

Paweł Gardziński
Łukasz Kamiński
Krzysztof Kowalak
Sławomir Maćkowiak
Katedra Telekomunikacji Multimedialnej i Mikroelektroniki
ul. Polanka 3, 61-131 Poznań,
tel. + (48) 6653900, fax: + (48) 6653899,
email: {pgardzinski, lkaminski, kkowalak, smack}@multimedia.edu.pl
web: www.multimedia.edu.pl

POPRAWA JAKOŚCI MODELU WOKSELOWEGO NA PODSTAWIE HISTOGRAMÓW OBRAZÓW REPROJEKCJI

Streszczenie: Niniejsza praca dotyczy metody rekonstrukcji modeli wolumetrycznych rzeczywistych obiektów przy pomocy systemów wykorzystujących układy wielokamerowe. Zaprezentowana została metoda poprawy jakości rekonstrukcji dzięki dopasowaniu wartości wybranych parametrów zewnętrznych i wewnętrznych. Zaproponowane rozwiązanie jest wbudowane w procedury rekonstrukcji wokselowej i bazuje na analizie histogramów obrazów wejściowych oraz obrazów reprojektacji modelu do płaszczyzny danego obrazu wejściowego.

1. WSTĘP

Trójwymiarowe modele komputerowe obiektów są kluczowymi elementami wielu multimedialnych systemów trójwymiarowych. Ostatnimi czasy bardzo popularna stała się rekonstrukcja realistycznych obiektów trójwymiarowych rzeczywistych obiektów, na podstawie serii zdjęć lub sekwencji wizyjnej. Przeważnie wymagane są modele o wysokiej jakości, które mogą być następnie wykorzystane w tworzeniu obiektów z użyciem drukarek 3D.

Najpopularniejszymi metodami rekonstrukcji 3D modeli obiektów bazującymi na obrazach są: techniki oparte na estymacji map głębi z dwóch lub większej liczby widoków obiektu [6, 10], rekonstrukcje polegające na pasowaniu cech charakterystycznych między obrazami [11, 12] oraz techniki rekonstrukcji wolumetrycznej z wielu widoków [8, 9]. Autorzy skupili się na trzeciej technice rekonstrukcji modeli trójwymiarowych.

Podczas przeprowadzania badań rekonstrukcji modeli trójwymiarowych obiektów z wykorzystaniem algorytmu rekonstrukcji wolumetrycznej, autorzy musieli zmierzyć się z problemem dokładności wewnętrznych i zewnętrznych parametrów kamery oraz wpływem tych parametrów na jakość tworzonego modelu. Parametry kamery (pozycja, rotacja i ogniskowa) przypisane do obrazów wejściowych, mogą zawierać drobne błędy, które mogą prowadzić do znaczących zmian w stworzonym modelu podczas wykonywania rekonstrukcji wolumetrycznej. W wielu przypadkach nie ma możliwości ponownego przeprowadzenia pomiaru parametrów z większą dokładnością, ponieważ treść (sekwencja wizyjna lub obrazy) została już stworzona i opublikowana. Błędy w parametrach kamery mogą być zaobserwowane przez porównanie jednego obrazu wejściowego (referencyjnego) z płaszczyzną obrazu stworzonego modelu trójwymiarowego, która jest wsteczną projekcją do określonego widoku.

Powszechnie używaną metodą wyznaczania wewnętrznych i zewnętrznych parametrów kamery między wszystkimi widokami jest technika estymacji parametrów globalnych *bundle adjustment*. Jest to algorytm optymalizacyjny, który wymaga bardzo liczego zbioru cech lokalnych, którymi zwykle są cechy wygenerowane przez transformację cech niezależnych od skali (ang. Scale-Invariant Feature Transform, SIFT [1]). Niestety, tego typu cechy lokalne są związane z wysoką złożonością obliczeniową ich wyznaczania. Ponadto, nie zawsze możliwe jest znalezienie wystarczającej liczby cech dla algorytmu, by zapewnić jego prawidłowe działanie. Algorytm *bundle adjustment* jest algorytmem iteracyjnym, który minimalizuje globalny błąd parametrów kamery, wykorzystując algorytm Levenberg-Marquardt [2], który również charakteryzuje się wysoką złożonością obliczeniową.

