AUTOMATYKA • 2006 • Tom 10 • Zeszyt 3

Piotr Pawlik*, Sławomir Mikrut**

Wyszukiwanie punktów charakterystycznych na potrzeby łączenia zdjęć lotniczych***

1. Wstęp

Artykuł dotyczy problemu dopasowania fotogrametrycznych zdjęć lotniczych. Istotą zagadnienia jest znalezienie i pomiar odpowiednich punktów charakterystycznych na dwóch zdjęciach lotniczych (fotogrametrycznych), wykonanych z odpowiedniej bazy. Pomiar tych punktów jest niezbędny do opracowań fotogrametrycznych (wykonywania map cyfrowych, cyfrowych modeli terenu itp.). Wyszukiwanie i pomiar tych punktów w sposób automatyczny pozwoli na podniesienie efektywności opracowań mapowych. W literaturze podanych jest kilka podejść pozwalających na wykonywanie tego typu prac [4]. W niniejszym artykule podjęto próbę wykorzystania metody *Scale Invariant Feature Transform* (SIFT) [1, 2, 3]. Metoda SIFT należy do metod detekcji stabilnych cech lokalnych i jest wykorzystywana m.in. do automatycznego tworzenia zdjęć panoramicznych, wyszukiwania wzorcowych obiektów na obrazach cyfrowych, nawigacji robotów (również w zastosowaniach przemysłowych). Przewagę tej metody nad innymi sposobami opisu punktów charakterystycznych potwierdzają m.in. badania Mikolajczyka i Schmid [5]. W celu rozwiązania problemu znalezienia odpowiadających sobie punktów na parze zdjęć lotniczych, opracowano algorytm oparty na metodzie SIFT oraz dokonano jego wstępnej implementacji i weryfikacji.

2. Zarys metody SIFT

Metoda SIFT, zaproponowana przez Davida Lowe [1, 2], znajduje stabilne cechy lokalne (punkty charakterystyczne), bazując na stworzeniu piramidy obrazów o coraz mniejszej rozdzielczości. Obrazy tworzące piramidę są obrazami różnicowymi. Są one uzyskiwane w wyniku odjęcia dwóch obrazów powstałych przez przefiltrowanie obrazu początko-

* Katedra Automatyki, Akademia Górniczo-Hutnicza w Krakowie; piotrus@agh.edu.pl
** Zakład Fotogrametrii i Informatyki Teledetekcyjnej, Akademia Górniczo-Hutnicza w Krakowie; smikrut@agh.edu.pl
*** Pracę wykonano w ramach grantu KBN nr 4 T12E 001 27

wego filtrami Gaussa o różnych parametrach σ. Operacja ta jest zatem rodzajem filtracji górnoprzepustowej, a uzyskane obrazy sa obrazami gradientowymi. Piramida jest podzielona na tzw. oktawy. Obrazy składające się na jedną oktawę posiadają tą samą rozdzielczość, a różnią się wielkością zastosowanego filtru Gaussa. W kolejnej oktawie znajdują się obrazy o zmniejszonej rozdzielczości przefiltrowane tymi samymi filtrami. Cechami lokalnymi są minima i maksima wyszukiwane w obrazach piramidy, poddane następującym ograniczeniom:

– jasność punktów musi przekraczać zadany próg,
– punkty nie mogą leżeć wzdłuż linii prostej.

Dla tak wyznaczonych punktów oblicza się wektor cech, będący rodzajem histogramu gradientów liczonych w ich otoczeniu. Takie cechy są w dużej mierze niewrażliwe na natężenie oświetlenia (ze względu na użycie gradientów), obrót (kierunek gradientów zapisuje się względem kierunku gradientu punktu charakterystycznego) i skalę (dzięki stosowaniu piramidy DOG (*Difference of Gaussian*).

Tak uzyskane wektory cech stanowią pewien opis obrazu. Porównując opisy dwóch obrazów, można (np. metodą najbliższego sąsiada) znaleźć wspólne cechy (punkty charakterystyczne) obu obrazów.

3. Algorytm wyszukiwania najlepszych punktów wspólnych pary zdjęć lotniczych

Proponowany algorytm bazuje na fakcie, iż metoda SIFT znajduje wiele wspólnych punktów charakterystycznych pary zdjęć lotniczych. Pozwala to na ich dokładną analizę statystyczną w celu odrzucenia (niewielu) błędnych detekcji metody SIFT oraz umożliwia wybór najlepszych punktów spośród wielu kandydatów (rys. 1).

