazów, natomiast istnieją prace, demonstrujące wyniki analiz obrazów zbliżonych treściowo. Dotyczą one przede wszystkim modnej ostatnio techniki endoskopii kapsułkowej przewodu pokarmowego (WCE – *Wireless Capsule Endoscopy*). Pilotażowy charakter niniejszej pracy usprawiedliwia zawężenie zakresu badań do jednej z ważnych zmian patologicznych, jaką jest krwawienie, aczkolwiek prezentowana metoda – w przypadku powodzenia – zostanie zastosowana do detekcji innych cech, charakteryzujących się odmiennym kolorem.

Podstawową cechą obszarów krwawień jest ich charakterystyczny kolor. Z tego powodu zdecydowano się na badania „od podstaw”, czyli wykorzystanie jako bazy do segmentacji jedynie koloru pojedynczych pikseli obrazu. Detekcja na poziomie pikseli powinna zapewnić większą precyzję lokalizacji w odróżnieniu od metod polegających na wykrywaniu cech w większych oknach, np. 30×30 pikseli [6]. Jest to także istotne w przypadku wykrywania mniejszych obszarów, na przykład wybroczyn w błonie śluzowej. Na ten aspekt de-

* Katedra Automatyki, Akademia Górniczo-Hutnicza w Krakowie

** Uniwersytet Jagielloński Collegium Medicum w Krakowie

*** Wykonano w ramach grantu MNiSW, umowa nr R13 011 03

tekcji zwracają uwagę także autorzy prac [3, 4]. Drugą ważną decyzją, jaką należy podjąć na wstępie, jest wybór przestrzeni barw dla reprezentacji pikseli obrazu. Modele barw, zaproponowane przez CIE (*Comission Internationale de l'Eclairage* – Międzynarodową Komisję Oświetleniową) i przeanalizowane w pracy [5], wydają się mieć głównie zastosowanie do analizy filmów i fotografii „analogowych”. Autorzy prac, w których analizuje się obrazy WCE [2, 3, 4, 6], wskazują na przewagę modelu HSV (*Hue Saturation Value* lub HSI – *Hue Saturation Intensity*) nad RGB (*Red Green Blue*). W pracy [6] stwierdzono, że składowe H i S – w odróżnieniu od RGB – są niezależne od kierunku oglądania oraz kierunku i intensywności oświetlenia, dlatego zdecydowano o zastosowaniu reprezentacji HSV w omawianych eksperymentach.