W niniejszej pracy zaproponowano podejście, które pozwala w prosty sposób dostosować estymowane parametry kamery do uzyskania trójwymiarowego modelu wolumetrycznego wysokiej jakości. Proponowana przez autorów rozwiązanie może być stosowane zamiast globalnej estymacji parametrów kamery w procesie rekonstrukcji wolumetrycznej. Prostota proponowanego rozwiązania leży w wykorzystaniu histogramów obrazów uzyskanych w procesie reprojektacji modelu trójwymiarowego na płaszczyznę obrazu. W ten sposób uzależnia się korekcję parametrów kamery od rotacji i ogniskowej kamery. Metoda ta wymaga podania wstępnych parametrów kamery, które mogą zawierać niewielkie błędy. Tak więc, zarówno wewnętrzne jak i zewnętrzne parametry kamery nie muszą być wyznaczane ponownie, jak to ma miejsce w przypadku algorytmu *bundle adjustment*. Proponowane rozwiązanie, w porównaniu do *bundle adjustment*, nie jest iteracyjne, dzięki czemu możliwe jest uzyskanie wymaganych parametrów kamery w bardzo krótkim czasie, co jest zaletą proponowanej metody.

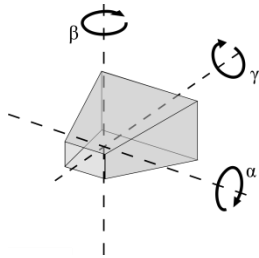
Proponowana metoda nie ma na celu poprawienia rzeczywistych parametrów kamery, a jedynie dostosować parametry obrazów wejściowych, których nie można dopasować do rekonstruowanego modelu.

Artykuł podzielony jest na 4 sekcje. W sekcji 2 przedstawiono ideę korekcji parametrów kamery z użyciem histogramów obrazu. Najpierw przedstawione zostało podejście do korekcji macierzy rotacji (wyrażonej przez kąty Eulera: α , β oraz γ). Następnie, przedstawiono sposób korekcji ogniskowej kamery. W sekcji 3

zaprezentowane zostały wyniki eksperymentów. Sekcja 4 zawiera podsumowanie i propozycje dalszych badań.

2. KOREKCJA NA PODSTAWIE HISTOGRAMÓW

W niniejszej pracy wykorzystano *Rekonstrukcję Wokselową* (ang. *Voxel Reconstruction* lub *Voxel Carving*), która została opisana w [3], konkretniej zastosowane zostało rozwiązanie przedstawione w [4]. Generalnie możliwe jest zastosowanie dowolnej innej techniki modelowania 3-D z tym zastrzeżeniem, żeby umożliwiała ona wygenerowanie obrazu reprojekcji odpowiadającego obrazowi wejściowemu w celu przeprowadzenia poprawy jakości modelu na podstawie histogramów. Rysunek 1 przedstawia oznaczenia kątów obrotu macierzy rotacji wykorzystywane w proponowanym rozwiązaniu (zazwyczaj zawarte w macierzy parametrów zewnętrznych).



Rys. 1. Oznaczenia parametrów macierzy rotacji w wykorzystywanym układzie odniesienia.

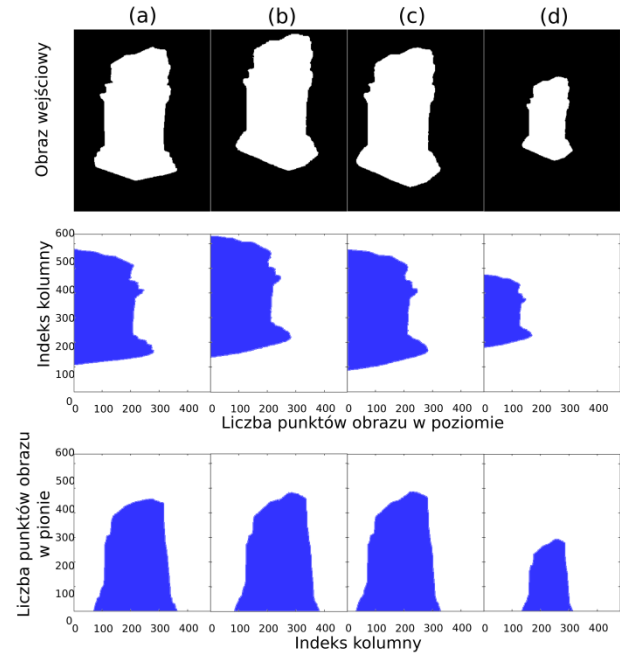
Wiadomym jest, że zmiana konkretnego parametru skutkuje znanymi zniekształceniami w obrazie reprojekcji dla konkretnego obrazu wejściowego. Na rysunku 2 pokazanych zostało kilka przykładowych zniekształceń obrazu reprojekcji (na podstawie modelu *Temple* – zbiór testowy z [6]). Zniekształcenia te są najlepiej dostrzegalne na podstawie obserwacji histogramów (wyznaczonych jako suma białych punktów w obrazie w rzędach i kolumnach). Łatwo zauważyć, że zmiany poszczególnych parametrów są dobrze widoczne na konkretnych rodzajach histogramów np. zmiana w kącie α widoczna jest jako przesunięcie histogramu pionowego, zaś błąd w kącie β widoczny jest jako przesunięcie histogramu poziomego itd. Metody estymacji poprawki parametrów przedstawione zostały w dalszej części pracy.

