# Pomiary przestrzenne w polu obserwacyjnym stereoendoskopu

**Streszczenie.** W pracy przedstawiono procedurę trójwymiarowej metrycznej rekonstrukcji powierzchni na podstawie zapisu wideo ze stereoendoskopu stosowanego w chirurgii minimalnie inwazyjnej. Metoda bazuje na dopasowaniu obszarami fragmentów obrazów. Wyniki rekonstrukcji porównano z danymi uzyskanymi dla tego samego obiektu metodą referencyjną. Średni błąd rekonstrukcji uzyskany dla poszczególnych klatek sekwencji wynosi od 2,1 do 4,2 mm.

**Abstract**. In this work, we present a procedure for 3-dimensional metric surface reconstruction based on video data from a stereoendoscope used in minimally invasive surgery. The reconstruction is based on stereo block matching algorithm. The results of the reconstruction were compared to a reference data set obtained simultaneously for the same object. The mean reconstruction error obtained for individual frames falls within the range of 2.1 to 4.2 mm. (**Spatial measurements in the field of view of a stereoendoscope**).

**Słowa kluczowe**: stereoendoskopia, metryczna rekonstrukcja trójwymiarowa, dopasowanie obszarami, algorytm MLS. **Keywords**: stereoendoscopy, 3-dimensional metric reconstruction, block matching, moving least squares.

## Wstęp

Techniki endoskopowe mają szeroki zakres zastosowań medycznych, m. in. w badaniach przewodu pokarmowego i dróg oddechowych, oraz w chirurgii minimalnie inwazyjnej. Mimo swojej użyteczności umożliwiają one obserwację trójwymiarowych struktur tylko na płaskim obrazie, w wąskim polu obserwacji. Manewrowanie miniaturowymi narzędziami chirurgicznymi widzianymi jedynie za pomocą endoskopu sprawia trudności nawet doświadczonym operatorom. Ograniczona jest także możliwość oceny rzeczywistej wielkości zmian patologicznych widocznych na obrazie, która może być istotna dla prawidłowej diagnozy.

W ostatnich latach prowadzone są intensywne badania nad zastosowaniem w endoskopii technik wizyjnych do przestrzennej rekonstrukcji obserwowanych struktur. Informacje przestrzenne dotyczące obiektów widzianych przez endoskop mogą być wykorzystane do wykonywania pomiarów, do wirtualnego rozszerzenia pola obserwacji poprzez fuzję danych z kilku obrazów zarejestrowanych w różnych miejscach [1] lub do wygenerowania częściowo wirtualnego widoku, w którym narzędzia nie przesłaniają operowanej tkanki [2]. Można je także wykorzystać do śledzenia i naprowadzania narzędzi chirurgicznych [3].

Wprowadzenie do użytku endoskopów z podwójnym torem optycznym pozwoliło na zastosowanie w endoskopii technik stereowizyjnych [4], [5]. Dzięki jednoczesnej akwizycji pary obrazów dają one możliwość rekonstrukcji powierzchni, której kształt zmienia się w czasie. Jest to ich główną zaletą w porównaniu z rozwiązaniami bazującymi na pojedynczym torze wizyjnym.

Tematem niniejszej pracy jest rekonstrukcja powierzchni narządu na podstawie obrazu stereoendoskopowego. Zaprezentowany algorytm rekonstrukcji wykorzystuje zapis wideo ze skalibrowanego zestawu dwóch kamer do wyznaczenia metrycznych współrzędnych punktów odpowiadających poszczególnym pikselom obrazu (rys 1).

Rekonstrukcję trójwymiarową na podstawie pary obrazów można przeprowadzić jedynie dla punktów, dla których określone jest odwzorowanie pomiędzy tymi obrazami, tj. dla punktów, dla których znane jest położenie ich odpowiednika w drugim obrazie pary. Większość metod opisywanych w literaturze określa to odwzorowanie jedynie dla charakterystycznych obszarów obrazu takich jak krawędzie i narożniki. Rekonstrukcja tego rodzaju ma selektywny charakter, ponieważ zagęszczenie krawędzi i narożników zależy od aktualnej zawartości sceny. W metodzie opisywanej w tej pracy podstawą do określenia odwzorowania pomiędzy parą obrazów jest algorytm matching) dopasowania obszarami (ang. block

wykorzystujący wzajemne podobieństwo regionów obrazu ograniczonych kwadratowym oknem. Podobieństwo to można określić niezależnie od występowania w oknie narożników i krawędzi. Zastosowana metoda daje więc teoretycznie możliwość obliczenia trójwymiarowych współrzędnych dla każdego piksela obrazu. W praktyce część zrekonstruowanych punktów obarczona jest jednak znacznym błędem i musi zostać wyeliminowana.

