

ALGORYTM MODELOWANIA 2D ZABUDOWY NA PODSTAWIE DANYCH LOTNICZEGO SKANOWANIA LASEROWEGO Z PROJEKTU ISOK

Krzysztof Sochiera, Andrzej Borkowski

Instytut Geodezji i Geoinformatyki, Uniwersytet Przyrodniczy we Wrocławiu

Streszczenie. Dane lotniczego skanowania laserowego (ALS) pozyskiwane są najczęściej na potrzeby budowy numerycznych modeli wysokościowych. W Polsce dane takie pozyskane zostały dla obszaru niemal całego kraju w ramach projektu ISOK, związanego z osłoną przed zagrożeniami naturalnymi. Dane te zostały wykorzystane w niniejszej pracy do modelowania obrysów budynków. W tym celu zaproponowano algorytm będący kombinacją algorytmu α -shape do detekcji konturów budynków oraz iteracyjnego ogólnego modelu wyrównawczego do aproksymacji rzutów ortogonalnych ścian budynków. Identyfikację punktów reprezentujących obrysy budynków wykonano na podstawie chmury punktów, z której odrzucono punkty powyżej zadanej wysokości progowej. Identyfikacja obrysów budynków jako otoczki pustych powierzchni reprezentujących budynki dokładniej przybliżyła rzeczywiste położenie przyziemi ścian budynków. Do weryfikacji algorytmu wykorzystano chmurę punktów o gęstości 12 pkt/m² reprezentującą miejski obszar zurbanizowany o zabudowie regularnej. Wyniki modelowania 2D budynków porównano z ich reprezentacją w bazie Ewidencji Gruntów i Budynków oraz obliczono odchyłki liniowe odpowiadających sobie narożników. Otrzymano średnią wartość odchyłki liniowej na poziomie 0,56 m. Wartość ta jest zgodna z nominalną dokładnością sytuacyjną danych ALS projektu ISOK. Błąd średniokwadratowy policzony na podstawie odchyłek liniowych wynosi 0,64 m. Otrzymane wyniki modelowania spełniają wymagania dokładnościowe Bazy Danych Obiektów Topograficznych 1:10000 (BDOT10k) i mogą być wykorzystane do jej weryfikacji, aktualizacji bądź zasilania.

Słowa kluczowe: lotniczy skaning laserowy, alpha shape, ogólny model wyrównawczy, ISOK, modelowanie 2D

WSTĘP

Technologia lotniczego skanowania laserowego (ALS) wykorzystywana jest w pierwszej kolejności do pozyskiwania informacji o ukształtowaniu topograficznej powierzchni terenu, zwłaszcza na obszarach trudno dostępnych. Potencjał informacyjny zawarty w danych ALS jest jednak zdecydowanie większy. Wraz ze wzrostem dokładności i gęstości skanowania dane ALS znajdują zastosowanie do modelowania innych elementów geometrycznych przestrzeni oraz zasilania baz danych. Dla przykładu, w pracy Mendel i Borkowskiego [2013] wykazano, że dane ALS mogą być wykorzystane jako wartościowe źródło danych do modelowania obiektów topograficznych i zasilania baz danych BDOT10K. Poza generowaniem topograficznych modeli wysokościowych głównym zastosowaniem danych ALS pozostaje jednak wciąż ich wykorzystanie do budowy trójwymiarowych numerycznych modeli zabudowy. Przegląd metod i możliwości wykorzystania danych ALS można znaleźć w pracach innych autorów [Vosselman i Maas 2010, Jarząbek-Rychard i Borkowski 2016]. W pracy Borkowskiego i Józkowa [2012] pokazano, że na podstawie danych skanowania laserowego można zbudować model 3D budynku z dokładnością na poziomie dokładności pomiarów geodezyjnych. Analizę przydatności lotniczego skaningu laserowego do opracowania modeli 3D budynków zgodnie ze standardem dyrektywy INSPIRE przeprowadzono w pracy Cisło-Lesickiej i in. [2014].

Innym obszarem wykorzystania danych ALS jest modelowanie obrysów budynków. Problematyka ta jest podejmowana w licznych publikacjach. W pracy Jarząbek-Rychard [2012] przedstawiono algorytm modelowania obrysów budynków w gęstej zabudowie na podstawie danych ALS wspomaganych informacją o punktach adresowych.

W pracy Tomljenovica i in. [2015] zebrano artykuły naukowe z lat 1998–2014 o tej tematyce oraz dokonano ich analizy. Autorzy wskazują na różnorodność metod użytych do wyznaczenia obrysu budynków na podstawie danych ALS. Część z nich wymaga dodatkowego źródła danych, jakimi są panchromatyczne zdjęcia satelitarne [Sohn i Dowman 2007] lub utworzony na podstawie danych ALS numeryczny model terenu [Priestnall i in. 2000]. Inna metoda wykorzystuje ulepszony model aktywnych konturów, który definiowany jest jako spline minimalizujący energię [Kabolizade i in. 2010]. Bazuje on na chmurze ALS pokolorowanej na podstawie zdjęć lotniczych. Do jego wykorzystania konieczna jest wiedza *a priori* dotycząca radiometrycznych oraz geometrycznych własności budynków na zdjęciach. Inne prace przedstawiają algorytmy bazujące na obróbce chmury punktów. Istnieją wśród nich opracowania, które koncentrują się na uzyskaniu obrysu budynków do tworzenia modeli 3D budynków zgodnie ze standardem OGC CityGML [OGC 2012].

Sampath i Shan [2004, 2007] przedstawili czteroetapowy algorytm wyznaczenia krawędzi budynków, składający się z separacji punktów tworzących budynki od pozostałych punktów ALS, identyfikacji poszczególnych budynków z wykorzystaniem algorytmu rosnących regionów oraz wyznaczenia punktów opisujących krawędzie budynków jako otoczki wypukłej. W ostatnim etapie punkty reprezentujące krawędzie budynku aproksymowane są liniami prostymi z wykorzystaniem metody najmniejszych kwadratów. W literaturze opisano także przebieg procesu w trzech etapach [Wei 2008] z użyciem algorytmu α -shape, który identyfikuje punkty opisujące kształt budynku. Następnie stosowany jest algorytm Improved Pipe Algorithm [Zhao 2002] do generalizacji obrysu oraz Clustering and Adjustemend Algorithm, który pozwala uzyskać regularne krawędzie budynku.

