

Zygmunt Paszotta

TEORETYCZNE PODSTAWY METODY SPASOWANIA OBSZARÓW OBRAZÓW CYFROWYCH

Streszczenie

Zagadnienie spasowania obszarów na obrazach cyfrowych związane jest ściśle z wyznaczaniem punktów homologicznych. Poprawność ich identyfikacji wpływa zasadniczo na dokładność wyznaczania elementów orientacji zdjęć, na dokładność NMT i ortofotomapy.

Metody spasowania obszarów są tematem wielu publikacji. Mimo to trudno znaleźć formalne teoretyczne podstawy, na których można budować lub weryfikować algorytmy.

Autor podejmuje próbę sformalizowania pewnych zagadnień spasowania obszarów, podaje warunki istnienia punktów i obszarów homologicznych oraz przedstawia pewne formuły i warunki dla spasowania obszarów. Teoretyczne rozważania były podstawą do zbudowania algorytmu wyznaczania punktów wiążących. Jego zastosowanie przedstawiono na przykładzie stereogramu lotniczego.

1. Wstęp

Ograniczenie udziału operatora w wyszukiwaniu punktów i obszarów homologicznych jest głównym celem automatyzacji orientacji zdjęć. Jest to również zagadnienie leżące u podstaw automatycznego wyznaczania NMT a także rekonstrukcji obrazów obiektów terenowych, gdy dane są ich rzuty.

Pojęciem ściśle powiązanim z terminem punkty homologiczne jest spasowanie (*matching*) obrazów. Różne metody spasowania są przedmiotem badań podejmowanym obecnie przez wielu autorów [Jahn, 2000; Xu, Li, 2000; Zhou, Dorrer, 2000]. Jednakże mając na uwadze aspekty naukowe jak i dydaktyczne istnieje potrzeba przedstawienia podstaw teorii spasowania obrazów. W przedstawionym artykule autor w sposób możliwie sformalizowany przedstawia rozważania dotyczące spasowania obszarów (*area matching*). To bardzo szerokie zagadnienie jest trudne do sformalizowania. Dlatego autor podejmuje się przedstawienia tylko wybranych, teoretycznych aspektów zagadnienia.

Przedstawiona teoria powstała w celu opisanego algorytmu wyszukiwania punktów wiążących dla celów aerotriangulacji [Heipke, Eder, 1998].

2. Punkty i obszary homologiczne

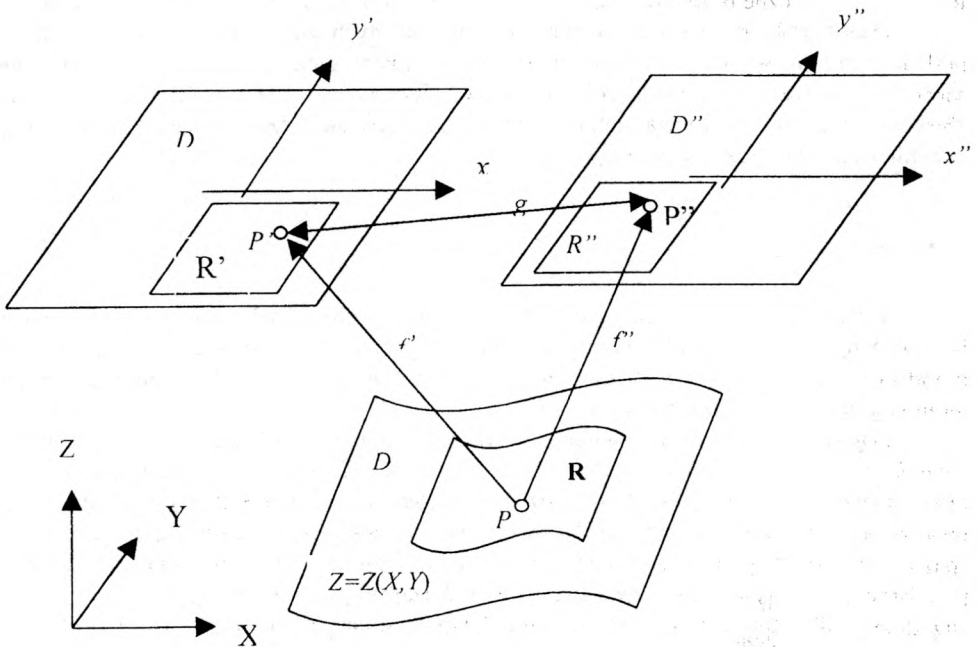
Dane są trójwymiarowy obraz terenu oraz dwa jego rzuty. Obrazy te oznaczmy odpowiednio przez