2.1. Estymacja korekty błędów kątów α i β

W celu wyznaczenia poprawki obrotu $\Delta\alpha$ i-tego obrazu wejściowego (i-tego widoku), należy zacząć od wyznaczenia poziomego i pionowego histogramu zarówno dla obrazu wejściowego jak i odpowiadającego mu obrazu reprojekcji. Następnie wyznaczana jest korelacja między tymi histogramami co pozwala wyznaczyć ich wzajemne przesunięcie (w punktach obrazu). Na koniec, wykorzystywany jest model kamery otworkowej, aby wyznaczyć kąt obrotu kamery, który odpowiada wartości przesunięcia gap przy pomocy następującego równania:

$$\Delta\alpha = \arctan\left(\frac{gap}{f}\right) \quad (1)$$

gdzie f – ogniskowa w punktach obrazu, gap – wyliczone przesunięcie histogramów.

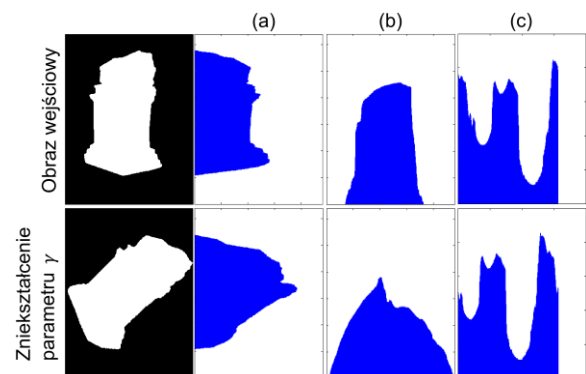


Rys. 2. Wizualizacja wpływu zmiany danych parametrów na obraz reprojekcji. W kolumnach: (a) – obraz wejściowy, (b) – reprojekcja ze zniekształceniem w kącie α , (c) – reprojekcja ze zniekształceniem w kącie β , (d) – reprojekcja ze zniekształceniem w ogniskowej f .

W przypadku wyznaczenia korekty rotacji $\Delta\beta$ i-tego obrazu wejściowego postępujemy analogicznie z tą różnicą, że do obliczeń wykorzystywane są poziome histogramy.

2.2. Estymacja korekty błędu kąta γ

W celu wyznaczenia obrotu i-tej kamery w osi γ nie jest już możliwe wykorzystanie danych histogramów tak jak to było w przypadku kątów z punktu 2.1, gdyż jak łatwo zauważyć na rysunku 3, rozłożenie energii histogramów znacznie się zmieniło. Widoczne jest, że zmiana w kształcie poziomych i pionowych histogramów zmieniła się znacząco przez co niemożliwe jest zastosowanie przedstawionego już podejścia. Z tego powodu konieczne jest zastosowanie histogramów radialnych (Rys.3c), aby umożliwić wyznaczenie popraw-



Rys. 3. Wizualizacja zmiany kształtu histogramów obrazu reprojekcji wynikająca z obrotu wokół osi γ dla i-tego obrazu wejściowego. (a) – histogramy pionowe, (b) – histogramy poziome, (c) – histogramy radialne wyznaczone z funkcji odległości $Distance(angle[rad])$.