Rys. 1. Fragmenty porównywanych zdjęć lotniczych wraz z wyznaczonymi metodą SIFT punktami charakterystycznymi

Warto zwrócić uwagę na fakt, iż zdecydowana większość znalezionych punktów leży poza obiektami o znacznej wysokości (np. budynki). Ma to niebagatelne znaczenie dla algorytmu łączącego zdjęcia, gdyż nie wprowadza zniekształceń spowodowanych kątem widzenia tych obiektów.

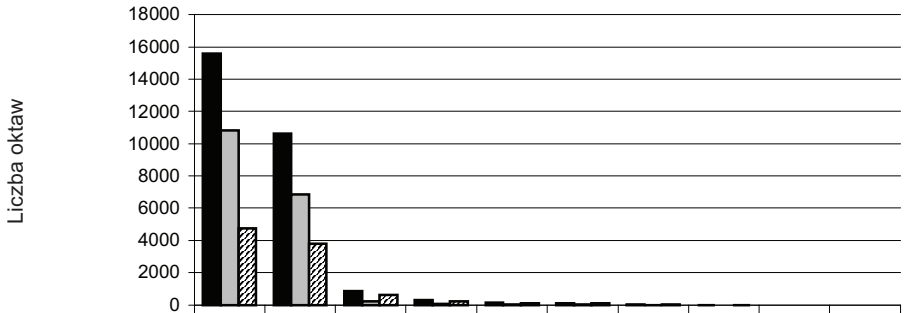
W pierwszej fazie algorytmu następuje ocena przesunięcia kątowego obu zdjęć. Pozwala to wstępnie odrzucić punkty, które zostały błędnie wykryte przez metodę SIFT (wykazujące dużą niezgodność kąta obrotu). Druga faza polega na sprawdzeniu dokładności wyznaczenia współrzędnych proponowanych przez SIFT par punktów za pomocą korelacji ich otoczeń. Warto zauważyć, że korelacja ta może zostać znacznie uproszczona ze względu na znajomość przesunięcia kątowego obu zdjęć wyznaczonego w pierwszej fazie algorytmu. Jeżeli przesunięcie to jest niewielkie (a tak zazwyczaj jest w wypadku zdjęć lotniczych), to wystarczy sprawdzić korelację otoczeń pary punktów. W przeciwnym wypadku należałoby dokonać korekty jednego z otoczeń (maski korelacji), obracając go o wyliczony kąt. W obu przypadkach unika się czasochłonnej korelacji uwzględniającej wszystkie możliwe obroty maski.

Realizacja pierwszej fazy polegała na obliczeniu dyspersji kąta obrotu dla wszystkich par punktów charakterystycznych. Wektor utworzony przez współrzędne punktów z pierwszej wyznaczonej przez metodę SIFT pary był brany jako wektor przesunięcia pomiędzy oboma obrazami. Natomiast współrzędne pierwszego punktu pary stanowiły środek obrotu. Dane te posłużyły do wyliczenia kątów obrotu w pozostałych parach punktów charakterystycznych i policzenia wartości oczekiwanej kąta obrotu oraz odchylenia standardowego. Duża dyspersja świadczyła o błędnym sparowaniu punktów przez metodę SIFT. Taka para została odrzucana, a obliczenia ponawiano dla kolejnej pary punktów. Powyższa metoda pozwala na skuteczne wyeliminowanie błędnych par punktów charakterystycznych.

W implementacji drugiej fazy algorytmu zdecydowano się na prostą korelację otoczeń odpowiadających sobie punktów charakterystycznych obu obrazów. Ze względu na to, że wzajemny obrót testowanych zdjęć nie przekraczał 3 stopni, zdecydowano się nie wprowadzać korekty obrotu do maski korelacji. Użyto maski korelacji o rozmiarze 10×10 pikseli.

4. Wyniki testowania poprawności wyznaczenia współrzędnych punktów charakterystycznych

Dla przetestowania algorytmu opisanego w poprzednim punkcie użyto danych pozyskanych z kilku par zdjęć lotniczych, z których wycięto fragmenty pokrywające się w około 60 procentach. Wycięcie fragmentów było spowodowane chęcią sprawdzenia algorytmu na obrazach testowych (zdjęć) w ich naturalnej rozdzielczości (tzn. 600 dpi, w wypadku użytych zdjęć lotniczych oznaczało to obrazy o rozdzielczościach około 5400×5400 pikseli). Do wyznaczania punktów charakterystycznych użyto oryginalnego oprogramowania autora metody SIFT, które nie posiada możliwości przetwarzania tak dużych obrazów. Z tego względu konieczny był podział zdjęć na fragmenty. Użycie własnej implementacji metody SIFT, pozbawionej tego ograniczenia, jest planowane w trakcie dalszych badań. W sumie uzyskano 102 pary obrazów o rozdzielczości 3980×200 pikseli.