$$c(X, Y, Z), \quad c'(x', y'), \quad c''(x'', y''), \quad (1)$$

ich dziedziny będące zbiorem współrzędnych punktów przez: D, D', D'' ; a punkty należące do tych dziedzin odpowiednio przez: P, P', P'' .

Wartości funkcji c, c', c'' należą do przestrzeni kolorów. Między argumentami tych funkcji zachodzą następujące związki (rys. 2):

$$\forall P' \in D': P' = f'(P), \quad \forall P'' \in D'': P'' = f''(P) \quad (2)$$



Rys.2. Graficzne przedstawienie związków między obszarami.

Będziemy zakładali, że f' i f'' mają odpowiednio postać następujących przekształceń rzutowych:

$$x' = -c'_k \frac{a_{11}(X - X'_0) + a_{12}(Y - Y'_0) + a_{13}(Z - Z'_0)}{a_{31}(X - X'_0) + a_{32}(Y - Y'_0) + a_{33}(Z - Z'_0)},$$

$$y' = -c'_k \frac{a_{21}(X - X'_0) + a_{22}(Y - Y'_0) + a_{23}(Z - Z'_0)}{a_{31}(X - X'_0) + a_{32}(Y - Y'_0) + a_{33}(Z - Z'_0)} \quad (3)$$

$$x'' = -c''_k \frac{b_{11}(X - X''_0) + b_{12}(Y - Y''_0) + b_{13}(Z - Z''_0)}{b_{31}(X - X''_0) + b_{32}(Y - Y''_0) + b_{33}(Z - Z''_0)},$$

$$y'' = -c''_k \frac{b_{21}(X - X''_0) + b_{22}(Y - Y''_0) + b_{23}(Z - Z''_0)}{b_{31}(X - X''_0) + b_{32}(Y - Y''_0) + b_{33}(Z - Z''_0)} \quad (4)$$

Przyporządkowanie punktom $P \in D$ punktów $P' \in D'$ nie jest funkcją, gdyż

$$\neg \forall P \in D \exists P' \in D': P' = f'(P) \quad (5)$$

Podobna relacja zachodzi między punktami P i P'' .

Punkty P' , P'' są punktami homologicznymi, jeżeli spełniony jest warunek:

$$\exists P \in D: (P' = f'(P)) \wedge (P'' = f''(P)) \quad (6)$$

Wyznaczanie punktów homologicznych jest jedną z najważniejszych czynności wykonywanych w procesie strojenia zdjęć. Powyższy warunek nie określa jednakże sposobu wyznaczania tych punktów. Jako metodę tradycyjną należy uznać wyznaczanie tych punktów przez obserwatora. Może to być realizowane poprzez obserwację monokularną lub stereoskopową.

Jednym z najbardziej aktualnych zagadnień współczesnej fotogrametrii cyfrowej jest automatyzacja wyznaczania punktów homologicznych. Stosowane metody naśladują metodę realizowaną przez człowieka a polegającą na identyfikacji pojedynczych punktów poprzez analizę ich otoczeń. Opiszemy je w sposób możliwie sformalizowany.

