

# Wykorzystanie otwartego oprogramowania do rekonstrukcji stanowisk pracy

Andrzej Grabowski

Centralny Instytut Ochrony Pracy, Państwowy Instytut Badawczy

**Streszczenie:** Systemy wizyjne są od wielu lat stosowane do rekonstrukcji 3D rzeczywistych obiektów różnego typu. Zwykle jednak ich zastosowanie wiąże się z pewnymi ograniczeniami, dotyczącymi m.in. rozmiaru skanowanych obiektów lub warunków oświetlenia. W artykule przedstawiono uniwersalną, a zarazem niewymagającą dużych nakładów finansowych, metodę rekonstrukcji stanowisk pracy opartą na oprogramowaniu o otwartym źródle takim jak OpenCV i PCL (Points Cloud Library). Opracowana metoda daje dużą dokładność rekonstrukcji, może być wykorzystana zarówno do małych jak i dużych obiektów oraz jest w dużej mierze niezależna od zewnętrznego oświetlenia - działa w zamkniętych pomieszczeniach oraz na zewnątrz. Podstawowym celem przedstawionej metody jest rekonstrukcja 3D stanowisk pracy w celu testowania wizyjnych systemów bezpieczeństwa z wykorzystaniem technik rzeczywistości mieszanej.

**Słowa kluczowe:** systemy wizyjne, rekonstrukcja 3D, rzeczywistość mieszana

## 1. Wprowadzenie

Zagadnienie prawidłowego określenia relacji przestrzennych pomiędzy przedmiotami jest tematyką badań, która rozwija się bardzo intensywnie w ostatnich latach. Rozwój ten jest związany głównie z szybko rosnącą mocą obliczeniową komputerów jak również powstawaniem coraz bardziej wyrafinowanych metod sztucznej inteligencji. Wszystkie współczesne metody pomiaru położenia przedmiotów w trójwymiarowej przestrzeni możemy podzielić na aktywne i pasywne. Metody aktywne związane są z koniecznością emisji promieniowania elektromagnetycznego o różnych zakresach fal bądź też emisji ultradźwięków. Najczęściej wykorzystywane jest promieniowanie laserowe, wówczas odległość przedmiotu od źródła promieniowania określana jest na podstawie pomiaru czasu nadejścia wiązki odbitej (skanery laserowe typu *time of flight*), pomiaru zmiany fazy odbitej wiązki laserowej (skanery typu AMCW – *Amplitude Modulated Continuous Wave*) lub analizy położenia oświetlonych punktów (triangulacja) [1-3]. Ciekawą techniką eliminującą konieczność wykorzystania promieniowania laserowego jest wykorzystanie tzw. światła strukturalnego. Metoda ta polega na wyświetlaniu różnego typu deseni (najczęściej składających się z umieszczonych naprzemiennie czarnych i białych pasków), a następnie badaniu zniekształconego w ten sposób obrazu [4].

Do wad metod aktywnych należy zaliczyć powodowane przez nie zagrożenie zdrowia ludzi (promieniowanie laserowe) lub konieczność zastosowania specjalnych warunków oświetlenia (światło strukturalne). Rozwiązaniem tych trudności jest wyświetlenie odpowiedniego desenia w podczerwieni. Podejście takie zastosowano m.in. w Microsoft Kinect.

Wzajemnie wymienionych wad nie mają również metody pasywne, które w przeciwieństwie do metod aktywnych, nie wymagają własnego źródła promieniowania – wykorzystywane jest wyłącznie oświetlenie zewnętrzne [5]. Charakterystyczną cechą większości metod pasywnych, takich jak strukturalna analiza cienia lub metody wykorzystujące zjawisko rozogniskowania [6, 7], jest długi czas działania uniemożliwiający ich zastosowanie do określenia położenia szybko przemieszczających się przedmiotów.

Najpowszechniejsze metody pasywne to metody stereoskopowe [8], w których do wnioskowania o właściwościach trójwymiarowych obserwowanej przestrzeni wykorzystuje się obrazy zarejestrowane przez różne (dwie lub więcej) kamery. Powszechność tych metod jest związana z faktem, że widzeniem stereoskopowym obdarzeni są zarówno ludzie, jak też i zwierzęta. Metody stereowizyjne działają na dwóch obrazach, uchwyconych przez system akwizycyjny w tej samej chwili. Następnie informacja dwuwymiarowa zawarta w tych obrazach zostaje przekształcona w informację trójwymiarową, mówiącą o przestrzennym układzie obserwowanych obiektów [6, 9-11].

W najprostszym przypadku układ wizyjny składa się z dwóch kamer przesuniętych względem siebie o odległość  $D$  w jednej osi. Wówczas problem analizy stereo redukuje się do problemu odnalezienia odpowiadających sobie punktów na obrazie z lewej kamery i obrazie z prawej kamery. Jednakże przed rozpoczęciem właściwej analizy konieczne jest programowe usunięcie zniekształceń związanych z niedoskonałościami soczewek znajdujących się w obiektywach kamer - dotyczy to zwłaszcza soczewek o małej wartości ogniskowej.