Pewną modyfikacją jest algorytm [Lee i in. 2011], który bazuje na otoczce wypukłej zmodyfikowanej, tak aby wyznaczała również obrys wewnątrz przyziemia budynków. Do generalizacji kształtu wykorzystano algorytm Sleeve-fitting [Zhao, Saalfeld 1997], który jest ekwiwalentny do algorytmu Douglasa–Peuckera [Douglas, Peucker 1973]. W pracy Yang i in. [2013] zaproponowali metodę bazującą na algorytmie Marked Point Process [Ripley i Kelly 1977], który wykorzystuje model energii Gibbsa.

Innym podejściem jest wyznaczenie budynków z danych ALS po uprzedniej klasyfikacji punktów pomiarowych. W publikacji Niemeyer i in. [2014] wykonano klasyfikację nadzorowaną. Polegała ona na integracji klasyfikatora Random Forest wraz z szablonem Conditional Random Field (CRF) [Kumar i Hebert 2006]. Chmura punktów była reprezentowana przez model graficzny, który był zbiorem relacji pomiędzy obiektami klas i danymi w kroku treningowym. W pracy porównano kilka wersji klasyfikatorów opartych na CRF. Wykazano, że są one dużo szybsze i dokładniejsze niż algorytmy działające na modelu liniowym. W wyniku tego uzyskano podzieloną na 7 klas chmurę punktów oraz dużą liczbę obiektów. Pośród tych obiektów znalazły się budynki. Do prawidłowego ich zidentyfikowania posłużono się prawdopodobieństwem klas z szablonu CRF, które następnie zostały parametrami Markow Random Field. Na koniec otrzymano obraz 2D z naniesionymi klasami obiektów, w tym m.in. budynkami. Uzyskano wysokie wartości kompletności i poprawności danych. W przypadku kryteriów powierzchniowego wartości wyniosły 90%, a dla kryterium obiektowego około 84%.

Autorzy przeglądu literatury [Tomljenovic 2015] zauważyli, że opisane publikacje charakteryzują się dużą dokładnością wyznaczenia budynków (powyżej 85%). Jednocześnie podkreślono brak standardu oceny dokładności i brak szeroko używanych metryk do porównania wyników różnych metod. Działanie algorytmów jest powszechnie testowane na danych generowanych przez samych autorów algorytmów, poprzez digitalizację mapy lub wycięcie rastra numerycznego modelu pokrycia terenu. Dlatego postulowano, by do testowania danych używać tylko zaufanych źródeł, takich jak urzędowe dane katastralne czy mapy Open Street Map. Należy zauważyć, że jakość tego ostatniego źródła bywa dyskusyjna i należy podchodzić do niego z ograniczonym zaufaniem.

Celem niniejszej pracy jest opracowanie i weryfikacja autorskiego algorytmu modelowania konturów 2D budynków na podstawie niesklasyfikowanych danych ALS Informatycznego Systemu Osłony Kraju (ISOK). Proponowany algorytm bazuje na modelu α -shape i zmodyfikowanym modelu regresji ortogonalnej zaproponowanym w pracy [Jazaeri i in. 2014]. Modelowanie wykonano na podstawie danych ALS z obszarów zurbanizowanych, w standardzie 12 pkt/m². Do weryfikacji algorytmu i oceny bezwzględnej jakości modelowania posłużyły dane państwowego rejestru Ewidencji Gruntów i Budynków (EGiB). Dane te uzyskane zostały w wyniku pomiarów klasycznymi metodami geodezyjnymi z zachowaniem odpowiednich standardów technicznych.

METODOLOGIA

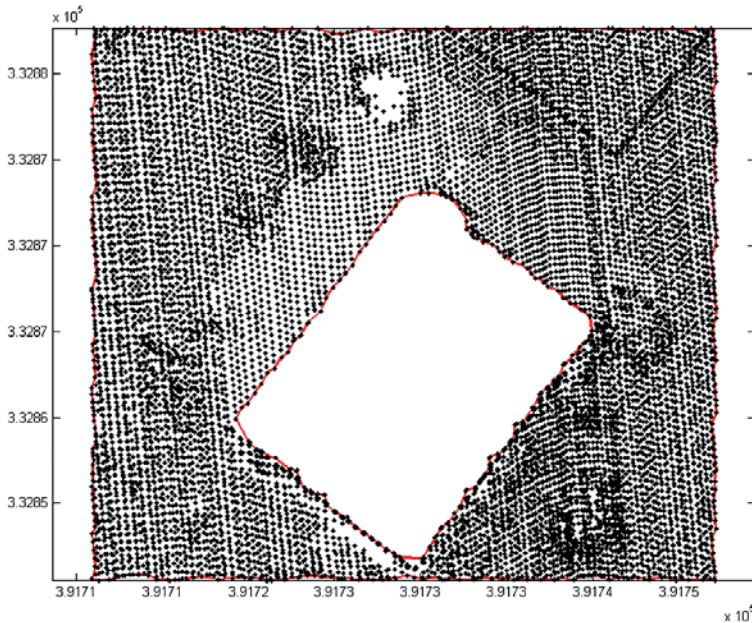
Zadanie automatycznej identyfikacji budynków w chmurze ALS i wyznaczenia ich obrysów podzielono na trzy etapy: identyfikacja punktów tworzących obrys budynku, segmentacja krawędzi oraz regulacja obrysów, przy założeniu prostopadłości ścian. Pierwszym krokiem jest usunięcie punktów reprezentujących dachy budynków. Odbywa się to poprzez ustalenie progu będącego sumą średniej arytmetycznej wysokości punktów oraz arbitralnego parametru w . Inicjalna wartość tego parametru (w) wynosi 8 m, ale operator algorytmu może dostosować jego wielkość do specyfiki zbioru danych. Algorytm α -shape posłużył do uzyskania punktów tworzących obrys budynków jako punktów reprezentujących obrys pustych obszarów w zbiorze danych ALS. Segmentację krawędzi przeprowadzono w dwóch etapach. W pierwszym etapie punkty zostały posortowane, tak by element $n+1$ zbioru był najbliższym do elementu n . W drugim etapie przeprowadzono właściwą segmentację na podstawie współczynnika kierunkowego prostej pomiędzy kolejnymi punktami zbioru. Ostatecznie, z wykorzystaniem algorytmu IWTLs (patrz pkt 2) przeprowadzono iteracyjny proces regulacji obrysów, wpasowując proste prostopadłe. Parametry opisujące przebieg prostej wyznaczone przez ten algorytm zostały dostosowane do warunku prostopadłości krawędzi. Liczba budynków i krawędzi może być dowolna. Danymi wchodzącymi do programu są punkty należące do krawędzi o znanych współrzędnych x i y .