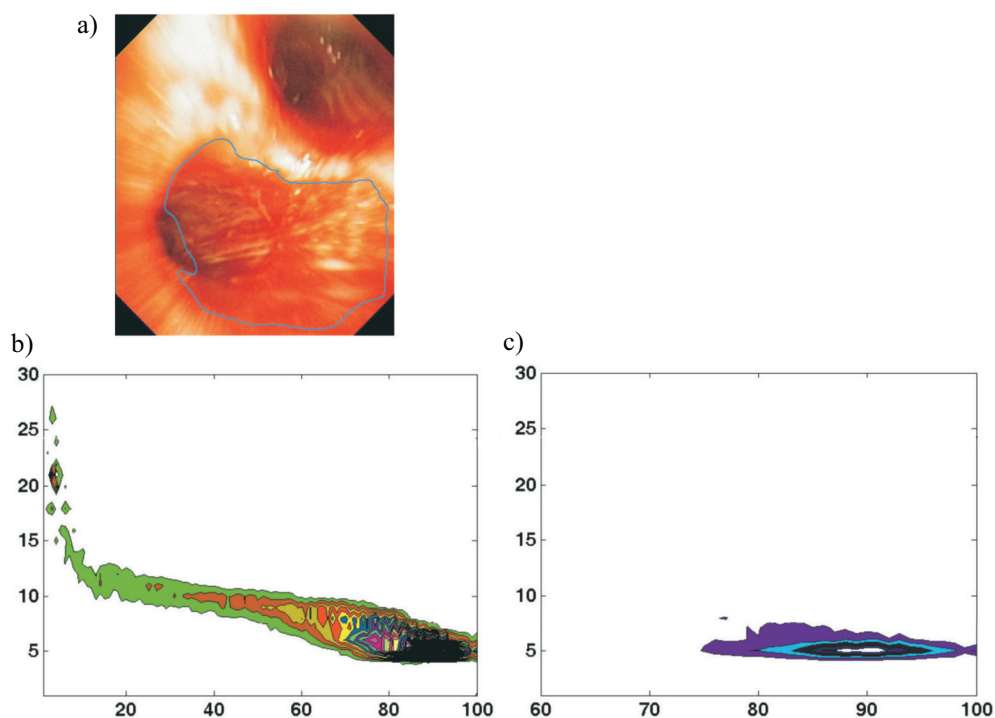
2. Dane pierwotne

Konstrukcja algorytmów automatycznej analizy filmów zarejestrowanych podczas badania bronchoskopowego zakłada wiodącą rolę lekarza – specjalisty, przeprowadzającego badanie. Na początkowych etapach tworzenia algorytmów głównym zadaniem lekarza jest wybranie z całego filmu kilkunastu ramek i zaznaczenie na nich interesujących – z diagnostycznego punktu widzenia – obszarów. Kolejną, niemniej ważną czynnością jest weryfikacja otrzymanych wyników, czyli stwierdzenie, że automatycznie wykryte obszary są istotne z diagnostycznego punktu widzenia.

Wybór ramek oraz zaznaczanie i opisywanie odbywa się za pomocą specjalnie stworzonego oprogramowania [1]. Zaznaczenie polega na narysowaniu wokół interesującego obszaru zamkniętego konturu linią o jednakowej grubości i jednolitym kolorze. Metoda ekstrakcji danych, opisana w tym artykule, zostanie przedstawiona na przykładzie kilku ramek wybranych i opisanych przez lekarza. Przykład jednej z ramek z zaznaczeniem został przedstawiony na rysunku 1a.

Wstępne badania możliwości ekstrakcji pikseli z zaznaczonego obszaru zakładały wykorzystanie do tego celu histogramów dwuwymiarowych. Po opracowaniu prostej metody wydzielania zaznaczonych obszarów generowano maski zaznaczeń i obliczano w nich dwuwymiarowe histogramy HS. Po wzięciu pod uwagę pierwotnego zakresu zmian składowych koloru (0–255) ustalono rozdzielczość histogramu 2D na 100×100 pól, uzyskując w ten sposób wstępną agregację wartości pikseli. Przykład histogramu 2D obliczonego dla całego obrazu i osobno dla obszaru zaznaczonego przedstawiono na rysunku 1b i 1c. Trzeci wymiar histogramu, czyli liczebności punktów dla współrzędnych H i S zostały zwizualizowane metodą wykreślenia poziomic. Z układu poziomic oraz z prób binaryzacji obrazu na kilku poziomach wynika, że wzięcie pod uwagę jedynie współrzędnych H i S może być niewystarczające.

W drugim wariantcie generacji danych reprezentację piksela uzupełniono o trzecią współrzędną (V). Wstępne porównanie wyników binaryzacji całego obrazu na podstawie histogramu 2D i 3D obszaru zaznaczenia z progiem równym zero (brano pod uwagę wszystkie punkty histogramu zaznaczonego obszaru) jest przedstawione na rysunku 2.



Rys. 1. Ramka z zaznaczonym obszarem „krew” (a), fragment histogramu 2D całego obrazu (b), fragment histogramu 2D obszaru zaznaczonego (c). Na rysunkach b) i c) oś pionowa odpowiada składowej H, oś pozioma – składowej S

frame₂12072.png: prog = 0.00 2D frame₂12072.png: prog = 0.00 3D



Rys. 2. W kolejności od lewej: oryginalny obraz, obraz zbinaryzowany na podstawie histogramu 2D, otrzymanego z obszaru zaznaczonego, obraz zbinaryzowany na podstawie histogramu 3D, otrzymanego z obszaru zaznaczonego. Obrazy binaryzowano na poziomie 0 ww. histogramów

Wyniki nieco się poprawiły. Równocześnie lepiej uwidocznił się wpływ niejednorodności zaznaczonego obszaru – zarówno na wielkość, jak i na rodzaj wykrytych w wyniku binaryzacji fragmentów obrazu. Powstała potrzeba eliminacji podobszarów, które musiały być zaznaczone przez lekarza, a nie są obserwowaną patologią (np. odblaski).