Odpowiedni dobór parametrów metody rekonstrukcji jest wynikiem kompromisu pomiędzy gęstością a precyzją wyznaczenia wynikowego zbioru punktów. W niniejszej pracy wykorzystano algorytm pozwalający na względnie gęstą rekonstrukcję. Celem jest odpowiedź na pytanie, czy jego dokładność jest wystarczająca do zastosowania w pomiarach przestrzennych w polu widzenia endoskopu.

W celu oceny dokładności zaproponowano procedurę ręcznego pomiaru długości obiektu widzianego na obrazie na podstawie wyników rekonstrukcji uzyskanych badaną metodą. Wyniki pomiarów są istotnie różne od wyników uzyskanych metodą referencyjną. Obliczono także średni błąd dla wszystkich zrekonstruowanych punktów sceny. Wartości tego błędu dla poszczególnych klatek sekwencji mieszczą się w zakresie od 2,1 do 4,2 mm.



Rysunek 1. Działanie przedstawionej metody na przykładowych danych: (a) para obrazów endoskopowych, (b) trójwymiarowa rekonstrukcja powierzchni, (c) wynik rekonstrukcji po usunięciu błędów i wygładzeniu powierzchni. Osie na rysunkach (b) i (c) wyskalowano w milimetrach



Rysunek 2. Schemat blokowy procedury rekonstrukcji sceny na podstawie stereoskopowej pary obrazów.

### Stereoskopowa rekonstrukcja sceny

Celem stosowania algorytmów stereowizyjnych jest trójwymiarowa rekonstrukcja sceny, tj. wyznaczenie zbioru punktów trójwymiarowych reprezentujących rzeczywiste obiekty widoczne na obrazie. Wynik rekonstrukcji nie jest jednak ani kompletnym, ani dokładnym odwzorowaniem rzeczywistości. Niektóre z punktów nie mają swojego fizycznego odpowiednika. Położenie innych, prawidłowo zrekonstruowanych, jest obarczone niepewnością, wskutek czego rekonstrukcja gładkiej powierzchni może mieć nieregularny kształt. Pierwszy typ błędów można usunąć stosując lokalną analizę statystyczną. Negatywne skutki błędów drugiego typu można zminimalizować stosując metody nieliniowej interpolacji. Przykładowy schemat algorytmu rekonstrukcji połączonej z korekcją wyników przedstawiono na rysunku 2. Najważniejsze jego elementy zostały omówione w dalszej części tego rozdziału.

Podstawą do rekonstrukcji stereoskopowej jest różnica w położeniu projekcji tego samego punktu sceny na obrazie lewej i prawej kamery. Funkcja przyporządkowująca tę różnicę każdemu z pikseli obrazu nosi nazwę mapy dysparycji. Rekonstrukcja wymaga także znajomości modelu matematycznego każdej z kamer oraz opisu geometrii ich wzajemnego układu. Model kamery opisuje związek pomiędzy położeniem obserwowanego punktu w przestrzeni, wyrażonym we współrzędnych metrycznych, a położeniem jego projekcji na obrazie, we współrzędnych pikselowych. Najczęściej wykorzystywany jest rozszerzony model kamery otworkowej [6]. Geometria układu kamer determinuje natomiast zależność pomiędzy położeniem projekcji tego samego punktu na prawym i lewym obrazie.