Algorytm Alpha shape

Algorytm α -shape został wykorzystany do identyfikacji punktów składających się na krawędzie budynków. Definicje wykorzystanego algorytmu przytaczamy za pracą Mendeli i Borkowskiego [2013]. Niech będzie dany skończony zbiór punktów $S \in R^2$. Dowolny punkt $p \in S$ jest punktem skrajnym w sensie α , jeśli istnieje pewien pusty dysk D o promieniu $1/\alpha$, niezawierający żadnego punktu z S z wyjątkiem punktu p na obrzeżu tego dysku. Dwa punkty skrajne $p, q \in S$ nazywane są punktami sąsiednimi w sensie α , jeśli leżą na obrzeżu tego samego dysku D . α -shape jest poligonem składającym się z prostoliniowych połączeń punktów sąsiednich.

Zaletami algorytmu są jego skuteczność i łatwość w użyciu. Jedynym parametrem sterującym algorytmu jest bezwymiarowy współczynnik α , który determinuje szczegółowość wynikowego poligonu. W niniejszej pracy wartość α została ustalona eksperymentalnie na zbiorze danych. W testach kierowano się kryterium szczegółowości, czyli uzyskania możliwie największej liczby punktów, a także dobrym i wyraźnym odwzorowaniem kątów prostych i uniknięciem nadmiernych szumów (rys. 1). Ostatecznie ustalono wartość 0,8.

Na potrzeby niniejszej pracy wykorzystano implementację algorytmu w środowisku Matlab [Mathworks 2013], która realizuje α -shape na podstawie triangulacji Delaunaya.



Rys. 1. Przykładowe wyznaczenie obrysu przez algorytm α -shape
 Fig. 1. Building boundary detected by α -shape

Algorytm ITWLS

Algorytm „Iterative algorithm for weighted total least squares” (ITWLS) [Jazaeri i in. 2014], zaimplementowany w ramach niniejszej pracy w środowisku Matlab, posłużył do wpasowania prostych w zbiory punktów reprezentujących rzuty poszczególnych ścian budynków. Idea tego podejścia bazuje na rozszerzonym modelu regresji ortogonalnej, w którym zarówno wektor obserwacji, jak i macierz współczynników przy niewiadomych są obciążone błędami losowymi. Macierz współczynników przy niewiadomych w metodzie ITWLS wykazuje cechy modelu probabilistycznego. Macierze wariancyjno-kowariancyjne dla błędów obserwacji (korekt) e oraz dla współczynników przy niewiadomych E_A tworzone są przy założeniu braku korelacji między zmiennymi (współzrzednymi) X i Y. Jest to metoda iteracyjna, w której wpasowanie prostych realizowane jest według poniższego algorytmu [Jazaeri i in. 2014]:

1. Obliczenie parametrów początkowych (x_0) metodą najmniejszych kwadratów.
2. Nadanie wag obserwacjom. W tym przypadku współzrzedne X i Y punktów reprezentujących obrysy budynków potraktowano jako jednakowo dokładne.

Dla współzrzednej X: $W_x = \text{diag}[w_{x1} \quad w_{x2} \quad \dots \quad w_{xn}] = \text{diag} [1 \quad 1 \quad \dots \quad 1]$

Dla współzrzednej Y: $W_y = \text{diag}[w_{y1} \quad w_{y2} \quad \dots \quad w_{yn}] = \text{diag} [1 \quad 1 \quad \dots \quad 1]$
 (diag oznacza macierz diagonalną).

3. Wyznaczenie macierzy kowariancji wektora błędów oraz macierzy A.

Z założenia wektor obserwacji i macierz współczynników przy niewiadomych \mathbf{A} są obarczone błędami:

$$\mathbf{y} = E(\mathbf{y}) = \mathbf{y} - \mathbf{e} \quad (1)$$

$$\mathbf{A} = E(\mathbf{A}) = \mathbf{A} - \mathbf{E}_A \quad (2)$$

gdzie E jest symbolem wartości oczekiwanej.

Macierz wariancyjno-kowariancyjną wektora błędów \mathbf{e} oznaczono przez \mathbf{Q}_e , a macierz wariancyjno-kowariancyjną rozwiniętej w wektor macierzy \mathbf{A} oznaczono przez \mathbf{Q}_E :

$$\mathbf{Q}_e = \text{diag}(w_{y_1}^{-1}, w_{y_2}^{-1}, \dots, w_{y_n}^{-1}) \quad (3)$$

$$\mathbf{Q}_E = \text{diag}(w_{x_1}^{-1}, w_{x_2}^{-1}, \dots, w_{x_n}^{-1}, 0, 0, \dots, 0) \quad (4)$$

Macierz \mathbf{Q}_e ma wymiar $n \times n$, a macierz \mathbf{Q}_E ma wymiar $nm \times nm$.

4. Obliczenie iloczynu Kroneckera parametrów początkowych oraz macierzy jednostkowej \mathbf{I}_n . Wymiar macierzy jednostkowej to $n \times n$.

$$\mathbf{x} = \mathbf{x}_0 \otimes \mathbf{I}_n \quad (5)$$

5. Rozpoczęcie obliczeń iteracyjnych. Użyta funkcja $\text{vec}()$ oznacza konwersję z macierzy do wektora. Wektor składa się z kolumn macierzy ułożonych jedna pod drugą. Funkcja $\text{reshape}()$ odwraca ten proces.

