

Proszę niczego nie wstawiać w nagłówku

WYZNACZANIE FAKTUR WE WSTĘPNYM PRZYGOTOWANIU OBRAZÓW DLA CELÓW OBRAZOWEJ BAZY DANYCH

Tatiana Jaworska

Instytut Badań Systemowych, Polska Akademia Nauk,
ul. Newelska 6, 01-447 Warszawa,

W artykule przedstawiono rozpoznawanie faktur elementów architektonicznych (cegły, dachówki, itp.) we wstępnym przygotowaniu obrazów do obrazowej bazy danych nieruchomości. Segmentacja takich elementów powoduje rozdzielenie poszczególnych elementów faktury, zamiast wyszczególnienia elementu fakturowanego jako jednego segmentu obrazu. Standardowe metody wydzielenia faktur oparte na sąsiedztwie pikseli wymagają zmienności (granulacji) faktury w mniejszym obszarze. Zatem w niniejszym artykule opisano metodę wydzielenia faktur opartą na transformacji falkowej. Odległości pomiędzy rzędami cegieł lub dachówek są charakteryzowane przez odpowiednie wielkości współczynników transformacji falkowej.

1. Wprowadzenie

Zalew obrazami cyfrowymi postępuje lawinowo wraz z rozwojem coraz szybszych i pojemniejszych łączy internetowych. Istniejąca ilość obrazów powoduje konieczność ich magazynowania w bazach danych, przetwarzania, rozpoznawania i wyszukiwania. Coraz częściej zaczynają powstawać duże bazy obrazowe. Bazy te różnią się między sobą w zależności od rodzaju przechowywanych obrazów i celów do jakich mają służyć. Inaczej będzie zatem konstruowana baza obrazów dla galerii i muzeów, a zupełnie inaczej, jeśli będzie to system informacji geograficznej (GIS) lub kolekcja zdjęć satelitarnych do celów geologicznych lub prognoz pogody. Jeszcze inne mechanizmy przeszukiwania będą wykorzystywane w policyjnych bazach odcisków palców i zdjęć twarzy przestępców, a inne przy obrazach kosmosu dla astronomów.

Dotychczas powstające bazy obrazowe ograniczały się jedynie do rozpoznawania rodzaju obrazu, czy jest to dom, pies, czy samolot (Ogle, 1995). Następnie, kiedy użytkownik takiej bazy zadawał pytanie do systemu np. o łąkę z żółto-pomarańczowymi kwiatami, system wybierał jeden lub kilka obrazów o podobnym kolorowym histogramie (Flickner, 1995).