Problem dopasowania obrazów leżący u podstaw metod rekonstrukcji polega na znalezieniu odwzorowania między dwiema przestrzeniami dwuwymiarowymi. Dzięki znajomości rozmieszczenia kamer, dla każdej projekcji punktu na obrazie lewym można znaleźć linię na obrazie prawym (tzw. linię epipolarną), na której powinna leżeć odpowiadająca projekcja tego samego punktu. Zadanie dopasowania można więc sprowadzić do pojedynczego wymiaru. Aby ułatwić poszukiwanie odpowiadających sobie punktów obrazy poddaje się zwykle tzw. rektyfikacji, w rezultacie której poszczególne wiersze pikseli są równoległe do linii epipolarnych. Najczęściej stosowaną metodą dopasowania obrazów pary stereoskopowej jest tzw. dopasowanie obszarami (ang. *block matching*) [7]. Jest to metoda lokalna bazująca na wartościach jasności pikseli. Zakłada ona, że dla wybranego wycinka obrazu  $I_i$  (określonego 2-wymiarowym oknem W) można w obrazie  $I_2$  znaleźć odpowiadający wycinek o wartościach pikseli minimalizujących pewną funkcję kosztu. Przykładem takiej funkcji kosztu jest suma wartości bezwzględnych różnic (SAD, ang. *Sum of Absolute Differences*), zdefiniowana dla okna w położeniu *i*, *j* jako:

(1) 
$$SAD(i, j, d) = \sum_{[m,n] \in W} |I_1[i+m, j+n] - I_2[i+m, j+n+d]|.$$

Dzięki wcześniejszej rektyfikacji, przeszukiwanie obrazów  $I_1$  i  $I_2$  można przyspieszyć, ograniczając przemieszczenie okna do wierszy pikseli o tym samym indeksie. Dysparycję można wtedy opisać za pomocą pojedynczego parametru *d*. Szukany wycinek obrazu  $I_2$ , odpowiadający analizowanemu fragmentowi obrazu  $I_1$ , znajdowany jest przez wyznaczenie minimum funkcji kosztu względem dysparycji.

Punkty nie posiadające fizycznego odpowiednika w zarejestrowanej scenie można usunąć, czyniąc pewne założenia co do lokalnego rozkładu wyników rekonstrukcji. Jedną z możliwości jest zastosowanie kryterium średniej odległości pomiędzy zbiorem *n* najbliższych punktów [8]. Dla każdego punktu rozpatrywany jest rozkład jego odległości od punków sąsiednich. Na podstawie rozkładu obliczana jest odległość średnia  $\mu_d$  i odchylenie standardowe odległości  $\sigma_d$ . Następnie wybierane są punkty sąsiedztwa, których odległość od rozpatrywanego punktu nie spełnia warunku  $\mu_d + a\sigma_d$ , gdzie *a* jest pewnym współczynnikiem skalującym. Z wynikowej chmury punktów wykluczane są wszystkie punkty, które co najmniej raz nie spełniły powyższego warunku przynależności

Z punktu widzenia analizy sceny bardziej praktyczne zastosowanie od chmury punktów ma odpowiadająca jej powierzchnia. Pierwszym etapem jej wyznaczenia jest wygładzanie chmury punktów, przy założeniu, że większość zrekonstruowanej sceny składa się z gładkich obszarów. Wykorzystuje się w tym celu metody lokalnej aproksymacji za pomocą funkcji wielomianowych. Przykładem takiej metody jest algorytm MLS (ang. *Moving Least Squares*) [9]. Jego idea została przedstawiona na rysunku 3. Na rysunku zaprezentowano wariant dwuwymiarowy, metodę można jednak łatwo dostosować do 3 wymiarów.



Rysunek 3. Działanie metody MLS (przypadek dwuwymiarowy): (a) liniowa aproksymacja sąsiedztwa punktu P, sąsiedztwo określone promieniem R, (b) projekcja punktu P na prostą H, (c) aproksymacja wielomianowa sąsiedztwa punktu P wielomianem g w dziedzinie określonej przez prostą H, (d) wyznaczenie punktu wynikowego  $P_g$  jako projekcji punktu  $P_H$  na wielomian g.