Dla $i \in N$:

$$\mathbf{Q}_1^i = \mathbf{Q}_e \mathbf{x}_{i-1}^T \mathbf{Q}_E \mathbf{x}_{i-1} \quad (6)$$

$$\text{vec}(\mathbf{E}_A^i) = -\mathbf{Q}_E \mathbf{x} (\mathbf{Q}_1^i)^{-1} (\mathbf{y} - \mathbf{A} \mathbf{x}_{i-1}) \quad (7)$$

$$\mathbf{E}_A^i = \text{reshape}(\mathbf{E}_A^i) \quad (8)$$

$$\mathbf{A}_i = \mathbf{A} - \mathbf{E}_A^i \quad (9)$$

Obliczanie nowych parametrów:

$$\mathbf{x}_i = (\mathbf{A}_i^T (\mathbf{Q}_1^i)^{-1} \mathbf{A}_i^T)^{-1} \mathbf{A}_i^T (\mathbf{Q}_1^i)^{-1} (\mathbf{y} - \mathbf{E}_A^i \mathbf{x}_{i-1}) \quad (10)$$

$$\lambda_i = (\mathbf{Q}_1^i)^{-1} (\mathbf{y} - \mathbf{A} \mathbf{x}_i) \quad (11)$$

Koniec działania pętli, gdy:

$$\mathbf{x}_i - \mathbf{x}_{i-1} < \varepsilon \quad (\text{w tym przypadku } = 10^{-5}) \quad (12)$$

6. Obliczenie błędu średniego σ_0^2 oraz macierz kowariancji \mathbf{C}_x :

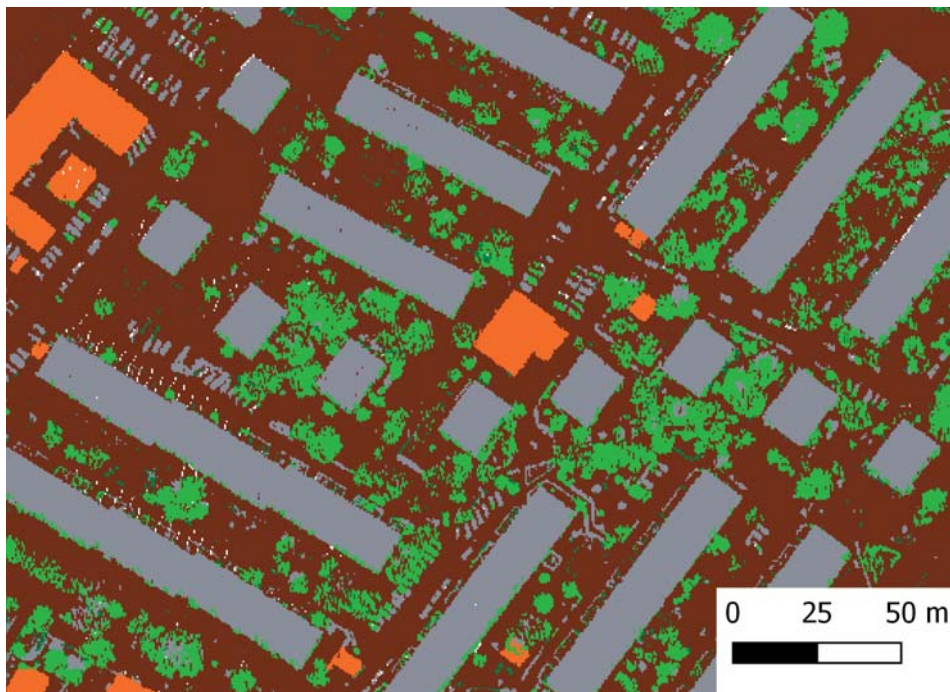
$$\sigma_0^2 = \left(\frac{\lambda^T \mathbf{y} - \mathbf{A}\mathbf{X}}{n - m} \right) \quad (13)$$

$$\mathbf{C}_x = \sigma_0^2 [\mathbf{A}_i^T \mathbf{Q}_1^{-1} \mathbf{A}_i]^{-1} \quad (14)$$

DANE TESTOWE

Dane ALS wykorzystane w pracy zostały pozyskane w trakcie projektu: Informatyczny System Osłony Kraju (ISOK). Chmura punktów charakteryzowała się gęstością nominalną 12 punktów/m². Dokładność wysokościowa nie powinna być gorsza niż 0,15 m, a planarna nie gorsza niż 0,5 m.

Na podstawie tych danych przeprowadzono testy dla 21 budynków o różnych wymiarach znajdujących się na osiedlu mieszkalnym w mieście Brzeg w woj. opolskim. Na obszarze testowym przeważają tereny płaskie z wyróżniającymi się wysokimi zabudowaniami oraz roślinnością wysoką. Obiekt testowy przedstawiono na rysunku 2. Kolorem szarym oznaczono zabudowę wysoką, pomarańczowym – zabudowę niską, zielonym – roślinność, a szarobrązowym – teren niezagospodarowany. Na potrzeby wizualizacji wykorzystano klasyfikację wykonaną w ramach ISOK, algorytm opracowany w niniejszej pracy działa na niesklasyfikowanej chmurze punktów.



Rys. 2. Obiekt testowy – fragment zabudowy miasta Brzeg, woj. opolskie
Fig. 2. Test site – a part of the city Brzeg

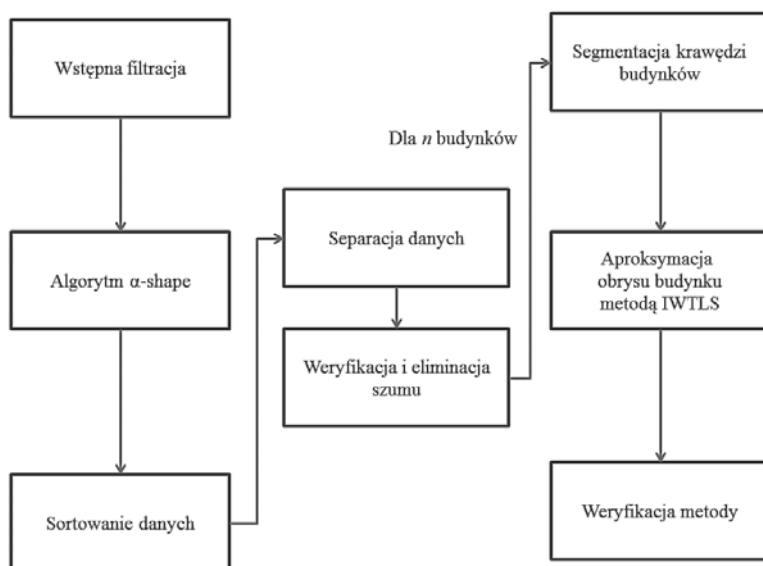
Do weryfikacji działania algorytmu posłużyły dane państwowego rejestru Ewidencji Gruntów i Budynków (EGiB). Dane te uzyskane zostały w wyniku pomiarów klasycznymi metodami geodezyjnymi z zachowaniem odpowiednich standardów technicznych i prawnych. Dane te zgromadzone zostały w pliku w formacie SHP, który pozwala na zapis geometrii obiektów, a także ich tekstowych atrybutów. Następnie poddano je obróbce w programie QGIS, aby uzyskać punkty załamania obrysów budynków i porównać ich położenie z punktami wyznaczonymi na podstawie danych ALS z wykorzystaniem proponowanego algorytmu.