	0	1	2	3	4	5	6	7	8	9
Łącznie	15570	10620	856	309	155	129	50	6	0	0
Do 2 oktawy	10811	6840	219	62	29	27	10	1	0	0
Pow. 2 okt.	4759	3780	637	247	126	102	40	5	0	0

Odległość od maksimum korelacji

Rys. 2. Histogramy odległości pomiędzy punktami wyznaczonymi przez metodę SIFT, a punktami maksimum korelacji

Jak to zostało opisane, poprawność dopasowania wykrytych par punktów sprawdzono za pomocą korelacji. Rysunek 2 przedstawia histogramy różnic pomiędzy współrzędnymi punktów wyznaczonymi przez metodę SIFT a współrzędnymi maksimum korelacji. Histogramy zostały liczone osobno dla oktaw do 2 (wysokie rozdzielczości) i pozostałych (niższe rozdzielczości). Widać tu gwałtowny spadek liczby punktów o odległościach od maksimum korelacji większych niż dwa piksele.

Tabela 1
Sumaryczne zestawienie punktów (par)
o odległościach do 2 pikseli od maksimum korelacji

	Łączna liczba par	Liczba par odl. do 2 pikseli	Procent par odl. do 2 pikseli
Do 2 okt.	17651	348	1,93
Pow. 2 okt.	8539	1157	11,93
Łącznie	26190	1505	5,43

Analizując otrzymane rezultaty (por. tab. 1), można zauważyć, że współrzędne większości punktów znajdowane są przez metodę SIFT poprawnie (odległość jest mniejsza niż dwa piksele). Pozostałe punkty, dla których odległości przekraczają dwa piksele, stanowią 5,75% liczby wszystkich punktów.

Jeżeli jednak uwzględni się podział punktów w zależności od oktawy, dla której zostały wyznaczone, okazuje się, że liczba punktów o odległości przekraczającej dwa piksele stanowi jedynie niespełna 2% liczby wszystkich punktów niskich oktaw. W przypadku

oktaw wyższych (powyżej drugiej), liczba punktów przekraczających graniczną odległość rośnie do 13,55%. Uzyskany wynik sugeruje, że dla wyznaczenia par punktów przydatnych do łączenia dwóch zdjęć lotniczych bardziej odpowiednie są pary wyznaczone dla niskich oktaw (wyznaczone w wysokich rozdzielczościach). W takich przypadkach wyznaczone pary powinny w zupełności wystarczyć do wyliczenia wzajemnego położenia dwóch zdjęć lotniczych.

5. Zakończenie

Dalsze prace skupią się na dokładnej analizie par punktów celem wyznaczenia tych par, które zajmują korzystne położenie (na obrzeżach wspólnego obszaru obu zdjęć) oraz zawierają w swoim otoczeniu unikalne wzorce pozwalające na dokładne wyznaczenie współrzędnych podpikselowych. Duża liczba pozostałych kandydatów powinna także pozwolić na dalszą, dokładniejszą analizę statystyczną, celem wychwycenia par leżących na podobnym poziomie ziemi. Planuje się również własną implementację metody SIFT ze względu na wspomniane ograniczenia oryginalnego programu dotyczące przetwarzania zdjęć dużej rozdzielczości. Własna implementacja pozwoli na skupienie się na parach wyznaczanych w niskich oktawach (co jest bezpośrednim wnioskiem z niniejszego artykułu). Ponadto będzie w niej można zastosować sieci neuronowe, celem przyspieszenia znajdowania par punktów charakterystycznych.

Literatura

[1] Lowe D.G.: *Object recognition from local scale-invariant features*. International Conference on Computer Vision, 1999, 1150–1157

[2] Lowe D.G.: *Distinctive image features from scale-invariant keypoints*. International Journal of Computer Vision, 60, 2, 2004, 91–110

[3] Ke Y., Sukthankar R.: *PCA-SIFT: A More Distinctive Representation for Local Image Descriptors*. Computer Vision and Pattern Recognition, 2004

[4] Kraus K.: *Photogrammetry*. Photogrammetry, Ferd. Dümmlers Verlag, Bonn, 1997

[5] Mikolajczyk K., Schmid C.: *A performance evaluation of local descriptors*. Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003