Dane są obszary R, R', R'' o następujących własnościach:

$$R \subset D, \quad R' \subset D', \quad R'' \subset D'' . \quad (7)$$

Obszary R', R'' są obszarami homologicznymi, jeżeli spełnione są następujące warunki:

$$\forall P' \in R \exists (P'' \in R'') \wedge (P \in R) : (P' = f'(P)) \wedge (P'' = f''(P)) \quad (8)$$

$$\forall P'' \in R'' \exists (P' \in R') \wedge (P \in R) : (P' = f'(P)) \wedge (P'' = f''(P))$$

Z warunków tych wynika, że relacja

$$g : R' \times R'' \quad (9)$$

jest wzajemnie jednoznaczna. Wynika stąd, że przekształcenia (3) i (4) muszą być funkcjami oraz, że muszą istnieć funkcje odwrotne do nich. Zachodzi to wtedy gdy przyporządkowanie $Z=Z(X,Y)$ jest funkcją.

Z trzech obrazów: c, c', c'' ; rejestrować można tylko dwa c' i c'' . Rejestracja ta dla danej kamery odbywa się w płaszczyźnie obrazowej odległej od środka rzutów o stałą c_k . Ponieważ funkcje f', f'' są ciągle w punktach, które nie leżą na płaszczyźnie zniknięcia, zatem przekształcenie $g : R' \rightarrow R''$ jest homeomorfizmem. Funkcję tą zapiszemy w postaci:

$$g : (x', y') \rightarrow f''(f'^{-1}(x', y')) . \quad (10)$$

Założymy dalej, że obszary R', R'' są homeomorficzne z prostokątem.

Wybór kształtu obszaru R' musi być podyktowany praktycznymi możliwościami realizacji algorytmu spasowania obszarów na obrazach cyfrowych.

Z tych względów przyjmiemy, że R' jest kwadratem. Z własności przekształceń rzutowych wynika, że jego obrazem w przekształceniu g jest figura homeomorficzna. Gdy funkcja $Z = Z(X, Y)$ będzie przedstawiała płaszczyznę to obrazem kwadratu będzie czworokąt.

Jeżeli wyrażenia (3) i (4) przyjmą postać przekształceń afinicznych

$$\begin{aligned} \mathbf{r}' &= \mathbf{A} \mathbf{r} + \mathbf{r}'_0 \\ \mathbf{r}'' &= \mathbf{B} \mathbf{r} + \mathbf{r}''_0 \end{aligned} \quad (11)$$

gdzie:

$$\mathbf{r}' = \begin{bmatrix} x' \\ y' \end{bmatrix}, \quad \mathbf{r}'' = \begin{bmatrix} x'' \\ y'' \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} X \\ Y \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

to funkcja g będzie przekształceniem afinicznym. Wynika to z równości:

$$\mathbf{r}'' = \mathbf{BA}^{-1}(\mathbf{r}' - \mathbf{r}'_0) + \mathbf{r}_0'' \quad (12)$$

Wtedy obrazem kwadratu będzie równoległobok.

Translację uzyskamy, gdy

$$\mathbf{BA}^{-1} = \mathbf{I} \quad (13)$$

Zachodzi to wtedy, gdy:

- obszar R leży w płaszczyźnie równoległej do obu rzutni,
- zdjęcia wykonane zostały kamerą o tej samej odległości obrazowej,
- odległość między płaszczyzną w której leży obszar R i rzutnią dla obu zdjęć jest jednakowa,
- kąt skręcenia zdjęć jest równy 0.

Obrazem kwadratu będzie kwadrat.

Warunek drugi i trzeci może być zastąpiony ogólniejszym warunkiem, że skale zdjęć są jednakowe.

3. Spaszowanie dwóch obrazów cyfrowych

Rozważmy przypadek dwóch obrazów cyfrowych c' , c'' określonych następująco:

$$c'(i', j') \text{ gdzie } n' \leq i' \leq N', \quad m' \leq j' \leq M'; \quad (14)$$

$$c''(i'', j'') \text{ gdzie } n'' \leq i'' \leq N'', \quad m'' \leq j'' \leq M''$$

W pierwszym obrazie dla piksela o indeksach (i', j') jego otoczenie oznaczymy przez:

$$c'(i_1', j_1', L) \quad (15)$$

gdzie L jest liczbą całkowitą. Zbiór indeksów jest określony następująco: $\{(i', j'); (i_1' - L \leq i' \leq i_1' + L), (j_1' - L \leq j' \leq j_1' + L)\}$.