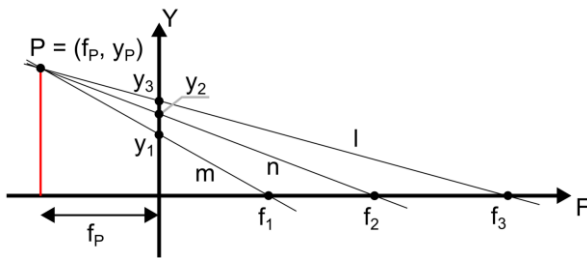
ki $\Delta\gamma$ dla i -tego obrazu wejściowego. Dzięki temu, że histogram radialny pozostaje niezmienny przy rotacji γ , zaś rotacja widoczna jest jako przesunięcie tychże histogramów, możliwe jest zastosowanie podejścia podobnego do tego przedstawionego w punkcie 2.1. Histogram radialny wyznaczany jest jako największa odległość konturu obiektu od środka optycznego obrazu (najczęściej oznaczonego jako punkt (C_x, C_y) w macierzy parametrów wewnętrznych) liczona w punktach obrazu w każdym kierunku od środka optycznego danego obrazu wejściowego. W eksperymentach przedstawionych w niniejszej pracy kierunki zostały skwantowane do 360 przedziałów.

Od tego momentu procedura wyznaczenia poprawki $\Delta\gamma$ dla i -tego obrazu wejściowego postępuje analogicznie do podejścia przedstawionego w punkcie 2.1. Liczona jest korelacja między histogramami radialnymi obrazów wejściowego oraz reprojekcji. Przesunięcie histogramów gap jest obrotem $\Delta\gamma$.

2.3. Korekcja ogniskowej f

W celu wyznaczenia korekty Δf autorzy opracowali następujące podejście 2-krokowe.

Po pierwsze, wyznaczane są histogramy radialne dla obrazów wejściowego oraz reprojekcji do płaszczyzny obrazu (wykorzystując początkową ogniskową f_1). Następnie obliczane są ich średnie obcinane (lub inaczej *średnie trymowane*), aby ustalić średnią odległość punktów $|y_1 - y_3|$ między histogramami. Jest to niezbędne ze względu na to, że korekcja wykonywana jest jeszcze w trakcie procesu rekonstrukcji wokselsej, więc część wokseli nie została jeszcze usunięta ze względu na efekt paralaksy i wpływa na wartość *overflow* (8) obrazu reprojekcji. Wyznaczone średnie trymowane y_1, y_3 zostały pokazane na Rysunku 4, przedstawiającym model kamery otworkowej z uwzględnieniem płaszczyzny obrazu Y oraz osi optycznej F .



Rys. 4. Wizualizacja 3 linii reprojekcji m, n, l rzutujących punkt P na płaszczyznę obrazu z uwzględnieniem 3 ogniskowych f_1, f_2, f_3 . Wizualizacja układu równań wykorzystanych do obliczenia poprawionej ogniskowej f_3 .

Po drugie, poprawka Δf obliczana poprzez estymację ogniskowej f_3 zgodnie z następującymi równaniami:

$$\begin{cases} m: y_P = \frac{y_1}{f_1} f_P + y_1 \\ n: y_P = \frac{y_2}{f_2} f_P + y_2 \end{cases} \quad (2)$$

Wykorzystując równania prostych rzutowania (2) wprowadzamy wzór na pozycję punktu P :

$$f_P = \frac{(y_1 - y_2)f_1 f_2}{(y_1 f_2 - y_2 f_1)}, \quad (3)$$

gdzie

$$f_2 = f_1 \cdot \frac{y_1 + y_3}{2y_1}. \quad (4)$$

Równanie (4) opisuje ogniskową pomocniczą f_2 wykorzystaną do wygenerowania próbnego obrazu reprojekcji, dla którego liczona jest średnia trymowana y_2 z odpowiadającego mu histogramu radialnego. Na podstawie tych wyników wyznaczana jest ostatnia prosta rzutowania:

$$l: 0 = \frac{y_P - y_3}{f_P} f_3 + y_3, \quad (5)$$

Wykorzystywana do wyznaczenia ostatecznej wartości ogniskowej f_3 z następującego równania:

$$f_3 = \frac{-y_3 f_P}{y_P - y_3}. \quad (6)$$

3. EKSPERYMENTY

Eksperymenty przeprowadzane były na modelu testowym *Temple* pochodzącym ze zbioru danych *Middlebury* [6]. Wyniki zostały uśrednione po 47 widokach, z których tworzony był powyższy model. Korekcja każdego błędu parametrów (parametrów rotacji i ogniskowej) były testowane oddzielnie.