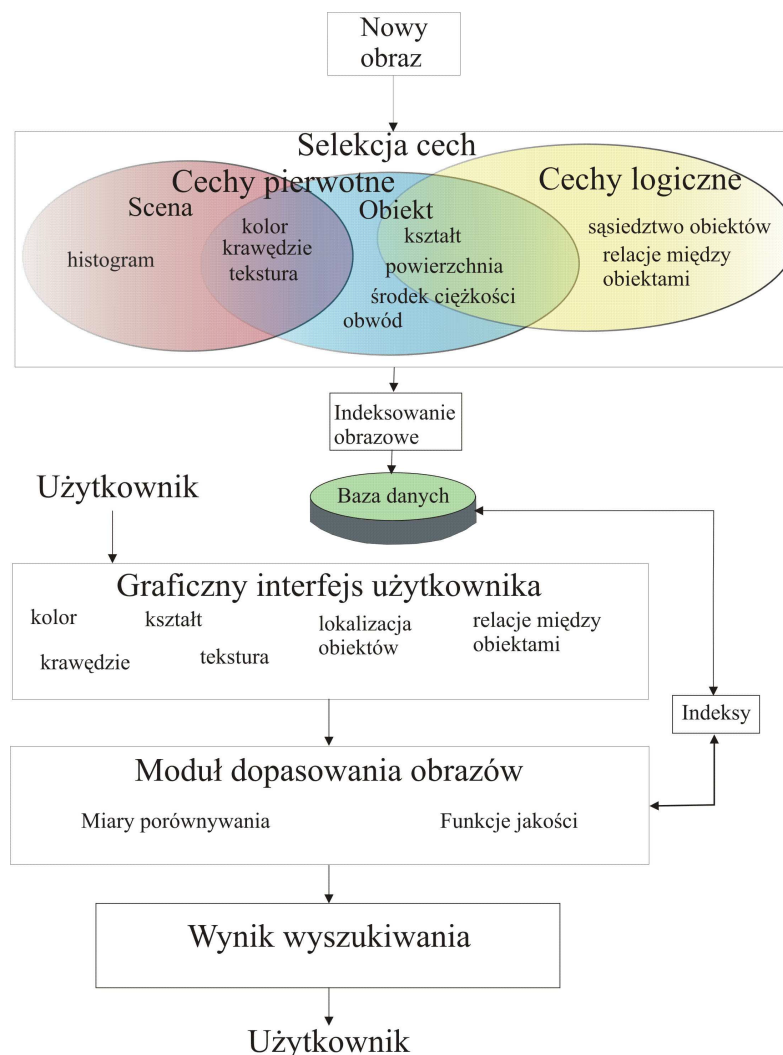
2. Struktura obrazowej bazy danych

W naszym projekcie zajmujemy się budową bazy zdjęć nieruchomości. W tym celu są wykorzystywane zdjęcia domów jednorodzinnych pobrane z Internetu w formacie JPEG. Baza danych i mechanizmy analizy obrazu są zaimplementowane za pomocą pakietu MatLab (wersja 7.1) ze wspomaganiami następujących ToolBoxów: Image Processing, Fuzzy, Statistics, Wavelet i Data Base. Baza danych jest dostępna

Ani w stopce

Proszę niczego nie wstawiać w nagłówku

w trybie interaktywnym przez Internet. Interfejs internetowy jest wykonany w Perlu (Jaworska, 2005).



Rysunek 1. Schemat obrazowej bazy danych

Każdy nowy obraz dołączany do bazy musi przejść obróbkę wstępną, która na rys. 1. jest przedstawiona jako selekcja cech. W ramach tego przygotowania muszą być wydzielone wszystkie interesujące nas elementy architektoniczne (takie jak okna, drzwi, dachy, itp.) oraz muszą one zostać oddzielone od otoczenia. Na tym etapie bardzo ważne jest, aby właściwie dokonać segmentacji obrazu i wyznaczyć cechy obiektów, ponieważ na tych cechach będzie się dalej opierać indeksowanie

Ani w stopce

Proszę niczego nie wstawiać w nagłówku

bazy i jej przeszukiwanie. Po wyznaczeniu wszystkich cech pierwotnych obiektów i ich cech logicznych (np. sąsiedztwa) następuje dołączenie obiektów do bazy z jednoczesnym ich indeksowaniem obrazowym, co oznacza, że indeksy opierają się na cechach obrazów, a nie na elementach tekstowych, np. nazwie pliku z obrazem.

Aby z tak utworzonej bazy obrazowej mogli korzystać użytkownicy wprowadzony został graficzny interfejs użytkownika, który jest już widoczny na stronie internetowej instytutu. „Graficzny interfejs użytkownika” (patrz rys. 1) jest podstawowym narzędziem do zadawania zapytań do bazy. Musi mieć on zatem możliwość łatwego pokazania użytkownikowi o jakie cechy może zapytać. Musi też istnieć możliwość wyboru (np. przesunięcia suwakiem) koloru, kształtu, tekstury, lub typu szukanego elementu, jego atrybutów, itp. W niektórych bazach istnieje nawet interfejs graficzny do zadawania pytań, czyli możliwość naszkicowania interesującego użytkownika kształtu lub podania kształtu przykładowego (Flickner, 1995). Można też podać obraz o interesującym użytkownika histogramie kolorów i wtedy nastąpi przeszukiwanie bazy w celu znalezienia obrazów o analogicznym rozkładzie kolorów, itd. W naszym systemie pokazują się poszczególne elementy architektoniczne uporządkowane w określone kategorie (oddzielnie dachy, okna, drzwi, itp.).