Problem spaszowania polega na znalezieniu na obrazie c'' takiego otoczenia które, uwzględniając cechy radiometryczne, jest podobne do otoczenia (15). Oznaczymy je przez

$$c''(i_1'', j_1'', L) \quad (16)$$

gdzie: $\{(i'', j''); (i_1'' - L \leq i'' \leq i_1'' + L), (j_1'' - L \leq j'' \leq j_1'' + L)\}$

Powyższy warunek można sprecyzować inaczej: szukamy takiego otoczenia (16), że odległość między obrazem (15) i (16) wyznaczona w przestrzeni kolorów jest najmniejsza.

Jakkolwiek problem poszukiwania rozwiązania w przypadku ogólnym jest bardzo ważny w artykule tym ograniczymy się do poszukiwania lokalnego rozwiązania przy założeniu, że spełnione są warunki wymienione na końcu poprzedniego rozdziału.

Zadanie nasze polegać będzie na wyznaczeniu współczynników translacji. Jeżeli ograniczymy rejon poszukiwań do otoczenia $\langle i_2' - K; i_2' + K \rangle$, gdzie K liczba całkowita, to formułę dla wyznaczenia homologicznych otoczeń można następująco.

Para liczb całkowitych $\Delta x, \Delta y$ wyznacza otoczenia homologiczne gdy spełniony jest warunek:

$$d(c', c''(\Delta x, \Delta y)) = \min_{\delta x, \delta y} (d(c', c''(\delta x, \delta y))) \quad (17)$$

gdzie:

$$c': (i', j') \rightarrow c'(i', j')$$

$$c'': (i'', j'') \rightarrow c''(i'', j'')$$

$$\{(i', j'); (i_1' - L \leq i' \leq i_1' + L), (j_1' - L \leq j' \leq j_1' + L)\}$$

$$\{(i'', j''); (i_1'' - L \leq i'' \leq i_1'' + L), (j_1'' - L \leq j'' \leq j_1'' + L)\}$$

$$i_1'' = i_1' + a_0 + \delta x, \quad j_1'' = j_1' + b_0 + \delta y$$

Liczy a_0, b_0 wyznaczają środek obszaru, w którym poszukiwane jest lokalne rozwiązanie. Mogą one być wskazane manualnie kursorem na obrazie. Wyznaczane są również iteracyjnie, gdy realizujemy tzw. metodę piramidy. Jeżeli posiadamy rozwiązanie $\Delta x, \Delta y$ dla pewnych otoczeń w obrazach to dla odpowiednich otoczeń w obrazach dwukrotnie powiększonych możemy przyjąć, że

$$b_0^{(2)} = 2(b_0^{(1)} + \Delta x)$$

$$a_0^{(2)} = 2(a_0^{(1)} + \Delta y)$$

W najprostszym przypadku, wyznaczenie minimum funkcji polega na obliczeniu wartości funkcji w regularnej siatce $2K+1$ punktów i wyznaczeniu wartości minimalnej z uzyskanych liczb. Zatem zagęszczenie siatki lub inne zwiększenie liczby jej punktów wiąże się ze znacznym zwiększeniem czasu obliczeń.

Można zatem rozpatrywać przypadek ogólny gdy $\delta x, \delta y$ są liczbami rzeczywistymi i zagęścić siatkę, w której obliczane są wartości funkcji. Założenie, że $\delta x, \delta y$ są liczbami całkowitymi jest przyjęte z uwagi na czas realizacji algorytmu. Wynikający stąd błąd spasowania jest nie większy od połowy długości piksela.