MODELOWANIE OBRYSU BUDYNKU

Algorytm modelowanie obrysu budynku przedstawiono schematycznie na rysunku 3. Poniżej omówione zostały poszczególne etapy realizacji algorytmu.

Wstępna filtracja i identyfikacja punktów obrysu

Celem wstępnej filtracji jest odrzucenie punktów powyżej założonej wysokości (usunięcie dachów budynków). W tym celu zakłada się niewielkie nachylenie terenu (teren płaski). W przypadku obszarów, gdzie założenie to jest nieskuteczne, należy obliczyć znormalizowany numeryczny model pokrycia terenu. Następnie oblicza się średnią wysokość punktów na rozpatrywanym obszarze i dodaje do niej ustaloną wartość parametru w . Parametr ten dobierany jest empirycznie. Wszystkie punkty powyżej tej granicy zostają usunięte. Efektem jest zbiór punktów zawierający puste miejsca odpowiadające lokalizacji budynków. W kolejnym kroku, algorytm α -shape wyodrębnia punkty reprezentujące obrysy obiektów. Wyznaczone punkty zapisane są do pliku tekstowego.



Rys. 3. Schemat działania algorytmu

Fig 3. Flowchart of the algorithm

Sortowanie danych

Po wczytaniu danych z pliku tekstowego należy zdefiniować takie parametry jak: minimalna długość pojedynczej krawędzi, minimalna liczba punktów tworzących krawędź oraz maksymalna odległość punktu przecięcia krawędzi od punktu początkowego lub końcowego krawędzi.

Wczytane dane najczęściej są ułożone w przypadkowej kolejności bądź posortowane według dowolnego kryterium. Aby rozdzielić dane na poszczególne budynki, punkty należy odpowiednio posortować.

Funkcja rozpoczyna swoje działanie od wyboru pierwszego punktu ze zbioru. Staje się on punktem głównym P. Automatycznie zostaje usunięty ze zbioru i umieszczony w macierzy N. Pozostałe punkty przenoszone są do macierzy M. Następuje obliczenie odległości od punktu P do wszystkich punktów zgromadzonych w macierzy M. Funkcja wyznacza punkt najbliższy do punktu P. Staje się on nowym punktem P i zostaje on jednocześnie skopiowany do macierzy N i usunięty z macierzy M. Funkcja zaczyna obliczenia od nowa i wykonuje kolejne kroki. Jej działanie kończy się, gdy w zbiorze M pozostaje 1 element. Uporządkowana macierz N jest zwracana do algorytmu.

Separacja danych na poszczególne budynki

Punktem wyjścia do wydzielenia poszczególnych budynków jest posortowany zbiór punktów. Separacja poszczególnych budynków odbywa się na podstawie odległości pomiędzy kolejnymi punktami w zbiorze. Jeśli odległość nie przekracza d_{max} (parametr ustalony empirycznie, przez użytkownika), to oznacza, że punkty tworzą ten sam budynek. Gdy warunek zostanie przekroczony, zbiór punktów tworzących aktualnie procedowany budynek zostaje zamknięty, a kolejne punkty zapisywane są do nowego zbioru. Ta procedura kontynuowana jest aż do wyczerpania wszystkich punktów w zbiorze.

Metoda jest weryfikowana na 2 sposoby: po pierwsze liczebność zbioru punktów tworzących pojedynczy budynek musi wynosić przynajmniej 3, a po drugie najdłuższa krawędź budynku nie może przekraczać ustalonej wartości d_{max} . W przeciwnym razie wszystkie punkty w danym zbiorze traktowane są jako szumy. Na koniec budynki są ponownie zliczane i otrzymują nowe nazwy, żeby uniknąć braków w numeracji macierzy.

Segmentacja na krawędzie budynku

W publikacji Sampatha i Shana [2007] zauważono, że segmentacja może odbywać się z wykorzystaniem właściwości współczynnika kierunkowego prostej. Dzięki temu można rozdzielić zbiór danych należących do jednego budynku na poszczególne krawędzie. Według algorytmu zaimplementowanego w niniejszej pracy odbywa się to dwuetapowo. W pierwszym etapie utworzone zostają dwie macierze, które zawierają punkty, w których współczynnik ma dodatnią wartość oraz w których współczynnik jest ujemny.

Na podstawie wykonywanych obliczeń stwierdzono, że powtórne rozdzielenie każdej z nowych macierzy na dwie kolejne z zachowaniem tej samej reguły pomaga w filtracji szumu. Dlatego w efekcie powstają jeszcze 4 macierze: 1 z elementami, w których wartość była dwukrotnie dodatnia, 1 z elementami, w których po obu działaniach były ujemne oraz 2 mieszane, które zostały uznane za szum. Ich liczebność wynosi do 7% całego zbioru [Sochiera 2015]. Po wykonaniu tego zadania robocze nazwy macierzy zostają zamienione na nowe z konkretnym opisem zawartości (po przejściu 2 etapów obliczeń) i numerem budynku.