Odpowiedzią na zapytanie użytkownika jest konkretny obraz, lub zbiór obrazów, który jest wynikiem przeszukiwania bazy. To, na ile ten obraz spełnia oczekiwania użytkownika zależy już od następnego modułu, wyspecyfikowanego na rys. 1, jako „Moduł dopasowania obrazów”. Zawiera on odpowiednie miary podobieństwa obrazów, które również bazują na cechach obrazów wyznaczonych w procesie obróbki wstępnej.

3. Segmentacja wstępna

W sytuacji, w której kolorowe obrazy domów są pobierane z Internetu i ich wstępna obróbka jest prowadzona w sposób nienadzorowany, kluczowym zagadnieniem staje się takie wyekstrahowanie obiektów z obrazu, żeby nie trzeba było ingerować w dalsze etapy włączania tych obrazów do bazy danych. W tym celu stosujemy dwu-etapową segmentację umożliwiającą dokładne wydzielenie interesujących nas obiektów z obrazu. Ze względu na fakt, że w niniejszym artykule koncentrujemy się na pokazaniu metody wyekstrahowania faktury, jako przykład został wybrany fragment domu z ceglana ścianą. Wszystkie etapy segmentacji będą pokazywane na tym właśnie przykładzie widocznym na rys. 2.

W pierwszym etapie obraz jest poddawany segmentacji przy użyciu rozmytej analizy skupień (*fuzzy c-means clustering*) (Melin, 2005), w wyniku której otrzymujemy wstępny podział obrazu. Dobre efekty uzyskuje się przy podziale na 5-6 skupień. Wtedy dobrze widoczne są takie elementy jak: dach, okna, itp. Rysunek 3 pokazuje właśnie obraz narożnika domu po podziale na 5 skupień. Widać wyraźnie, że poszczególne cegły są tu traktowane jako pojedyncze obiekty, co uniemożliwia wydzielenia ściany jako jednego obiektu.

Ani w stopce

Proszę niczego nie wstawiać w nagłówku

Drugim etapem obróbki wstępnej było rozdzielanie poszczególnych skupień tak, żeby dostać obraz binarny – jedno białe skupienie, czarna reszta obrazu. W ten sposób dostajemy tyle obrazów binarnych, ile było wyznaczonych skupień. Następnie znajdujemy krawędzie obiektów algorytmem Canny (Canny, 1986). Algorytm ten został wybrany spośród innych algorytmów do wyszukiwania krawędzi ponieważ z jego użyciem dostajemy zamknięte kontury obiektów (Davis, 1975). Na podstawie uzyskanych zamkniętych konturów jesteśmy w stanie wyszczególnić pojedyncze obiekty. Po takiej segmentacji dostajemy zbiór oddzielnych ponumerowanych obiektów dla każdego skupienia. Problemem jest jednak wydzielenie każdej cegły jako pojedynczego obiektu oraz odbicia w szybach, które uniemożliwiają wydzielenie każdej szyby jako pojedynczego obiektu.



Rysunek 2. Przykładowy obraz fragmentu domu poddawany obróbce wstępnej

4. Wydzielenie elementów z fakturą

Jednym z narzędzi służącym do charakteryzowania faktury jest dwuwymiarowa transformacja częstotliwościowa. Poza klasyczną transformacją Fouriera, są też stosowane różne operatory „wrażliwe” na fakturę, np. Laplasjan 3x3 lub 5x5, Gaussian 5x5, Hursta, Haralicka, lub Frei&Chena (Russ, 1995). Wszystkie one są przydatne przy fakturach o małej granulacji.

Ani w stopce

Proszę niczego nie wstawiać w nagłówku

image labeled by cluster index



Rysunek 3. Wynik segmentacji obrazu na 5 skupień

Następną metodą rozpoznawania faktury jest histogram dwuwymiarowy. Opiera się on na prawdopodobieństwie wystąpienia podobnego piksela w podobnej odległości. Jest on jednak kłopotliwy w implementacji. Metoda progowania histogramu bywa przydatna do obrazów monochromatycznych jeśli chcemy wyróżnić 2-3 rejony z fakturą.