Dla kolorowych obrazów można przyjmować różne miary podobieństwa między obrazami. Jeżeli ograniczyć rozważania do obrazów z poziomami szarości to można wykorzystać miarę zależności liniowej między poziomami cyfrowymi czyli współczynnik korelacji r . Biorąc pod uwagę jego własności można przyjąć jako odległość między obrazami wielkość

$$d = 1 - r \quad (18)$$

Zastosowanie współczynnika korelacji można uzasadnić następująco.

Niech

$$c'(x', y') = T'(c(f'^{-1}(x', y')))) + \varepsilon',$$

gdzie T' jest transformacją afiniczną w przestrzeni cech spektralnych (kolorów) a ε' błędem wartości tej cechy. Wtedy przyjmując odpowiednią równość oraz oznaczenia dla obrazu c'' , zależności między cechami spektralnymi punktów homologicznych (x', y') , (x'', y'') możemy zapisać:

$$c''(x'', y'') = T''(c(f''^{-1}(x'', y'')))) + \varepsilon'' = T''(c(X, Y, Z)) + \varepsilon'' = T''(T'^{-1}(c'(x', y')))) + \varepsilon \quad (19)$$

Jest to zależność liniowa, której miarą, w przypadku przestrzeni jednowymiarowej, jest współczynnik korelacji. Zastosowanie tej czy innej miary podobieństwa nie przesądza o poprawności wykonania konkretnego spasowania. W algorytmie spasowania obszarów bardzo istotne jest jak zachowuje się w danym przypadku odległość między obrazami dla obszarów niehomologicznych w pobliżu punktu spasowania.

Ze względu na czas pracy algorytmu istotne znaczenie ma liczba parametrów, dla których wyznaczamy minimum odległości. Zastosowanie bardziej ogólnego przekształcenia niż translacja powoduje zwiększenie liczby parametrów a zatem zwiększenie wymiaru przestrzeni, w której poszukujemy minimum. Dla dwóch dodatkowych parametrów czas zwiększy się $(2K + 1)^2$ razy. Zatem jeżeli poszukiwanie odbywa się w siatce 5×5 wartości argumentów to czas obliczeń zwiększy się 25 razy. Gdy weźmiemy pod uwagę, że obrazy posiadają wymiary rzędu 10000×10000 pikseli a obliczenie współczynnika korelacji wymaga wielu działań arytmetycznych to może się okazać, że na wolniejszych komputerach 25 krotne zwiększenie czasu realizacji obliczeń nie jest realne.

Jak wspomniano wcześniej wyznaczamy pewne minimum lokalne dla odległości między obrazami. Zgodnie z równością (18) dla współczynnika korelacji wyznaczamy wartość maksymalną. Obszar rozwiązań dopuszczalnych wyznaczony jest przez pewne warunki. Uwzględnienie tych ograniczeń w procesie realizacji algorytmu oraz przyjęcie odpowiednich wartości progowych ma decydujący wpływ na poprawność wyznaczania obszarów, które możemy uznać jako homologiczne.

1. Warunek dla współczynnika korelacji

$$r(\Delta x, \Delta y) \geq r_0$$

2. Warunek dla wielkości obszaru, w którym poszukiwane jest ekstremum

$$\Delta x_i < \Delta x_0, \quad \Delta y_i < \Delta y_0$$

3. Warunek na odchylenie standardowe współczynnika korelacji

$$s_r \geq s_0$$

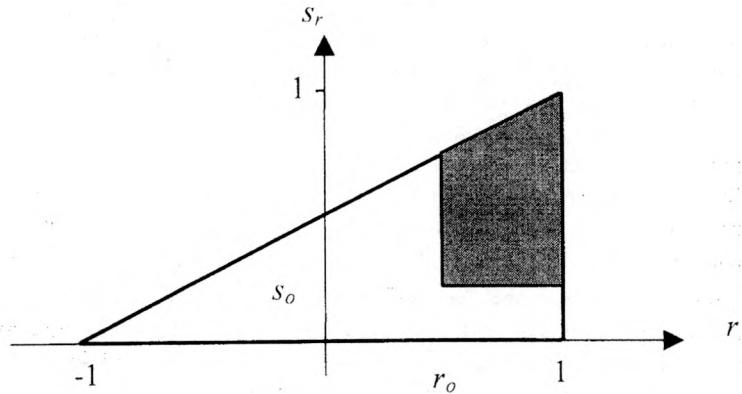
4. Warunek na wartość minimalną współczynnika korelacji