Wpasowanie prostej z użyciem metody IWTLS

Funkcja w pierwszym etapie wyznacza iteracyjnie współczynnik kierunkowy prostej w przypadku poszczególnych krawędzi z użyciem metody IWTLS. Następnie wyznacza średnią ważoną tego współczynnika. Należy podkreślić nadany wcześniej warunek o prostopadłości krawędzi, który dla przecinających się prostych przyjmuje postać $a_1 a_2 = -1$, gdzie a_1 i a_2 to współczynniki kierunkowe poszczególnych prostych. Uzyskano w ten sposób parametry pozwalające na wpasowanie prostych i prostopadłych względem siebie krawędzi. Obliczone parametry posłużyły do wyznaczenia punktów załamania obrysów. Włączenie kolejnych punktów załamania obrysu pozwoliło na najbardziej dokładną aproksymację przebiegu krawędzi budynków.

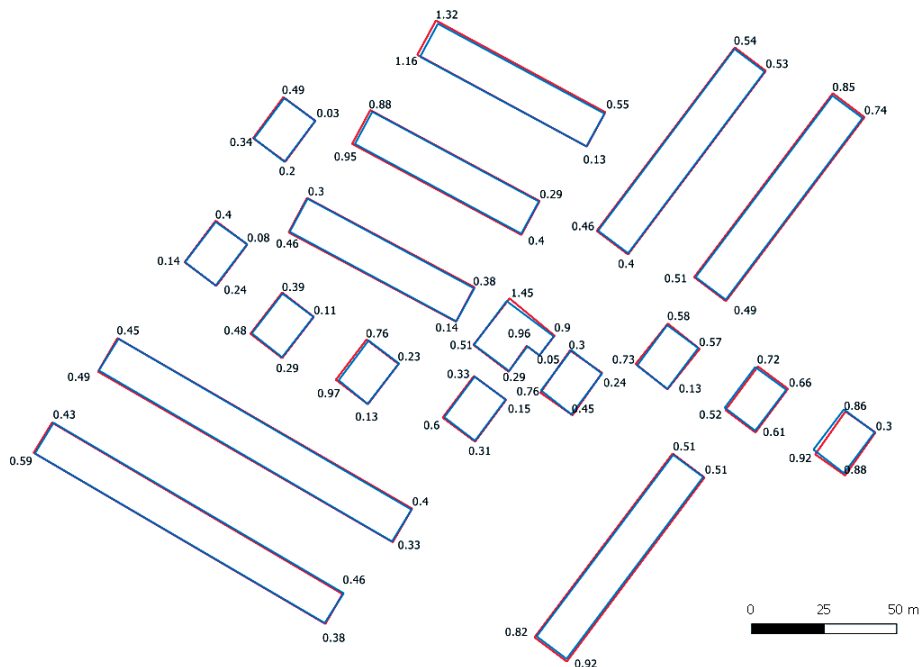
OCENA DOKŁADNOŚCI MODELOWANIA

Opracowany algorytm został poddany weryfikacji poprzez wyznaczanie obrysów budynków pochodzących z danych testowych. Ogółem otrzymano 86 punktów załamania obrysu budynków, z których algorytm złożył poprawnie 17 na 21 budynków istniejących w zbiorze danych (rys. 4). Tym samym skuteczność tego algorytmu wyniosła 81%. Trzy budynki, które nie zostały wyznaczone przez algorytm, znajdowały się na obrzeżach terenu poddanego testom, w wyniku tego obrys budynków był niekompletny. Zauważono, że wpływ na jakość wyników ma gęstość chmury punktów, a w konsekwencji – punktów składających się na krawędzie budynków [Sochiera 2015]. Większa liczba takich punktów pozwala na iteracyjną zmianę wagowania w przypadku punktów odstających ze zbioru danych i ustalenie poprawnej granicy budynku.

Dane z EGiB przyjęto jako dane referencyjne (wartości prawdziwe i bezbłędne). Na ich podstawie wyznaczono punkty załamania krawędzi budynków, które porównano z punktami załamania obrysów budynków otrzymanymi z algorytmu. Miara dokładności były odległości (rezydua) pomiędzy odpowiadającymi sobie punktami. Mieszają się one w przedziale od 0,04 do 1,45 m. Średnia wartości odchyłek liniowych wyniosła 0,56 m z odchyleniem standardowym 0,33 m. 90% obserwacji nie przekracza różnicy 1 m w stosunku do danych z EGiB. Obliczony na podstawie odchyłek liniowych błąd średniokwadratowy (RMSE) kształtuje się na poziomie 0,64 m.

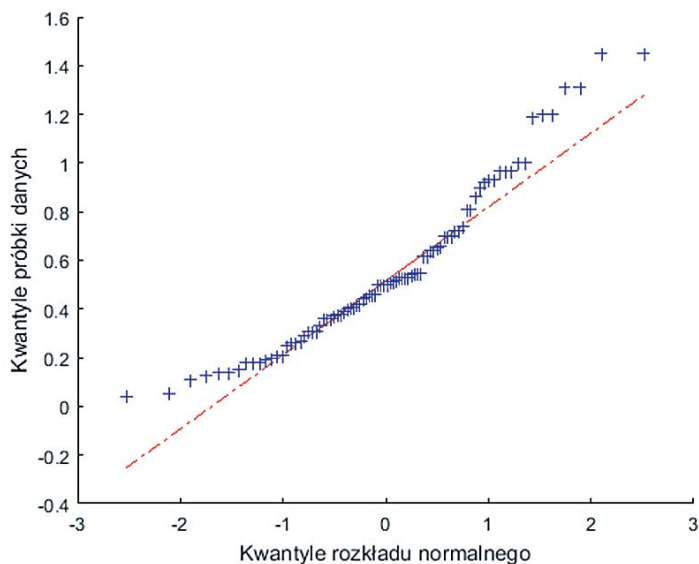
Na rysunku 5 przedstawiono wykres kwanty-kwantyl w przypadku obliczonych odchyłek. Jest to graficzna reprezentacja danych w porównaniu z teoretycznymi kwantylami o rozkładzie normalnym. Wykres przedstawia rozmieszczenie obserwacji względem prostej reprezentującej oczekiwane wartości wynikające z rozkładu normalnego. Przedstawione na wykresie odchyłki są zbliżone do wykreślonej prostej, w części środkowej. Na końcach przedziałów widoczne są znaczne odchylenia od rozkładu normalnego, co świadczy o występowaniu odchyłek zarówno bardzo małych, jak i dużych, przy czym te ostatnie (7 wartości) świadczą o występowaniu błędów grubych.