Zastosowanie analizy falkowej do opisu tekstur i wzorów zaproponowali Balmelli i Mojsilović w 2001 (Balmelli, 2001). Znajdowali oni cechy charakterystyczne obrazów regularnych, ornamentów roślinnych i wzorów geometrycznych i dla każdego rodzaju faktury konstruowali wektor cech dystynktywnych. Nie zaproponowali jednak żadnego połączenia swojej metody z obrazami rzeczywistymi.

Ani w stopce

4.1 Wstęp do transformacji falkowej obrazów

Dwuwymiarowe transformacje falkowe obrazu sprowadzają się do rzutowania obrazu na falki ustalonego ortonormalnego układu falek (Daubechies, 1992), (Mallat, 1998). Możemy wyróżnić dwuwymiarowe transformaty ciągłe, diadyczne i dyskretne. Ze względu na dyskretny charakter obrazów, będziemy się dalej posługiwać falkami dyskretnymi, zaimplementowanymi w Toolboxie Wavelet programu MatLab wersja 7.1.

Dwuwymiarowy separowany układ falek ortonormalnych utworzony konwencjonalnie na podstawie układu jednowymiarowego ma postać:

$$\{\psi_{j,k}(x), \psi_{l,m}(y)\}_{(j,k,l,m) \in \mathbb{Z}^4} \quad (1)$$

gdzie $\psi_{j,k}(x)$ jest k -tą falką jednowymiarową na poziomie j -tym analizy wielorozdzielczej.

Wykorzystując założenia analizy wielorozdzielczej można pokazać (co zrobił już Mallat w 1989r. (Mallat, 1989)), że wyżej zaproponowany układ jest nadmiarowy i można go zredukować do następującego układu:

$$\{\psi_{j,k,m}^1(x, y), \psi_{j,k,m}^2(x, y), \psi_{j,k,m}^3(x, y)\}_{(j,k,m) \in \mathbb{Z}^3} \quad (2)$$

gdzie:

$$\begin{aligned} \psi^1(x, y) &= \phi(x) \psi(y) \\ \psi^2(x, y) &= \psi(x) \phi(y) \\ \psi^3(x, y) &= \psi(x) \psi(y) \end{aligned} \quad (3)$$

Podzbiory falek z indeksami górnymi 1,2,3 nazywane są odpowiednio dwuwymiarowymi falkami horyzontalnymi, wertykalnymi i diagonalnymi. Podobnie rzuty obrazu na te falki nazywane są horyzontalnymi, wertykalnymi i diagonalnymi detalami obrazu. Konwencję wizualizacji tych detali przedstawia rysunek 4.

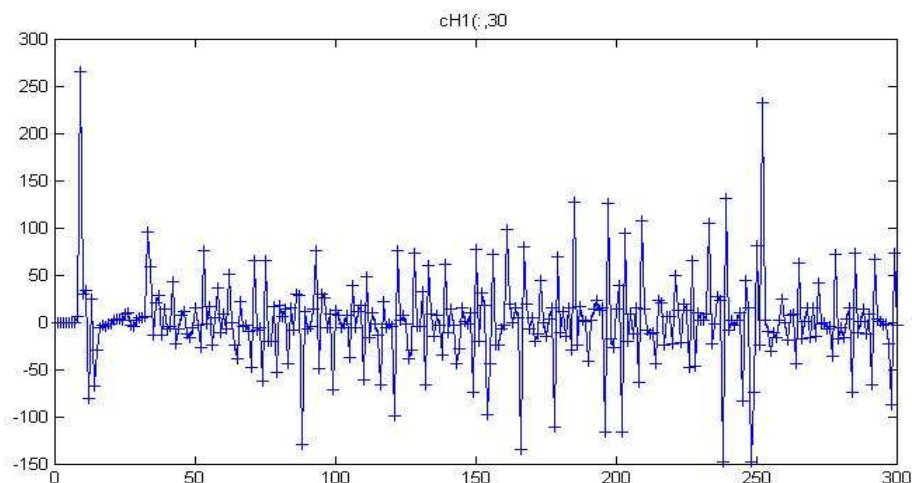
a_3	d_3^2	d_3^2	
d_3^1	d_3^3	d_3^2	
d_2^1	d_2^2	d_1^2	
d_1^1	d_1^1	d_1^1	