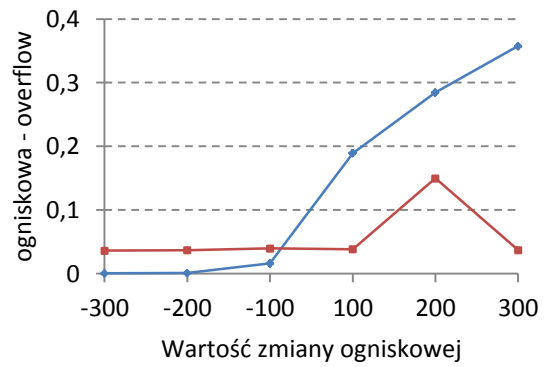
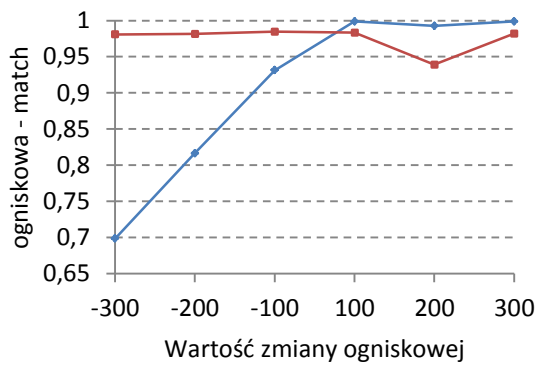
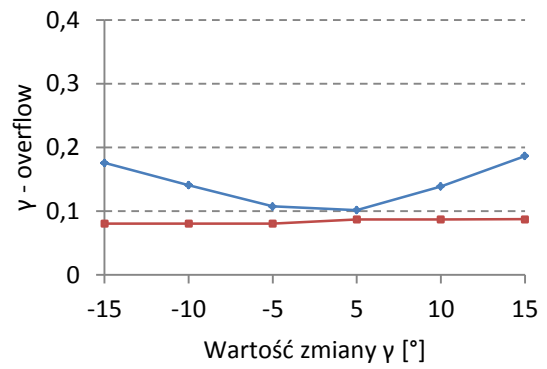
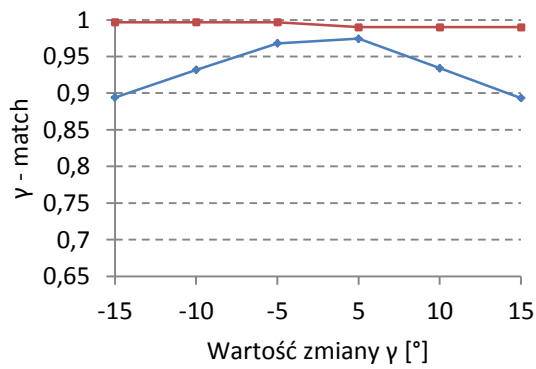
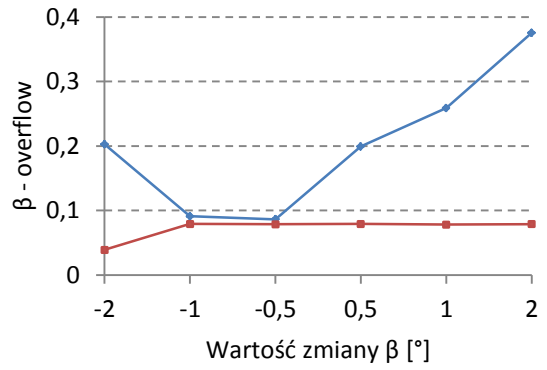
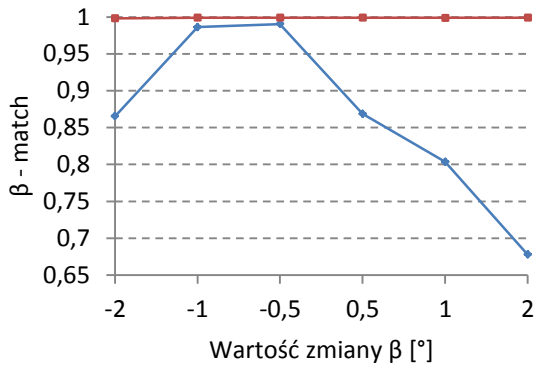
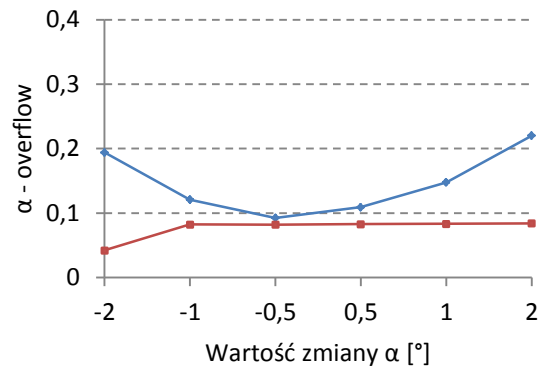
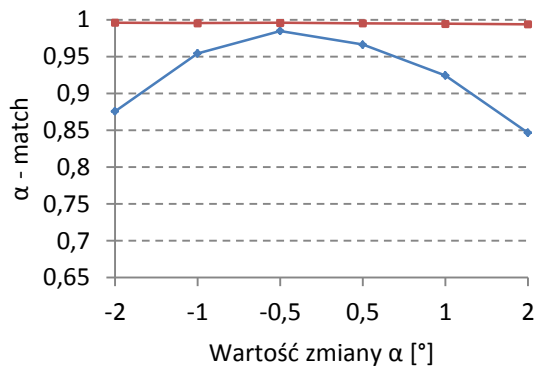
Najpierw, z niezniekształconych parametrów wejściowych zostały wybrane niektóre z nich i zmodyfikowane o znaną wartość. Następnie wykonywana jest rekonstrukcja wokselsej, która generuje zniekształcony model. W kolejnym kroku, wykorzystywane jest podejście zaprezentowane w sekcji 2 do estymacji korekcji parametrów. Na koniec, przeprowadzana jest ocena korekcji parametrów przez porównanie obrazu reprojekcji modelu z obrazem referencyjnym (obrazem wejściowym). W tym celu wykorzystane zostały dwie metryki:

$$match = \frac{mutual_area}{reference_area}, \quad (7)$$

$$overflow = \frac{reproj_area}{reproj_area + mutual_area}, \quad (8)$$

gdzie *reference_area* jest obszarem modelu w obrazie referencyjnym, *mutual_area* jest obszarem wspólnym obu obrazów (obrazu referencyjnego i obrazu reprojekcji), *reproj_area* jest obszarem obrazu reprojekcji, który nie należy do części wspólnej z obrazem referencyjnym. Najlepszy rezultat korekcji uzyskany jest wtedy, kiedy metryki *match* i *overflow* osiągną odpowiednio wartość 1 i 0.

Warto nadmienić, że przedstawione metryki zostały wprowadzone, ponieważ powszechnie stosowane metody porównywania zrekonstruowanego modelu



Rys. 5. Wyniki eksperymentów dla wybranych parametrów. Kolorem niebieskim zostały oznaczone wyniki po wprowadzeniu modyfikacji parametrów o pewną wartość. Kolorem czerwonym oznaczono wyniki po przeprowadzeniu korekcji parametrów.

z modelem referencyjnym nie działają przed osiągnięciem końca procesu rekonstrukcji.

Uzyskanie idealnej wartości metryki $match = 1$ jest zwykle niemożliwe, ponieważ dysponujemy ograniczoną liczbą wokseli w przestrzeni modelu, która jest mniejsza niż rozdzielczość obrazu. Oznacza to, że pojedynczy woxel składa się z więcej niż jednego punktu obrazu. Można to interpretować jako trójwymiarowy błąd kwantyzacji. Z drugiej strony, idealna wartość metryki $overflow = 0$ jest również nieosiągalna, ponieważ jest ona obliczana przed procesem rekonstrukcji, a więc przed usunięciem nadmiarowych wokseli. Jednakże, wartość metryki $overflow$ spada z każdym obrazem wejściowym dodanym do procesu rekonstrukcji.