Rezydua zostały poddane testowi Shapiro–Wilka na poziomie istotności $\alpha = 0,05$, który miał na celu weryfikację tezy H_0 o rozkładzie normalnym tych rezyduów [Pawłuszek i in. 2014]. Test nie został spełniony. Należy mieć na uwadze, że test Shapiro–Wilka jest czuły na pojedyncze odstające obserwacje. Następnie przeprowadzono symulację testu po usunięciu ze zbioru odstających rezultatów, w tym przypadku wszystkich równych lub większych niż 1 m, test został spełniony. Liczba usuniętych obserwacji wyniosła 8.



Rys. 4. Obrisy wyznaczone na podstawie danych ALS (kolor niebieski) oraz prawidłowe, reprezentowane w EGB (kolor czerwony) wraz z podanymi wartościami odchyłek

Fig. 4. Building outlines detected in ALS point cloud (blue lines) and reference building outlines represented in cadaster (red lines) with linear deviation values between them



Rys. 5. Wykres kwantyl-kwantyl dla otrzymanych wyników

Fig. 5. Q-Q plot of linear deviations

Uzyskane dokładności mieszczą się w granicach dopuszczalnych dokładności określonych dla BDOT10k [Rozporządzenie 2011]. Należy zauważyć, że dokładności te osiągnięto w przypadku standardowej procedury opracowania danych ALS, jaka była stosowana podczas realizacji projektu ISOK. Stosując odpowiednie, zaawansowane procedury, związane z pozyskaniem punktów kontrolnych i wyrównaniem pasów skanowania można osiągnąć dokładność sytuacyjną danych ALS na poziomie kilku centymetrów [Csanyi i Toth 2007, Li i in. 2012, Pyka i in. 2012].

PODSUMOWANIE

Dane lotniczego skanowania laserowego pozyskane w trakcie projektu ISOK obejmują obszar całego kraju. Dane te pozyskane zostały na potrzeby budowy numerycznych modeli wysokościowych wykorzystywanych do modelowania hydrologicznego, jednak potencjał informacyjny zawarty w tych danych jest znacznie większy.

W pracy zaimplementowano i zweryfikowano algorytm modelowania obrysów budynków na podstawie danych z projektu ISOK. Algorytm ten wykorzystuje metodę α -shape oraz iteracyjny ogólny model wyrównawczy do aproksymacji obrysów przyziemi budynków. Identyfikację punktów reprezentujących obrysy budynków wykonano z wykorzystaniem α -shape na chmurze punktów, z której usunięto punkty powyżej zadanej wysokości. W literaturze zdecydowanie częściej stosowane są algorytmy wykorzystujące sklasyfikowaną chmurę, z której ekstrahowane są budynki [Alberts 2016]. Wyznaczone w ten sposób obrysy budynków są tożsame z obrysami dachów. Zastosowana w niniejszej pracy identyfikacja budynków jako otoczki wewnętrznej powinna lepiej przybliżać rzeczywiste położenie i kształt przyziemi ścian budynków, zwłaszcza w przypadku budynków wysokich, kiedy występują liczne odbicia impulsów laserowych od ścian budynków i jednocześnie budynków posiadających znaczne okapy. Aspekt ten będzie przedmiotem dalszych badań autorów.

Wyniki modelowania obrysów budynków porównano z danymi Ewidencji Gruntów i Budynków. Otrzymana średnia wartość odchyłki liniowej pomiędzy narożnikami budynków z EGB a modelowanymi z danych ALS na poziomie 0,56 m jest zgodna z nominalną dokładnością planarną danych ALS projektu ISOK. Jednocześnie błąd średniokwadratowy obliczony na podstawie odchyłek liniowych otrzymano na poziomie 0,64 m. Taka dokładność geometryczna pozwala na aktualizację Bazy Danych Obiektów Topograficznych (BDOT10k) na podstawie danych ALS. Oczywiście, dane ALS są aktualne w momencie wykonania skanowania i wtedy są najbardziej przydatne na potrzeby modelowania obiektów 2D. Wydaje się jednak, że nawet po upływie pewnego czasu, jak w przypadku danych ISOK, dane ALS mogą być wciąż przydatne na potrzeby weryfikacji obiektów, np. pod kątem identyfikacji możliwych błędów (grubych) w bazie obiektów topograficznych.

W zaimplementowanej aktualnie wersji algorytm identyfikuje obrysy przyziemi budynków, pod warunkiem że poszczególne ściany są do siebie prostopadłe. W kolejnej implementacji algorytm powinien modelować budynki o dowolnym kształcie, z wykorzystaniem metody IWTLs w wariantcie odpornym na błędy grube. Powinno to poprawić skuteczność modelowania, w przypadku gdy w sąsiedztwie budynków występują drzewa.