Rysunek 4. Konwencja wizualizacji detali $d_{j,k,l}^p = \langle f, \psi_{j,k,l}^p \rangle$ dwuwymiarowej transformaty falkowej, gdzie a_j oznacza aproksymację na poziomie j -tym

Proszę niczego nie wstawiać w nagłówku

Dzięki temu, że szybka transformacja falkowa (FWT) jest zaimplementowana w Toolboxie Wavelet programu MatLab wersja 7.1 narzędzie to staje się odpowiednio wydajne do częstego użycia i dlatego postanowiliśmy zastosować je do naszych celów.

4.2 Mapa odległości oparta na transformacji falkowej



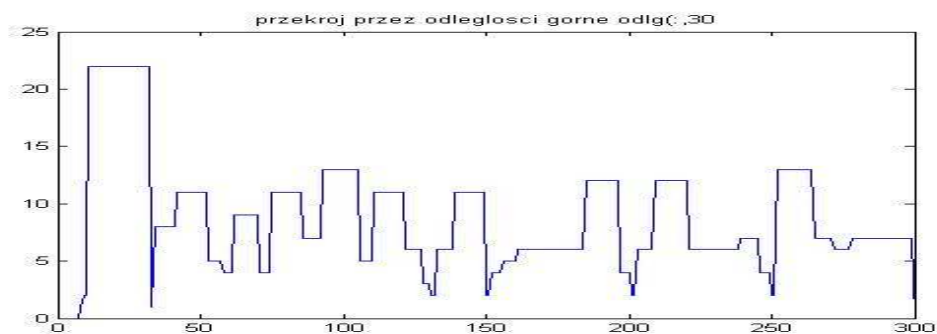
Rysunek 5. Przekrój przez detale horyzontalne dla 30-tej kolumny (falka Haara, $j=1$)

Należy zauważyć, że jedną z własności detali jest ich kierunkowość. Oznacza to, że detale horyzontalne mają duże wartości, gdy spłot falki następuje z poziomymi elementami obrazu i odwrotnie, detale wertrykalne mają duże wartości, gdy spłot falki następuje z pionowymi elementami obrazu (stąd pokazywane kierunki na rys. 4). Jeśli skorzysta się z tej cechy splatając obraz regularnych dachówek, czy cegieł z odpowiednimi falkami otrzyma się transformatę dwuwymiarową, której wartości maksymalne będą się pokrywały z połączeniami tych dachówek lub cegieł. Do wszystkich naszych obliczeń stosujemy falki Haara (Haar, 1910). Są one najprostsze, ale najlepiej się sprawdzają w przypadku elementów geometrycznych. Przekrój przez 30-ty wiersz macierzy detali horyzontalnych (oznaczanych przez program MatLab jako cH1) pokazuje rysunek 5. Na osi odciętych odłożone są numery falek Haara dla pierwszego poziomu analizy wielorozdzielczej, natomiast na osi rzędnych - wartości współczynników cH1. Maksima i minima na tym wykresie to punkty pomiędzy cegłami na omawianym przez nas rysunku 2.

Można teraz stworzyć wykres odległości między tymi maksymalnymi lub minimalnymi wartościami i okazuje się, że odległości (które odpowiadają odległościom między cegłami) stają się dobrym parametrem wyróżniającym ten właśnie obszar. Taki wykres odległości pomiędzy wartościami maksymalnymi jest pokazany na rysunku 6. Próg odcięcia jest przyjęty dla 2% wartości maksymalnej z całej macierzy.