3. Interaktywna ekstrakcja danych

Drugi etap ekstrakcji danych polega na wykorzystaniu maski zaznaczenia oraz histogramu 2D lub 3D tego obszaru, obliczonego w przestrzeni HSV. W tym miejscu należy wyraźnie podkreślić, że proces ten nie musi obciążać lekarza – bez problemu wykona to zadanie odpowiednio przyuczony pracownik techniczny.

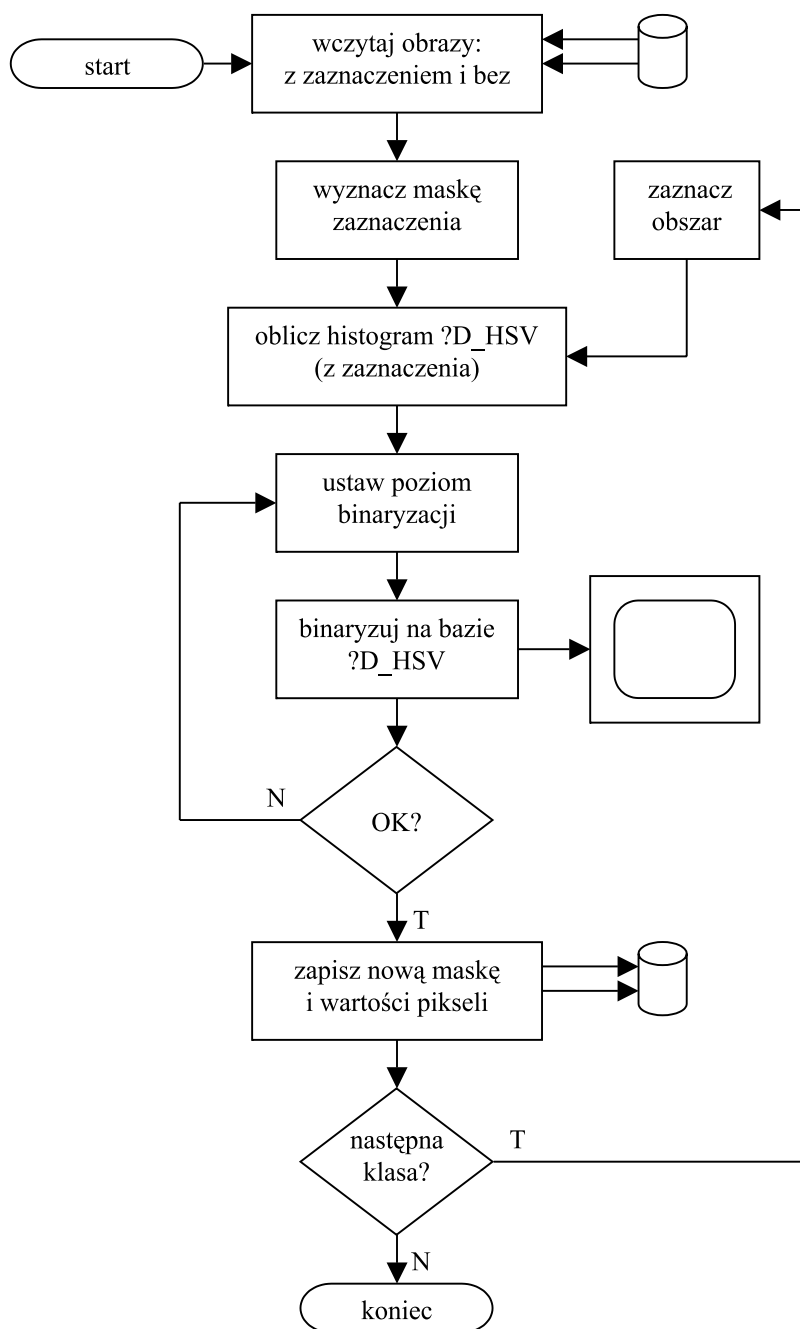
Głównym założeniem, leżącym u podstaw algorytmu ekstrakcji danych jest ilościowa dominacja pikseli danej klasy względem pikseli innych klas w analizowanym obszarze (zaznaczeniu). Schemat blokowy algorytmu przedstawiono na rysunku 3. Podstawą algorytmu jest proces binaryzacji, w wyniku którego powstają kolejne maski – mniej lub bardziej dokładnie pokrywające obszary zawierające piksele rozważanej klasy. Z kolei bazą dla binaryzacji jest dwu- lub trójwymiarowy histogram pikseli (HSV), zawartych w obszarze zaznaczonym przez lekarza. Interaktywnie zmienianym progiem binaryzacji jest liczba pikseli wchodzących w skład poszczególnych elementów histogramu.