Korekcja położenia punktu P metodą MLS rozpoczyna się od wyznaczenia jego sąsiedztwa, czyli zbioru punktów oddalonych co najwyżej o zdefiniowany promień R. Sąsiedztwo punktu P jest aproksymowane płaszczyzną Hmetodą najmniejszych kwadratów. Następnie określana jest projekcja punktu P na płaszczyznę H i wyznaczana jest aproksymacja wielomianowa sąsiedztwa, przy czym dziedziną dla szukanego wielomianu jest płaszczyzna H. Aproksymacja ma charakter ważony, tzn. im bliżej punktu Pznajduje się dany punkt sąsiedztwa, tym większy ma on wpływ na końcowy wynik. Ostatecznie skorygowane położenie punktu P obliczane jest jako projekcja punktu  $P_H$ na powierzchnię g. Opisana operacja powtarzana jest dla każdego z elementów wejściowego zbioru punktów.

#### Materiały i metody

W charakterze danych wejściowych dla badanej metody wykorzystano sekwencję wideo z bazy Hamlyn Centre [4], zawierającą ruchomy fantom serca (rys 1a). Sekwencja została zarejestrowana za pomocą laparoskopu należącego wyposażenia systemu chirurgicznego da do Vinci. Laparoskop posiada podwójny tor wizyjny przekazujący do sensora stereoskopowe pary obrazów. Osie optyczne kanałów wizyjnych są wzajemnie równoległe i oddalone od siebie o 5,5 mm. Rozdzielczość każdego z obrazów wynosi 360 na 288 pikseli. Oprócz laparoskopowych sekwencji wideo w bazie Hamlyn Centre dostępne są także parametry kalibracyjne kamer laparoskopu oraz dane referencyjne, umożliwiające wyznaczenie wzorcowej geometrii sceny dla 19 klatek sekwencji stereo. Dane referencyjne pochodzą z fuzji pomiarów stereowizyjnych z tomografią komputerową wykonaną ze środkiem kontrastowym.

Algorytm rekonstrukcji uruchomiono na obrazach przekształconych do skali szarości. Do wyznaczenia odwzorowania pomiędzy obrazami wykorzystano opisaną wcześniej metodę dopasowania obszarami, zadając zmienność dysparycji w przedziale od 0 do 64 pikseli. Dane wyjściowe oczyszczano z błędów rekonstrukcji poprzez lokalną analizę statystyczną odległości pomiędzy punktami. Pozostałe punkty wygładzano metodą MLS.

Jako miarę błędu rekonstrukcji przyjęto średnią odległość euklidesową pomiędzy zrekonstruowanymi punktami a ich odpowiednikami w zbiorze referencyjnym. Przy wyznaczaniu tej wartości brano pod uwagę tylko prawidłowo zrekonstruowane punkty oraz ich odpowiedniki w danych wzorcowych

Przed wykonaniem właściwych badań zaproponowanej metody rekonstrukcji określono rozmiar okna wyszukiwania, dla którego obliczana jest wartość parametru *SAD*. Dla okna kwadratowego o wielkości zmienianej w zakresie od 39 do 69 pikseli monitorowano średni błąd rekonstrukcji oraz liczbę pikseli, dla których algorytm był w stanie określić trójwymiarowe współrzędne. Wyniki przedstawiono na rysunku 4. Zaobserwowano, że średni błąd przyjmuje wartość minimalną dla okna o rozmiarze 49 pikseli. Jednocześnie stwierdzono spadek liczby poprawnie zrekonstruowanych punktów wraz ze zwiększaniem okna.

Zaproponowana metoda została następnie zastosowana do prostych pomiarów liniowych na zrekonstruowanej powierzchni trójwymiarowej. Zadanie pomiarowe polegało na recznym zaznaczeniu na obrazie endoskopowym ścieżki wzdłuż widocznej zmiany. Znajac trójwymiarowe współrzędne każdego z pikseli obrazu wyznaczano długość ścieżki, a tym samym rzeczywista rozpietość zmiany wzdłuż płaszczyzny rekonstrukcji. Wyniki pomiarów wykonanych na zrekonstruowanych danych proponowaną metoda porównano z wynikami uzyskanymi tą samą metodą na danych referencyjnych. Dla każdej z metod wykonano serie 30 pomiarów na 3 wybranych klatkach sekwencji.





#### Wyniki rekonstrukcji

Dla wyznaczonego rozmiaru okna obliczeniowego wykonano rekonstrukcję wszystkich klatek sekwencji wejściowej. Wartość średnią błędu rekonstrukcji wyznaczoną dla poszczególnych klatek przedstawiono na rysunku 5. Dla środkowej części sekwencji (klatki 8-14) zaobserwowano wysokie wartości błędu. Przyczyną pogorszenia się jakości rekonstrukcji mógł być fakt, że w środkowej części zapisu rejestrowany fantom znajdował się w fazie skurczu, wskutek czego wystąpiło rozmycie obrazu.