Ani w stopce

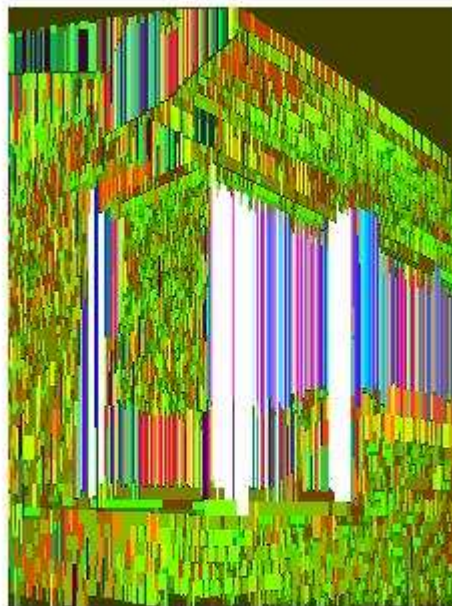
Proszę niczego nie wstawiać w nagłówek



Rysunek 6. Odległości wyznaczone dla progu odcięcia 2% wartości maksymalnej dla całej macierzy detali

Teraz tworzymy mapy odległości między poszczególnymi maksymalnymi lub minimalnymi wartościami dla całej macierzy cH1. Takie mapy odległości górnych i dolnych są pokazane odpowiednio na rysunkach 7 i 8.

Mapa odleglosci gornych dla cH1

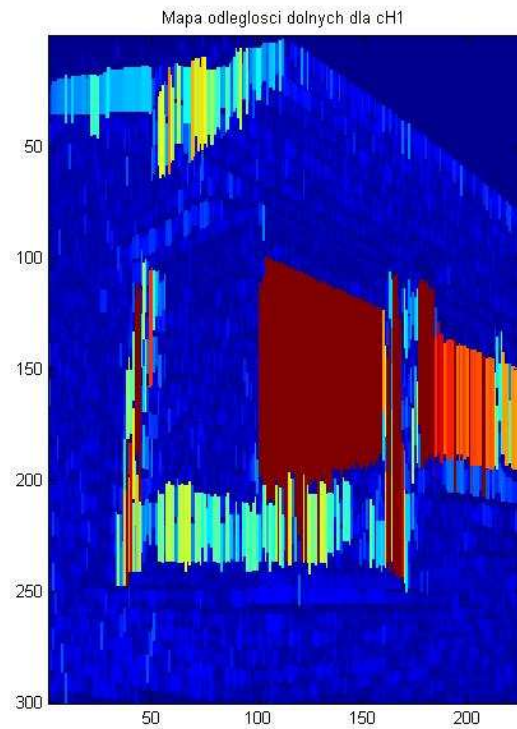


Rysunek 7. Mapa odległości górnych, czyli policzonych na podstawie maksymalnych wartości detali horyzontalnych

Ani w stopce

Proszę niczego nie wstawiać w nagłówku

Po stworzeniu takich map trzeba wybrać obszary nas interesujące, czyli takie które zawierają w naszym przypadku cegły. Pomaga w tym policzenie histogramu odległości pokazanego na rysunku 9. Nie pokazuje on liczby odcinków o poszczególnych odległościach, lecz liczbę pikseli, którym przyporządkowano poszczególne odległości. Zatem pierwsza wartość 0 odpowiada niebu na rysunku 2. A dalej są już piksele należące do poszczególnych odległości. W omawianym przypadku długość cegieł została oszacowana na 5-9 długości falek dla $j=1$, czyli 10-18 pikseli obrazu oryginalnego. (Należy pamiętać, że transformaty są liczone przez podpróbkowanie.)

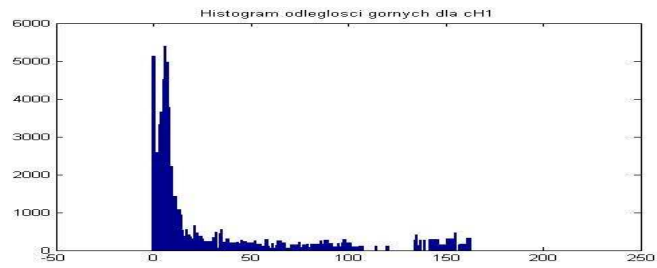


Rysunek 8. Mapa odległości dolnych, czyli policzonych na podstawie minimalnych wartości detali horyzontalnych. W tym wypadku próg odcięcia został ustawiony na 5% wartości minimalnej dla całej macierzy współczynników cH1.