Wyniki binaryzacji są na bieżąco obserwowane obok obrazu oryginalnego z zaznaczeniem lekarza – podobnie jak przedstawiono to na rysunku 2. Umożliwia to osobie zmieniającej próg podjęcie decyzji o zakończeniu korekty i zapisaniu nowej maski klasy „krew” na dysku. Równocześnie zapisywane są dane – wartości składowych pikseli ze zdefiniowanego obszaru, uzupełnione etykietą klasy (por. trzeci od dołu blok schematu z rysunku 3).

Przedostatni blok logiczny na schemacie daje użytkownikowi możliwość generacji danych dla innych klas. Obecnie możliwe jest określenie maski odbłasków, a następnie zdefiniowanie dopełnienia obu masek, czyli utworzenia maski i wygenerowania danych klasy „inne”. W środowisku MATLAB [8], w którym zrealizowano omawiany algorytm, wykorzystano wygodne funkcje do tworzenia menu oraz – testowo – do zaznaczania dodatkowych obszarów (innych klas), które mogą być następnie korygowane podobnie jak opisany wyżej obszar klasy „krew”.

4. Analiza porównawcza wygenerowanych danych

Ze zbioru 20 obrazów z zaznaczeniami lekarza wybrano 6 obrazów wizualnie różnych. W późniejszym etapie badań utworzą one tak zwany zbiór uczący dla klasyfikatora. Pozostałe obrazy zostaną wykorzystane jako zbiór testowy – do sprawdzenia wyników uczenia i określenia procentowego rozpoznania. Wszystkie obrazy zostały przetworzone w sposób opisany w punkcie 3 – utworzono reprezentacje HS i HSV (bazujące odpowiednio na histogramach 2D i 3D). Dla każdego obrazu powstały więc dwie maski klasy „krew” oraz cztery zestawy danych: dwa dla klasy „krew” oraz dwa dla klasy „inne”.



Rys. 3. Interaktywne wybieranie pikseli należących do różnych klas.
Znak „?” oznacza „2” lub „3”

Porównanie wygenerowanych danych miało dwa cele. Po pierwsze starano się porównać dwie reprezentacje pikseli. Drugim celem było porównanie kilku obrazów ze sobą, a dokładniej sprawdzenie, czy składowe pikseli należących do klasy „krew” są takie same dla różnych obrazów. Do porównania wykorzystano 6 obrazów przeznaczonych na zbiór uczący. Wyniki zestawiono w tabelach 1 (reprezentacja HSV) i 2 (dla HS).