W dalszej części przeprowadzonych testów badano przydatność zaproponowanej metody do ręcznych pomiarów struktur anatomicznych obserwowanych na obrazie endoskopowym. Sposób wykonywania pomiarów przedstawiono na rysunku 6a. Na odcinku pomiarowym o długości 9,25 mm wartość bezwzględna różnic pomiędzy wartościami średnimi pomiarów wykonanych metodami testowaną i referencyjną wynosiła od 0,46 do 1,12 mm (rys 6b). Zakładając normalność rozkładów dla wyników pomiarów uzyskanych obiema metodami stwierdzono, że różnice te są istotne statystycznie na poziomie  $\alpha = 0,05$ .



Rysunek 5. Wartość średnia błędu rekonstrukcji dla kolejnych par obrazów badanej sekwencji.



Rysunek 6. Ręczny pomiar długości ścieżki na obrazie endoskopowym: (a) sposób wykonania pomiaru, (b) wartości średnie wyników pomiarów zaproponowaną metodą (test.) i metodą referencyjną (ref.) dla trzech wybranych klatek sekwencji. Słupki błędów zaznaczono jako granice przedziału ufności przy p = 0.05.



Rysunek 7. Przestrzenny rozkład błędu rekonstrukcji w płaszczyźnie równoległej do płaszczyzny obrazowania endoskopu. Wartości błędów przedstawiono w postaci izolinii i wyrażono w milimetrach. Na rysunku przedstawiono tylko poprawnie zrekonstruowane piksele obrazu. Strzałkami zaznaczono nieregularności rozkładu błędu spowodowane występowaniem rozbłysków.

Ostatnim etapem badania zaproponowanej metody było określenie w jakim stopniu lokalne cechy obrazu wpływają na dokładność odtworzenia sceny w wybranym regionie. Porównując otrzymane wyniki rekonstrukcji z danymi referencyjnymi dla jednej z ramek sekwencji wyznaczono przestrzenny rozkład błędu rekonstrukcji w płaszczvźnie równoległej do płaszczyzny obrazowania. Zaobserwowano rozkład niejednorodny (rys. 7). W centralnej części obrazu błąd przyjmuje niższe wartości niż w obszarach oddalonych od środka. W regionach występowania rozbłysków pochodzących od oświetlenia występują dodatkowe zaburzenia rozkładu. Potwierdzają one negatywny wpływ kierunkowego odbicia światła od powierzchni na dokładność rekonstrukcji.

### Podsumowanie

W niniejszej pracy opisano procedurę trójwymiarowej rekonstrukcji powierzchni na podstawie zapisu wideo ze stereoendoskopu, wraz z metodami usuwania błędnie zrekonstruowanych punktów i wygładzania powierzchni. Dokładność przedstawionej metody badano poprzez wykorzystanie wyników rekonstrukcji w prostej procedurze pomiarowej, mającej na celu określenie długości ścieżki narysowanej na obrazie endoskopowym. Podjęto także próbę określenia zależności pomiędzy wartościami błędu metody a lokalnymi cechami obrazu.

Pierwsze przeprowadzone testy zaproponowanej metody, miały na celu dobór odpowiedniej wielkości okna obliczeniowego algorytmu dopasowującego dla stereoskopową parę obrazów. Zaobserwowano, że o ile istnieje pewien rozmiar okna, dla którego błąd rekonstrukcji przyjmuje minimum, o tyle każde zwiększenie okna powoduje spadek liczby pikseli dla których możliwa jest rekonstrukcja trójwymiarowa (rys. 4). Na monotoniczny charakter zależności wpływają dwa czynniki. Po pierwsze, im większy fragment sceny zawiera badany obszar, tym trudniejsze jest znalezienie jego odpowiednika w drugim z obrazów. Po drugie, im większy jest rozmiar okna, tym większy staje się margines, dla którego wartość mapy dysparycji nie jest zdefiniowana, ponieważ część okna przeszukiwania wykracza poza obszar obrazu.