Aby wyznaczyć obszary z fakturą wybieramy z całej mapy odległości obszary z długościami nas interesującymi i generujemy następną mapę już częściowo zagregowaną, ale jeszcze ciągle w przestrzeni transformaty – rysunek 10. Następnie interesujące nas obszary nakładamy już na oryginalny obraz i otrzymujemy rysunek 11. Jak można na nim zauważyć wszystkie obszary z cegłami są zasłonięte, czyli dobrze wyznaczone.

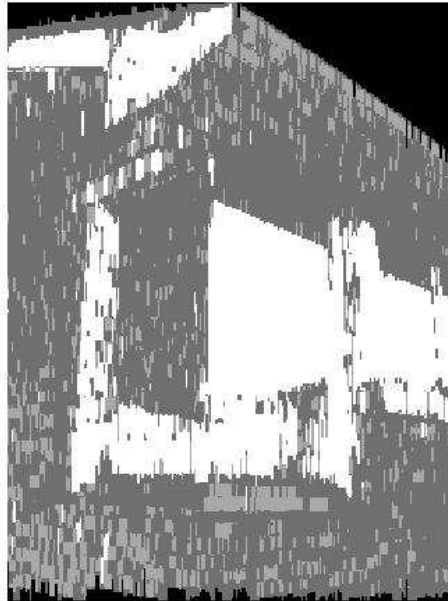
Ani w stopce

Proszę niczego nie wstawiać w nagłówku



Rysunek 9. Histogram odległości górnych

image labeled by cluster index average distance cv1



Rysunek 10. Mapa interesujących nas odległości. Odległości zbyt małe są zaznaczone na czarno, a odległości większe od progu na biało.

5. Analiza powstających błędów

Rysunek 11 pokazuje zasłonięte nie tylko obszary z cegły, ale również deski pod dachem i część odbijających się w szybie okna drzew oraz fragmenty dachu (prawa część okna). Odbicia różnych nieliniowych elementów też dają duże wartości

Ani w stopce

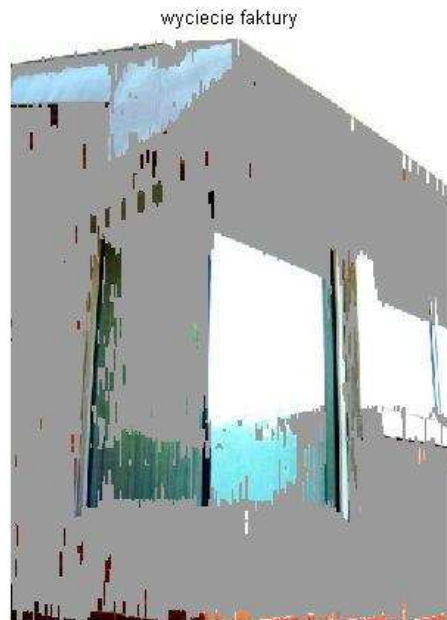
Proszę niczego nie wstawiać w nagłówku

detali, a nie należą bezpośrednio do interesującego nas obszaru. Gdyby to były elementy roślinne, można by je rozpoznać innymi metodami [Jaworska 2006] (choćby po kolorze), ale to jest okno, które chcemy wydzielić jako jeden element architektoniczny. Jak na razie nie opracowaliśmy metody wyznaczania elementów z refleksami, bo w przeciwieństwie do grafiki komputerowej nie mamy otoczenia rozpoznawanych obiektów.

W przypadku desek na dolnej stronie dachu, można je odróżnić od cegieł używając drugiego parametru, a mianowicie odległości poziomej, który uzyskuje się analogicznie, licząc tym razem detale wertykalne i tworząc mapy odległości poziomych.