Tabela 1

Porównanie zgodności: pikseli klasy „krew” (HSV) na obszarach zaznaczonych interaktywnie (górna część tabeli) i poza zaznaczeniami (dolna część tabeli) [%]

Kod ramki	1854765	1881874	1928222	200833	212072	235324	Uwagi
1854765	100	43,03	42,11	8,49	29,92	0	w_zazn.
1881874	73,13	100	38,45	15,4	30,9	0	w_zazn.
1928222	57,87	37,68	100	12,49	48,55*	0	w_zazn.
200833	7,54	11,1	17,2	100	38,37	0	w_zazn.
212072	66,55	51,17	56,32*	66,72	100	0	w_zazn.
235324	0	0	0	0	0	100	w_zazn.
1854765	0	17,7	15,95	13,38	4,44	0	poza_zazn.
1881874	2,86	0	7,52	21,22	6,74	0	poza_zazn.
1928222	5,96	23,52	0	24,97	5,24	0	poza_zazn.
200833	0,14	6,09	0,15	0	0,79	0	poza_zazn.
212072	3,84	31,99	14,57	53,48	0	0,14	poza_zazn.
235324	1,78	10,84	1,02	7,07	4,99	0	poza_zazn.

* odpowiednie obrazy pokazano na rysunku 4

W pierwszej kolumnie tabel znajdują się kody ramek tworzących reprezentację klasy „krew”, czyli piksele odpowiadające interaktywnym zaznaczeniom w tych ramkach. Z tymi pikselami porównywano piksele tych samych ramek (kody w pierwszych wierszach tabel) metodą „każdy z każdym”. Tabele składają się z dwóch części (patrz ostatnia kolumna – „uwagi”). W części górnej znajduje się zestawienie procentowych zgodności reprezentacji pikseli dla obszaru zaznaczonego. W części dolnej pokazano analogiczne zestawienie dla pikseli znalezionych poza obszarem zaznaczonym. Wysokie wartości w części górnej świadczą o wykrytej zgodności reprezentacji pikseli. Równocześnie niskie wartości w części dolnej oznaczają, że poza obszarem zaznaczonym wykryto niewiele pikseli poszukiwanej klasy. Całkowitej zgodności odpowiadają wartości na przekątnych: 100% dla górnej części tabeli i 0% dla części dolnej. Jest to przypadek idealny, osiągany jedynie podczas porównywania tych samych obrazów.

Z dwóch rozważanych reprezentacji pikseli postać HS powinna być bardziej „uogólniona” a postać HSV bardziej „precyzyjna”. Powinno się to odzwierciedlić większymi

wartościami w tabeli 2 w stosunku do wartości w tabeli 1. Dla większości pól tabel to stwierdzenie jest prawdziwe. Szczególnie duże różnice obserwuje się dla klatki oznaczonej numerem 235324, dla której w tabeli 1 (HSV) otrzymano zerowe wartości zgodności z pozostałymi obrazami, a w tabeli 2 (HS) wartości te są dość wysokie. W tabeli 1 zaznaczono kolorem szarym pola, dla których wyżej opisana reguła nie zachodzi, tzn. wartości z tabeli 1 są większe od odpowiednich pól tabeli 2. Większość tych pól jest zlokalizowana w wierszu piątym, odpowiadającym sytuacji, w której z zaznaczeniem klatki 212072 (por. rys. 1 i 2) porównywane są pozostałe obrazy.

Tabela 2

Porównanie zgodności: pikseli klasy „krew” (HS) na obszarach zaznaczonych interaktywnie (górna część tabeli) i poza zaznaczeniami (dolna część tabeli) [%]