$$\min_{\Delta x, \Delta y} (r(\Delta x, \Delta y)) > r_1$$

5. Warunek na różnicę między dwoma największymi wartościami współczynnika korelacji

$$|r_{\max} - r_{r_{\max}1}| > \Delta r_0$$

Każdy z tych warunków ogranicza obszar rozwiązań. Na rysunku 1 przedstawiono graficznie warunek 1 i 3.



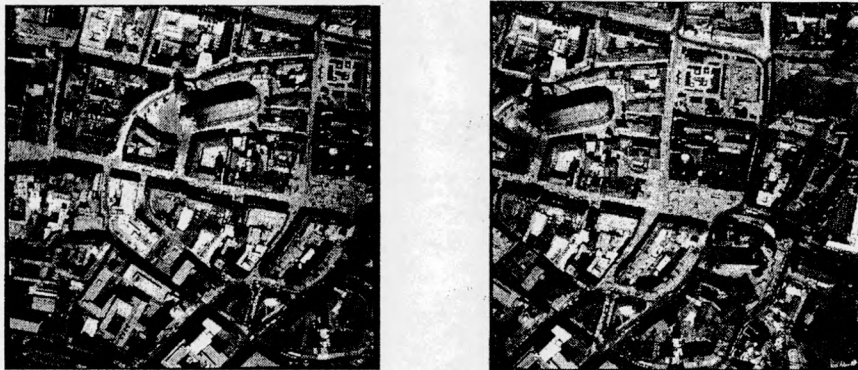
Rys.1. Warunek 1 i 3 w postaci graficznej.

Warunki 1, 3, 4, 5 są warunkami radiometrycznymi i dlatego ich spełnienie nie gwarantuje homologiczności w sensie określenia (9). Interpretacja tych warunków oraz dobór odpowiednich wartości progowych będzie przedmiotem dalszych badań. Przykładowo w warunku 2 wartości progowe można uzależnić od nachylenia terenu, a warunki 4 i 5 pozwalają uniknąć pewnych przypadków szczególnych dotyczących treści obrazów.

Podsumowując nasze rozważania należy stwierdzić, że warunki 1, 2, 3, 4, 5 wyznaczają obszary decyzyjne dla zaklasyfikowania porównywanych obszarów jako homologiczne lub nie. Obszary te są wyznaczone dla pewnych warunków geometrycznych określonych przy końcu rozdziału 3 i warunków radiometrycznych wynikających z definicji współczynnika korelacji. W przypadku, gdy te założenia nie są spełnione wnioskowanie jest trudne, gdyż trudno zbudować regułę decyzyjną, w której wyznaczając obszary decyzyjne uwzględnia się miarę odstępstwa od wspomnianych założeń.