6. Podsumowanie i kierunki dalszych prac

Jak widać przygotowanie wstępne obrazu z wydzieleniem znajdujących się na nim elementów jest zadaniem trudnym i musi składać się z wielu etapów. Zadanie jest tym trudniejsze, że staramy się stworzyć system pracujący automatycznie, bez pomocy człowieka nawet w trudniejszych przypadkach. W wyniku naszych prac część elementów teksturowanych i roślinnych jest już możliwa do rozpoznania i wydzielenia z tła. Elementy z refleksami wymagają dalszych prac, które są obecnie prowadzone.



Rysunek 11. Obszary z fakturą zaznaczone na jednolicie szaro i nałożone na oryginalny obraz.

Ani w stopce

Proszę niczego nie wstawiać w nagłówku

Po wydzieleniu poszczególnych obiektów wyznacza się następnie cechy tych obiektów, takie jak: kolor, obwód, powierzchnia, środek ciężkości, momenty główne bezwładności, osie bezwładności, mimośród, kod granic oraz położenie w obrazie. Po ich policzeniu każdy obiekt jest charakteryzowany przez wektor cech. Poza cechami pojedynczych obiektów opisuje się również cechy logiczne, takie jak: położenie obiektów w całym obrazie, położenie obiektów względem siebie, itp. Na podstawie powyżej wyznaczonych cech obraz jest dołączany do bazy danych i odpowiednio indeksowany - też w sposób obrazowy.

Dotychczas wykonana segmentacja nie jest jeszcze zgodna z klasyfikacją w rozumieniu ludzkiego mózgu, który przypisuje od razu desygnaty widzianym obiektom. Jeśli widzimy obiekt trójkątny to zostanie on częściej zaklasyfikowany jako dach, a okrągły i ciemny jako okno.

Głębsza analiza semantyczna jest planowana w przyszłości. Będzie wtedy użyte rozumowanie oparte na przypadkach i wnioskowanie rozmyte. Zapewni ona łatwiejszą komunikację z użytkownikiem dzięki stworzeniu możliwości zadawanie rozmytych zapytań do bazy.

Literatura

- Balmelli L., Mojsilović A. (2001) Wavelet domain features for texture/pattern description, classification and replicability analysis. W: A. Petrosian i F. Meyer, red. *Wavelets in Signal and Image Analysis*. Kluwer Academic Publishers, Londyn, rozdz. 7, 193-212.
- Canny J. (1986) A Computational Approach to Edge Detection, W: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, Nr. 6, 679-698.
- Daubechies I. (1992) Ten lectures on wavelets, Society for Industrial and Applied Math., Filadelfia.
- Davis L. S. (1975) A Survey of Edge Detection Techniques, *CGIP* Vol. 4, 248-270
- Flickner M., Sawhney H., et al. (1995) Query by Image and Video Content: The QBIC System, W: *IEEE Computer*, Vol. 28, Nr. 9, sierpień, 23-32.
- Haar A. (1910) Zur theorie der orthogonalen funktionensysteme. *Math. Annal.*, 69, 331-371
- Jaworska T. (2006) Przygotowanie wstępne obrazu do obrazowej bazy danych, W: *Materiały konferencyjne IWSE06*, czerwiec, 1-10, Wrocław, (w druku).
- Jaworska T., Partyka A. (2005) Obrazowe bazy danych, Raport Badawczy RB/37/2005, IBS PAN.
- Mallat S. (1998) A wavelet tour of signal processing, Academic Press, Londyn.
- Mallat S. (1989) Multiresolution approximations and wavelet orthonormal basis of $L^2(\mathbb{R})$. W: *Trans. Amer. Math. Soc.*, 315, 69-87, wrzesień.
- Melin P., Castillo O. (2005) Hybrid Intelligent Systems for pattern Recognition Using Soft Computing, Springer, Berlin, 169-184.
- Ogle V., Stonebraker M. (1995) CHABOT: Retrieval from a Relational Database of Images, *IEEE Computer*, Vol. 28, No. 9, sierpień, 40-48.
- Russ J. C. (1995) The image processing. Handbook, CRC, Londyn, 361-385.

Ani w stopce