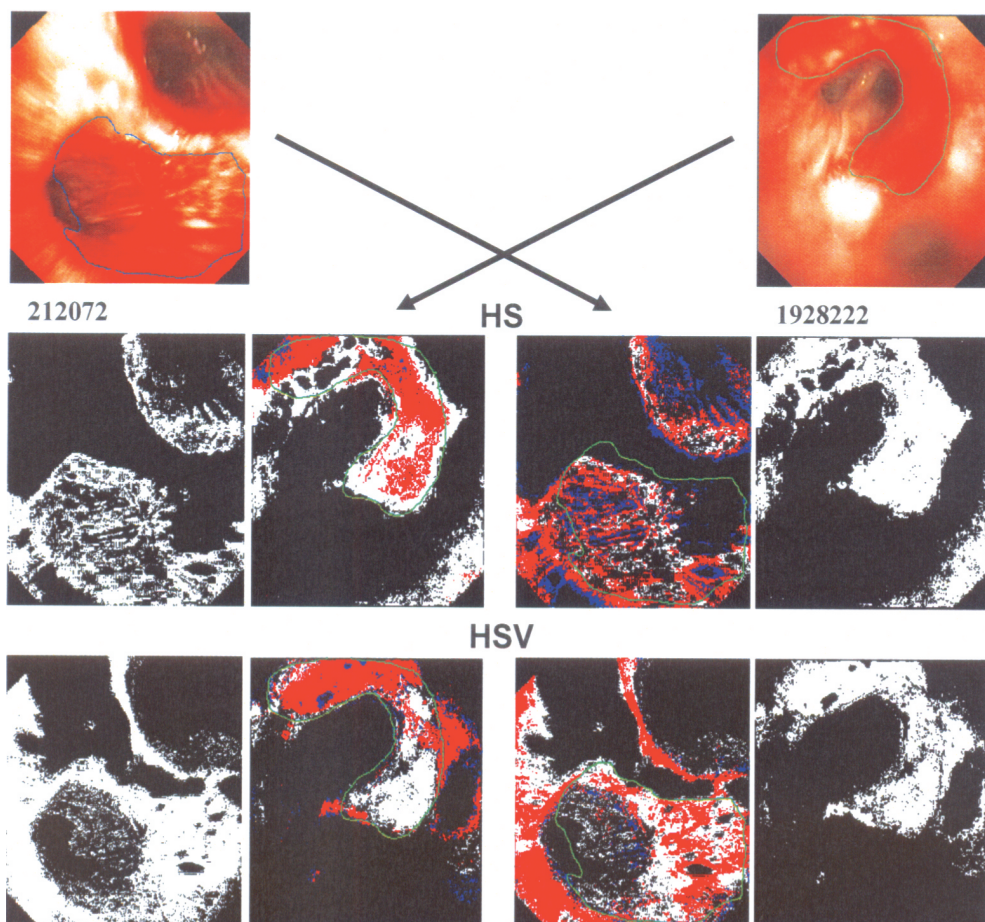
Kod ramki	1854765	1881874	1928222	200833	212072	235324	uwagi
1854765	100	47,04	85,01	7,6	69,93	36,92	w_zazn.
1881874	57,68	100	57,88	27,58	80,65	60,6	w_zazn.
1928222	80,58	47,04	100	8,07	69,93*	36,92	w_zazn.
200833	20,01	59,43	19,02	100	69,82	40,88	w_zazn.
212072	40,23	56,77	35,99*	22,4	100	32,1	w_zazn.
235324	63,24	86,02	60,26	38,57	90,93	100	w_zazn.
1854765	0	17,23	9,27	42,18	34,53	17,45	poza_zazn.
1881874	4,88	0	2,42	27,88	42,32	4,51	poza_zazn.
1928222	6,09	15,99	0	36,7	46,71	15,08	poza_zazn.
200833	6,31	15,7	4,56	0	54,11	15,18	poza_zazn.
212072	2,97	7,32	2,24	9,23	0	1,97	poza_zazn.
235324	10,14	28,12	2,83	33,14	54,95	0	poza_zazn.