4. Przykład zastosowania metody do wyznaczania punktów wiążących stereogramu

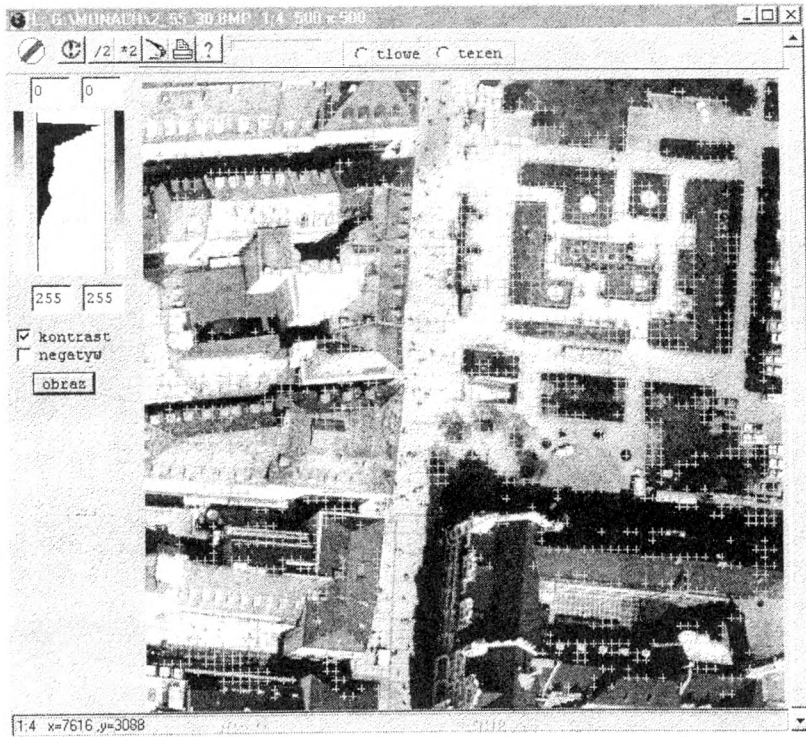
W przedstawionym dalej przykładzie przedstawione będą zdjęcia Monachium udostępnione do badań w ramach testu OEEPE-ISPRS. Zdjęcia wykonano kamerą RMK w skali 1:3000 w roku 1971. Skanowano je skanerem PS1, przyjmując wielkość piksela $25 \mu m$.



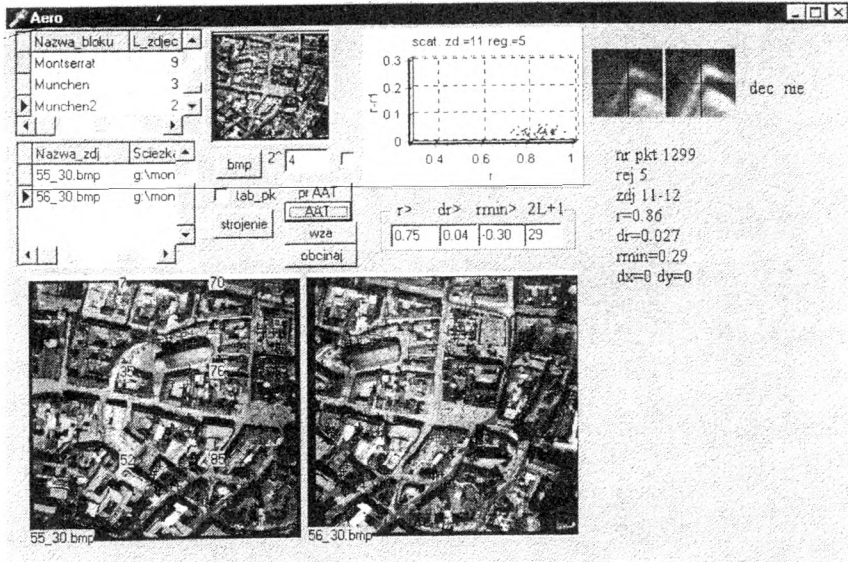
Rys.3. Stereogram zdjęć Monachium.

Algorytm poszukiwania obszarów homologicznych realizowany jest w kilku etapach.