PIŚMIENNICTWO

- Albers B., Kada M., Wichmann A., 2016. Automatic extraction and regularization of building outlines from airborne LiDAR point clouds. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLI-B3, 555–560.
- Borkowski A., Józków G., 2012. Accuracy assessment of building models created from laser scanning data. *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXIX-B3, 253–258.
- Cisło-Lesicka U., Borowiec N., Marmol U., Pyka P., 2014. Analiza przydatności lotniczego skaningu laserowego do opracowania modelu budynku 3D zgodnie ze specyfikacją INSPIRE. *Archiwum Fotogrametrii, Kartografii i Teledetekcji*, vol. 26, 39–52
- Csanyi N., Toth C., 2007. Improvement of lidar data accuracy using lidar-specific ground targets. *Photogrammetric Engineering & Remote Sensing*, 73 (4), 385–396.
- Douglas D.H., Peucker T.K., 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 10 (2), 112–122.
- Jarząbek-Rychard M., 2012. Reconstruction of building outlines in dense urban areas based on LiDAR data and address points. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX, B3, 121–126.
- Jarząbek-Rychard M., Borkowski A., 2016. 3D building reconstruction from ALS data using unambiguous decomposition into elementary structures. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 118, 1–12.
- Jazaeri S., Amiri-Simkooei A.R., Sharifi M.A., 2014. Iterative algorithm for weighted total least squares adjustment. *Survey Review* 46, 334, 19–27.
- Kabolizade M., Ebadi H., Ahmadi S., 2010., An improved snake model for automatic extraction of buildings from urban aerial images and LiDAR data. *Computers, Environment and Urban Systems*, 34 (5), 435–441.
- Kumar S., Hebert M., 2006. Discriminative random fields. *International Journal of Computer Vision*, 68 (2), 179–201.
- Lee J., Han S., Byun Y., Kim Y., 2011. Extraction and Regularization of Various Building Boundaries with Complex Shapes Utilizing Distribution Characteristics of Airborne LIDAR Points, *ETRI Journal*, Vol 33, No 4, 547–557.
- Li D., Guo H., Wang C., Dong P., Zuo Z., 2012. Improved bore-sight calibration for airborne light detection and ranging using planar patches. *J. Appl. Remote Sens.* 10 (2), 024001.
- Mathworks, 2013. Adres: <http://www.mathworks.com/matlabcentral/fileexchange/28851-alpha-shapes>, funkcja `alphavol.m`, autor: Jonas Lundgren (dostęp: 20.12.2013).
- Mendela M., Borkowski A., 2013. Identyfikacja zbiorników wodnych, jako obiektów BDOT10K, w zbiorze danych lotniczego skaningu laserowego z wykorzystaniem algorytmu `alpha shape`. *Acta Sci. Pol. Geod. Descr. Terr.*, 12 (4), 13–26.
- Niemeyer J., Rottensteiner F., Soergel U., 2014. Contextual classification of lidar data and building object detection in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87, 152–165.
- OGC City Geography Markup Language CityGML Encoding Standard, 2012. <http://www.opengeospatial.org/standards/citygml> (dostęp: 07.11.2016).
- Pawluszek K., Ziaja M., Borkowski A., 2014. Ocena dokładności opracowania danych lotniczego skaningu laserowego systemu ISOK na obszarze doliny rzeki Widawy, *Acta Sci. Pol., Geodesia et Descriptio Terrarum*, 13 (3–4), 27–38.
- Priestnall G., Jaafar J., Duncan A., 2000. Extracting urban features from LiDAR digital surface models. *Computers, Environment and Urban Systems*, 24 (2), 65–78.

- Pyka K., Borowiec N., Poręba M., Słota M., Kundzierewicz T., 2012. Airborne laser scanning data for railway line survey. *Pomiary, Automatyka, Kontrola*, vol. 58 (3), 260–263.
- Ripley B.D., Kelly F.P., 1977. Markov point processes. *Journal of the London Mathematical Society*, 2 (1), 188–192.
- Rozporządzenie Ministra Spraw Wewnętrznych i Administracji, dnia 17 listopada 2011 r. w sprawie bazy danych obiektów topograficznych oraz bazy danych obiektów ogólnogeograficznych a także standardowych opracowań kartograficznych (Dz.U. Nr 279, poz. 1642).
- Sampath A., Shan J., 2007. Building boundary tracing and regularization from airborne LiDAR point clouds. *Photogrammetric Engineering and Remote Sensing* 73, 7, 805–812.
- Sampath A., Shan J., 2004. Urban modeling based on segmentation and regularization of airborne LIDAR point clouds. *Proc. ISPRS 20th Congress Commission III*, 937–941.
- Sochiera, K., 2015. Modelowanie obrysu budynku na podstawie danych lotniczego skaningu laserowego. Praca magisterska, Uniwersytet Przyrodniczy we Wrocławiu, 81.
- Sohn G., Dowman I., 2007. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62 (1), 43–63.
- Tomljenovic I., Höfle B., Tiede D., Blaschke T., 2015. Building extraction from airborne laser scanning data: An analysis of the state of the art. *Remote Sensing*, 7 (4), 3826–3862.
- Vosselman G., Maas H-G., 2010. *Airborne and terrestrial laser scanning*. Whittles Publishing, Dunbeath. 58–75, 317.
- Wei S., 2008. Building boundary extraction based on lidar point clouds data. *Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37, 157–161.
- Yang B., Xu W., Dong Z., 2013. Automated extraction of building outlines from airborne laser scanning point clouds. *Geoscience and Remote Sensing Letters, IEEE*, 10 (6), 1399–1403.
- Zhao Z., 2002. Line Simplification. <http://ra.cfm.ohiostate.edu/~zhao/algorithms/linesimp/linesimp.html> (dostęp: 01.03.2015).
- Zhao Z., Saalfeld A., 1997. Linear-time sleeve-fitting polyline simplification algorithms. *Proceedings of AutoCarto*, Vol. 13, 214–223.

AN ALGORITHM OF 2D BUILDING MODELING IN AIRBORNE LASER SCANNING POINT CLOUD OF THE ISOK PROJECT

Abstract. Airborne laser scanning data (ALS) are acquired mostly for the purpose of digital elevation models generation. In Poland, ALS data have been obtained for the whole country within the ISOK project, established for natural hazards risk mitigation. These data were used in this study to model the outlines of buildings. For this purpose an algorithm is proposed, that is a combination of α -shape algorithm and iterative total least squares adjustment. α -shape is used to detect points representing building outlines while the total least squares method is performed to receive regularized 2D building vector models. Identification of points representing outlines of buildings was performed on the point cloud thresholded at the given height with rejection of points above that height. Identification of a building as a gap (internal hull) in ALS data set is a better approximation of real building shape. For the algorithm verification a point cloud with a density of 4 points /m² is utilized. This point cloud represents a city urban area, covering 21 large buildings. The results of 2D modeling of buildings have been compared with their representation in the cadaster data base. The linear deviation between corresponding corners of modeled and represented in

cadaster data base buildings have been measured. The received mean value of the deviation equal 0.56 m is consistent with the nominal planar accuracy of ISOK ALS data. RMSE of building outline modelling calculated on the basis of linear deviations was equal 0,64 m. The results of modeling meet the requirements of Topographic Database Objects 1: 10000 (BDOT10k) and can be used for verification and updating of this data base.

Key words: airborne laser scanning, alpha shape, total least squares, ISOK, 2D-modelling

Zaakceptowano do druku – Accepted for print: 30.11.2016

Do cytowania – For citation: Sochiera K., Borkowski A., 2016. Algorytm modelowania 2D zabudowy na podstawie danych lotniczego skanowania laserowego z projektu ISOK, *Acta Sci. Pol. Geod. Descr. Terr.* 15 (1-4), 19–33.