* odpowiednie obrazy pokazano na rysunku 4

Podobieństwo pikseli klasy „krew” dla analizowanych obrazów powinno przejawiać się zarówno wysoką procentową zgodnością, jak i ich symetrią względem przekątnej. Jak już wspomniano, wyższe wartości obserwuje się dla reprezentacji HS. Generalnie jednak wartości te rzadko przekraczają 80%. O wiele mniejsze wartości zawiera tabela 1 (reprezentacja HSV). Średnie wartości dla górnych części tabel (bez przekątnej) wynoszą 25% i 50% odpowiednio dla reprezentacji HSV i HS (37% i 48%, jeśli nie weźmie się pod uwagę klatki 235324). W celu wyznaczenia symetrii reprezentacji przyjęto, że rozważane będą klatki o zgodności większej od około 50% oraz o różnicy zgodności poniżej ~10%. Dla reprezentacji HS warunki te spełniły jedynie trzy pary klatek, które zaznaczono w tabeli 2 kolorem szarym. W tabeli 1 oznaczono także podwójną ramką jedną parę klatek HSV, która spełniła

powyższe warunki. Położenia klatek o symetrycznych reprezentacjach w tabelach 1 i 2 nie pokrywają się.

Wyniki procentowe, zestawione w tabelach 1 i 2, dają ogólny pogląd na wzajemne podobieństwo pikseli klasy „krew” wyekstrahowanych z różnych klatek. Późniejsza analiza obrazów powinna uwzględniać zarówno fakt wykrycia pikseli tej klasy, jak i ich rozmieszczenia. W szczególności istotne jest występowanie spójnych obszarów, a nie rozproszonych małych grup pikseli. W większości przypadków ten postulat jest spełniony dla obu typów reprezentacji pikseli.



Rys. 4. Porównanie wzajemnej zgodności pikseli klasy „krew” dla dwóch ramek: 212072 i 1928222.

Kolumny 1 i 4 – od góry: obraz oryginalny, zaznaczenie interaktywne dla reprezentacji HS, zaznaczenie interaktywne dla reprezentacji HSV. Kolumny 2 i 3 – wizualizacja rozłożenia pikseli: kolor biały – zaznaczenie pierwotne, kolor czerwony: piksele z drugiego obrazu zgodne z pierwszym, kolor niebieski: piksele obrazu drugiego poza zaznaczeniem

Przykład wzajemnej zgodności wykrytych pikseli dla dwóch obrazów pokazano na rysunku 4. Do wizualizacji wybrano obrazy o średnim stopniu zgodności, które było w przybliżeniu symetryczne dla reprezentacji HSV i niesymetryczne dla HS. Procentowe wyniki porównań tych obrazów (kody: 212072 i 1 928222) oznaczono w tabelach 1 i 2 gwiazdkami. Interesujące jest porównanie wyników interaktywnego zaznaczania dla obu obrazów i obu typów reprezentacji, pokazane w pierwszej i ostatniej kolumnie rysunku 4. Dla ramki 1 928222 (ostatnia kolumna) są one bardzo podobne. Nieco większe różnice można zaobserwować dla klatki 212072 (pierwsza kolumna). Częściowo wynikają one z doboru progów, ale w tym przypadku uwidacznia się większa precyzja reprezentacji HSV. Osoba dokonująca interaktywnej binaryzacji dążyła do wyeliminowania obszaru wlotu drzewka oskrzelowego, w większej części zaznaczonej przez lekarza, co udało się jedynie dla reprezentacji HSV. W obu przypadkach i dla obu obrazów wykryte piksele klasy „krew” tworzą spójne obszary lub nieco mniej spójne rejony, które łatwo można dopełnić przy pomocy podstawowych operacji morfologicznych [7].

5. Podsumowanie i wnioski

Celem prezentowanych eksperymentów było sprawdzenie, która z reprezentacji pikseli – HS czy HSV – bardziej nadaje się do utworzenia reprezentacji wybranej klasy (w tym wypadku: klasy „krew”), przydatnej w późniejszym etapie rozpoznawania. Pierwszym etapem badań było stworzenie interaktywnego algorytmu korekty obszarów, zaznaczonych na klatkach badania bronchoskopowego przez lekarza. Za pomocą tego algorytmu dokonano ekstrakcji pikseli zadanej klasy z sześciu wizualnie różniących się klatek. Ekstrakcji dokonano dla dwóch reprezentacji pikseli (HS i HSV) przy założeniu, że w zaznaczonym przez lekarza obszarze znajduje się najwięcej pikseli klasy „krew”. Reprezentacje porównano ze sobą, a wyniki zamieszczono w tabelach 1 i 2. Przykład wizualizacji wyników pokazano na rysunku 4.