Wyniki rekonstrukcji dla kolejnych klatek zapisu z endoskopu wskazują na podatność badanego rozwiązania na rozmycie spowodowane ruchem rejestrowanego obiektu. Pomimo tego, że metoda jest stereoskopowa, dzięki czemu nie dotyczą jej błędy związane z niesynchroniczną akwizycją scen zmiennych w czasie, pozostaje jednak w pewnym stopniu wrażliwa na dynamikę sceny.

Dzięki wyznaczeniu przestrzennego rozkładu błędu rekonstrukcji na płaszczyźnie obrazowania kamerv stwierdzono wyraźną zależność pomiędzy wartością błędu a odległościa od środka obrazu. Za duże wartości błędów w pobliżu krawędzi obrazu mogą odpowiadać zarówno zniekształcenia wprowadzane przez optykę endoskopu jak i niski stopień oświetlenia tych regionów, utrudniający znalezienie dopasowania. Fakt, że algorytm rekonstrukcji uwzględniał kompensację modelu zniekształceń optycznych sugeruje, że większe znaczenie miał drugi z wymienionych czynników. Oprócz globalnej zależności pomiędzy jasnością sceny a dokładnością metody, zaobserwowano także widoczny wpływ rozbłysków spowodowanych kierunkowym odbiciem światła od powierzchni fantomu.

Porównanie z danymi referencyjnymi pokazało, że wartości błędów uzyskiwane za pomocą prezentowanej metody są zbyt duże, aby można ją było zastosować nawet w prostych zadaniach pomiarowych takich jak wyznaczenie długości wzdłuż płaszczyzny rekonstrukcji. W celu poprawy dokładności konieczne wydaje się zastąpienie algorytmu dopasowania obszarami innymi metodami dopasowania obrazów, wykorzystującymi większą liczbę informacji zawartych w obrazie, w tym barwą pikseli oraz punktami charakterystycznymi, bazującymi na gradiencie jasności pikseli.

#### LITERATURA

- Soper T., Porter M., Seibel E., Surface Mosaics of the Bladder Reconstructed from Endoscopic Video for Automated Surveillance, *IEEE Transactions on Biomedical Engineering*, 59 (2012), n. 6, 1670-1680.
- [2] Morugues F., Devernay F., Coste-Manière E., 3D Reconstruction of the operating field for image overlay in 3-D endoscopic surgery, *IEEE and ACM International Symposium* on Augmented Reality, (2001), 191-192.
- [3] Wengert C., Bossard L., Häberling A., Baur C., Székely G., Cattin P., Endoscopic Navigation for Minimally Invasive Suturing, *Medical Image Computing and Computer Assisted Interventions*, 10 (2007), 620-627.
- [4] Stoyanov D., Visentini Scarzanella M., Pratt P., Yang G.-Z., Real-Time Stereo Reconstruction in Robotic Assisted Minimally Invasive Surgery, *Medical Image Computing and Computer* Assisted Interventions, (2010) 275-282
- [5] Yang B., Wong W., Liu C., Poignet P., 3D soft tissue tracking using spatial-color joint probability distribution and thin-plate spline model, *Pattern Recognition*, 47 (2004), n. 9, 2962-2973.
- [6] Heikkilä J., Silvén O., A four-step camera calibration procedure with implicit image correction, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (1997) 1106-1112.
- [7] Konolige K., Small Vision Systems: Hardware and Implementation, *Robotics Research: The Eighth International Symposium* (1998) 203-212.
- [8] Rusu R., Marton Z., Blodow N., Dolha M., Beetz M., Towards 3D Point Cloud based Object Maps for Household Environments, *Robotics and Autonomous Systems*, 56 (2008), n. 11, 927-941.
- [9] Alexa M., Behr J., Cohen-Or D., Fleishman S., Levin D., Silva C., Computing and Rendering Point Set Surfaces, *IEEE Transactions on Visualization and Computer Graphics*, 9 (2003), n. 1, 3-15.

**Autor**: mgr inż. Adrian Goral, AGH Akademia Górniczo-Hutnicza, Katedra Metrologii i Elektroniki, Al. Mickiewicza 30, 30-059 Kraków, E-mail: <u>adrian.goral@agh.edu.pl</u>.