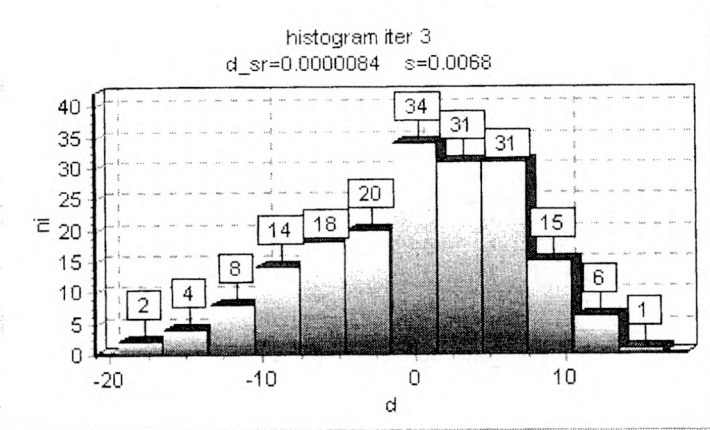
1. Na obszarze wspólnego pokrycia stosując metodę piramidy dokonuje się spasowania w regularnej siatce otoczeń. Przesunięcia Δx , Δy zapisywane są do zbioru, gdy spełnione są warunki 1, 2 z poprzedniego rozdziału.
2. Na poziomie zdjęcia oryginalnego (w piramidzie) informacje o spasowaniu kolejnych otoczeń zapisywane są do tabeli bazy danych, gdy spełnione są warunki 1, 2, 3. Zdjęcie podzielone jest na 6 obszarów będących otoczeniami punktów Grubera. Numer obszaru jest jednym z pól w tabeli. Na rysunku 4 krzyżykami zaznaczone są miejsca, w których spasowanie uznane zostało jako poprawne.
3. Następnym etapem jest weryfikacja obszarów przy uwzględnieniu warunku 1, 2, 4, 5 (rys.5). Na tym etapie operator ma możliwość obserwacji otoczeń oraz podjęcia decyzji w przypadkach wątpliwych. Operator ma możliwość obserwacji rozkładu wartości dwóch parametrów, których dotyczy rysunek 1. Rozkład ten (*scattergram*) jest przedstawiony na rysunku 5 w środkowej, górnej części.
4. Wyznaczane są elementy orientacji wzajemnej stereogramu. Daje to możliwość obserwacji modelu przestrzennego w postaci anagliflu.
5. Kolejna weryfikacja następuje w oparciu o rozkłady odległości punktów homologicznych od płaszczyzn rdzennych. Oblicza się odległość punktu będącego środkiem otoczenia na prawym zdjęciu. Granice wyznacza się na podstawie reguły 3σ lub ręcznie. Odpowiedni wykres przedstawiony jest na rysunku 6.



Rys. 4. Obszary homologiczne wybrane w drugim etapie.



Rys.5. Weryfikacja obszarów homologicznych.



Rys.6. Rozkład odległości punktów od płaszczyzny rdzenej.

Podsumowanie

Z przedstawionej pracy wynika, że sformułowanie ogólnej teorii dotyczącej wyznaczania obszarów homologicznych nie jest zadaniem łatwym. Oczywiście, intuicyjnie przyjmowane strategie bywają trudne do formalnego uzasadnienia. Rozważania przedstawione w pracy dotyczą stereogramu, jednakże były one podstawą do opracowania algorytmu wyznaczania punktów wiążących dla bloków zdjęć. Opracowanie tego algorytmu, wykonanie odpowiedniego oprogramowania oraz wykonanie koniecznych obliczeń związane było z realizacją zadań w ramach testu OEEPE-ISPRS [Heipke, Eder, 1998].

Literatura

1. Heipke C., Eder K., 1998, *Performance of tie point extraction in automatic aerial triangulation*, OEEPE Official Publication No 35.
2. Jahn H., 2000, *Stereo matching for pushbroom stereo cameras*, International Archives of Photogrammetry and Remote Sensing, Vol XXXIII, Part B3.
3. Xu Z., Li Z., 2000, *Last median of squares matching for automated detection of surface deformation*, International Archives of Photogrammetry and Remote Sensing, Vol XXXIII, Part B3.
4. Zhou X., Dorrer E., 2000, *Non-linear scale and orientation free correlation matching algorithm based on edge correspondence*, International Archives of Photogrammetry and Remote Sensing, Vol XXXIII, Part B3

Recenzował: prof. dr hab. inż. Józef Jachimski