Wyniki otrzymane dla dwóch reprezentacji różnią się. Reprezentacja HSV zapewnia większą precyzję różnicowania pikseli. Reprezentacja HS dokonuje uogólnienia. Można to zaobserwować zarówno na rysunku 4, jak i w tabelach 1 i 2, na przykładzie wyników dla klatki 235324 (dla reprezentacji HSV otrzymano zerowe procentowe zgodności z pozostałymi obrazami, dla reprezentacji HS – dość wysokie wartości). Przeprowadzono także eksperyment dodatkowy polegający na porównaniu wybranych sześciu obrazów z klatką nie zawierającą krwawień. Otrzymano minimalne wartości pokryć (dla HSV rzędu 0,1%, dla HS rzędu 1,6%), przy czym wykryte piksele nie tworzyły spójnych obszarów. Zróżnicowane wartości zestawione w tabelach 1 i 2 oraz mała liczba obszarów symetrycznych świadczą o prawidłowym wyborze obrazów, które powinny być różnorodne, ponieważ – na etapie rozpoznawania – będą tworzyć zbiór uczący dla klasyfikatorów. Natomiast nie jest możliwe stwierdzenie, która reprezentacja bardziej nadaje się do klasyfikacji, a więc eksperymenty dotyczące rozpoznawania należy przeprowadzić dla obu.

Perspektywy dalszych badań związane są z algorytmem interaktywnej korekty obszarów zaznaczonych przez lekarza. Obecnie korekta jest oparta na stopniowym progowaniu histogramów 2D (HS) i 3D (HSV), przy czym zmiennym progiem jest liczba pikseli w poszczególnych polach odpowiednich histogramów. Celowe byłoby sprawdzenie efektów korekty dla algorytmu zmodyfikowanego, w którym zmiana progu wiązałaby się dodatkowo z wybieraniem pikseli należących do najbliższego otoczenia już wybranych pól histogramu. W ten sposób w większym stopniu uwzględniono by wzajemne podobieństwo reprezentacji pikseli, a nie tylko ich liczbę.

Literatura

- [1] Duplaga M., Leszczuk M., Przelaskowski A., Janowski L., Zieliński T., *BRONCHOVID – zintegrowany system wspomagający diagnostykę bronchoskopową*. Przegląd Lekarski t. 64, wyd. spec.: Inżynieria medyczna, 2007, 42–48.
- [2] Giritharan B., Yuan X., Liu J., Buckles B., Oh J.H., Tang S.J., *Bleeding Detection from Capsule Endoscopy Videos*. 30th Annual International IEEE EMBS Conference, Vancouver, British Columbia, Canada, August 20–24, 2008.
- [3] Jung Y.S., Kim Y.H., Lee D.H., Kim J.H., *Active Blood Detection in a High Resolution Capsule Endoscopy using Color Spectrum Transformation*. Proc. International Conference on BioMedical Engineering and Informatics BMEI 2008, May 28–30, Sanya, Hainan, China, 859–862.
- [4] Lau P.Y., Correia P.L., *Detection of bleeding patterns in WCE video using multiple features*. Proceedings of the 29th Annual International Conference of the IEEE EMBS Cité Internationale, Lyon, France, August 23–26, 2007,
- [5] Lee H.C., *Introduction to Color Imaging Science*. Cambridge University Press, Cambridge, 2005.
- [6] Li B., Meng M.Q-H., *Computer Aided Detection of Bleeding in Capsule Endoscopy Images*. IEEE Canadian Conference on Electrical and Computer Engineering, Niagara Falls, 2008.
- [7] Tadeusiewicz R., *Systemy wizyjne robotów przemysłowych*. WNT, Warszawa, 1992.
- [8] Thompson C.M., Shure L., *Image Processing Toolbox for use with Matlab*. The MathWorks, Inc., Natick, 1995.