Ponadto, przeprowadzono dodatkowe eksperymenty dotyczące łączenia błędów (np. jednocześnie błąd w kątach α i β). Udowodniono, że kształt proponowanych histogramów dla poszczególnych kątów jest niezależny od rotacji. Histogramy obrazów w przypadku rotacji obserwowane są jako przesunięcie. Jednakże, różne kombinacje błędów parametrów wewnętrznych oraz błędy parametrów zewnętrznych wymagają dodatkowych badań.

4. PODSUMOWANIE

W niniejszym artykule przedstawiono metodę poprawy jakości trójwymiarowego modelu wokselowego, opartą na histogramach obrazów reprojekcji modelu. W trakcie przeprowadzanych badań, autorzy zaobserwowali podatność rekonstrukcji wokselowej nawet na najmniejsze zmiany parametrów kamery, co przekładało się na uzyskany model.

Na podstawie wyników eksperymentów, autorzy zaproponowali nieoptymalizacyjne podejście do poprawy jakości rekonstrukcji wokselowej. Proponowana metoda wykorzystuje informacje, że dane wejściowe dla rekonstrukcji wolumetrycznej mogą zawierać pewne małe błędy, które mogą zostać poprawione bez potrzeby stosowania skomplikowanych metod. Do oceny zaprezentowanego rozwiązania wykorzystano metryki $match$ i $overflow$, które pokazały wysoką efektywność metody.

Kombinacja indywidualnych błędów rotacji przedstawiona została na rys. 5, gdzie widać skuteczność proponowanej metody. Jednakże, badania dotyczące wszystkich możliwych kombinacji wymagają dodatkowej pracy i oddzielnej analizy. Dodatkowym problemem jest poprawa parametrów macierzy translacji (X, Y, Z), która jest jeszcze bardziej złożona, ze względu na zmianę perspektywy w obrazie reprojekcji i również wymaga dodatkowych badań.

SPIS LITERATURY

[1] D.G. Lowe, "Object recognition from local scale-invariant features", *Proceedings of the International Conference on Computer Vision*, str. 1150–1157, 1999
[2] K. Levenberg, "A Method for the Solution of Certain Non-Linear Problems in Least Squares", *Quarterly of Applied Mathematics*: str. 164–168, 1944
[3] A.O. Balan, "Voxel Carving and Coloring – Constructing a 3D Model of an Object from 2D Images", Brown University, Grudzień 2003

[4] S. Maćkowiak, P. Gardziński, Ł. Kamiński, K. Kowalak, "Human Activity Recognition in Multiview Video", *11th IEEE International Conference on Advanced Video and Signal-Based Surveillance*, South Korea, str. 148–153, 2014
[5] Z. Zhang, "A flexible new technique for camera calibration", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11): str. 1330–1334, 2000
[6] S.M. Seitz, B. Curless, J. Diebel, D. Scharstein, R. Szeliski, "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms", *IEEE Conf. on Computer Vision and Pattern Recognition*, str. 519–528, 2006
[7] B. Triggs, P. McLauchlan, R. Hartley, A. Fitzgibbon, "Bundle Adjustment – A Modern Synthesis", *ICCV'99: Proc. of the International Workshop on Vision Algorithms*, Springer-Verlag, str. 298–372, 1999
[8] B. Goldlücke, M. Magnor, "Real-time, free-viewpoint video rendering from volumetric geometry", *Proc. Visual Computation and Image Processing (VCIP)*, Lugano, Switzerland, 5150(2): str. 1152–1158, 2003
[9] H. Baker, "Three-dimensional modelling", *Proc. of the 5th International Joint Conference on Artificial Intelligence*, Cambridge, MA, USA, str. 649–655, 1977
[10] J. Xiao, J. Chen, D.-Y. Yeung i L. Quan, "Learning two-view stereo matching", *In Proceedings of the 10th European Conference on Computer Vision: Part III, ECCV '08*, Springer-Verlag, Berlin, Heidelberg, str. 15–27, 2008
[11] N. Snavely, S. Seitz, i R. Szeliski, "Photo tourism: exploring photo collections in 3d", *In ACM Transactions on Graphics (TOG)*, volume 25, ACM, str. 835–846, 2006
[12] N. Snavely, S. Seitz, i R. Szeliski, "Modeling the world from internet photo collections", *International Journal of Computer Vision*, 80(2): str. 189–210, 2008