Ponieważ w najprostszym przypadku kamery nie są względem siebie obrócone oraz przesunięte w pionie algorytm polega na porównaniu otoczenia punktu  $P_1=(x_1,y)$  z obrazem z lewej kamery z otoczeniem punktów znajdujących się w tym samym wierszu w obrazie z drugiej kamery. Punkt  $P_2=(x_2,y)$ , dla którego obserwowane różnice są najmniejsze jest właśnie punktem odpowiadającym punktowi  $P_1$ . Znając przesunięcie w poziomie pomiędzy tymi dwoma punktami  $d = |x_1 - x_2|$  można wyznaczyć odległość tego punktu od kamer:

$$z = f_k \cdot D / d \quad (1)$$

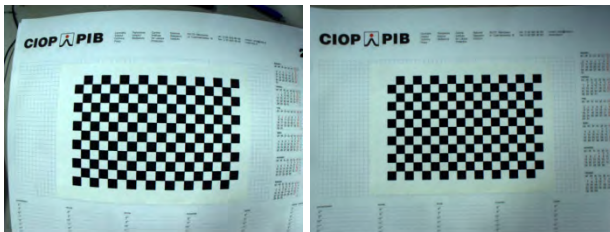
oraz wartości  $x, y$

$$x = (x_1 - c_x) f_x z, \quad y = (y_1 - c_y) f_y z \quad (2)$$

gdzie  $f_x, f_y$  to parametry określające właściwości obiektywu kamery, a  $c_x, c_y$  to współrzędne punktu znajdującego się w środku obrazu. Po wykonaniu obliczeń dla wszystkich punktów obrazu uzyskujemy mapę głębi.

## 2. Procedura rekonstrukcji 3D

Do odtworzenia informacji o położeniu punktów w trójwymiarowej przestrzeni wykorzystano analizę stereo. Jednakże przed przeprowadzeniem właściwej analizy konieczne było usunięcie zniekształceń wynikających z niedoskonałości obiektywu (rys. 1). W celu zachowania dużej dokładności pomiaru, przy jednoczesnym uniknięciu problemów z przesłanianiem (*occlusion*) zastosowano szereg ujęć z kamer o niewielkim przesunięciu w osi  $X$  (rzędu kilku centymetrów). Ponadto podejście takie zwiększa szanse na prawidłowe wykrycie par odpowiadających sobie punktów ze względu na stosunkowo niewielkie różnice w obrazach. Do przygotowania zdjęć wykorzystano zestaw siedmiu kamer internetowych zamocowanych na jednej osi (rys. 2) oraz aparat fotograficzny Canon 550D z osiemnastomegapikselową matrycą zamontowany na suwnicy (rys. 2). W pierwszym przypadku możliwe było jednoczesne wykonanie wszystkich zdjęć (co jest przydatne gdy mamy do czynienia z ruchomymi obiektami lub zmiennym oświetleniem), w drugim zdjęcia wykonywane były sekwencyjnie, tzn. po wykonaniu zdjęcia konieczne było przesunięcie aparatu na suwnicy.



**Rys. 1.** Obraz oryginalny (z lewej) oraz ten sam obraz po usunięciu zniekształceń wynikających z niedoskonałości obiektywu (z prawej). Do kalibracji kamer oraz usunięcia zniekształceń wykorzystano bibliotekę OpenCV

**Fig. 1.** Original image (left) and the same image after the removal of distortions (right). For calibration of cameras and to remove distortions OpenCV library was used

Procedura analizy obrazów wyglądała następująco. Przyjmując, że obraz centralny ma numer  $c$ , pierwszym krokiem było odnalezienie odpowiadających punktów znajdujących się w obrazie  $c+1$ . Następnie dla wszystkich punktów w obrazie  $c+1$  wyszukano punkty odpowiadające w obrazie  $c+2$ . Operację tą powtarzano, aż do ostatniego zdjęcia. Sumując przesunięcia dla wszystkich kolejnych

par obrazów uzyskano mapę przesunięć dla obrazu  $c$  oraz ostatniego obrazu  $n$ . W analogiczny sposób została wyznaczona mapa przesunięć dla obrazu środkowego i pierwszego. W rezultacie otrzymano zbiory punktów odpowiadającym punktom w obrazie środkowym ( $c$ ) znajdujących się w obrazach najbardziej wysuniętych na lewo i na prawo (rys. 3).

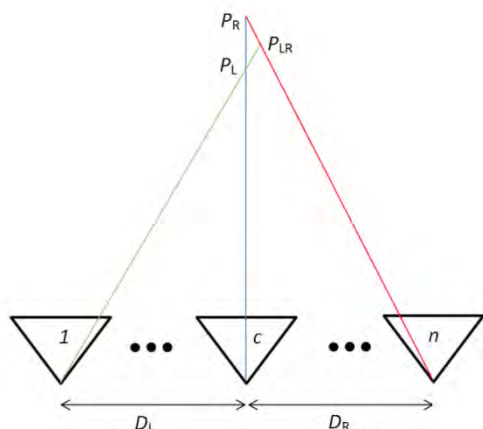
Ze wzorów (1) i (2) wyznaczane jest położenie punktu  $P_L$  (będącego efektem porównania obrazu środkowego i najbardziej wysuniętego w lewo) i  $P_R$  (będącego efektem porównania obrazu środkowego i najbardziej wysuniętego w prawo). Kolejną zaletą wykorzystanej metody jest możliwość weryfikacji par odpowiadających sobie punktów, jeżeli odległość pomiędzy punktami  $P_L$  i  $P_R$  jest zbyt duża punkt ten nie jest uwzględniany w dalszej analizie. W przeciwnym przypadku położenie punktu wyznaczane jest na podstawie przesunięcia pomiędzy odpowiadającymi sobie punktami w obu skrajnych obrazach (lewym i prawym). Współrzędne punktu  $P_{LR}$  wyliczamy wstawiając w równaniach (1) i (2)  $D = D_L + D_R$ . Ze względu na dużą wartość sumy  $D_L + D_R$  możliwe jest uzyskanie dużej precyzji pomiaru. Dla  $D_L + D_R = 0,5$  m oraz zdjęć wykonanych aparatem Canon 550D (rozdzielczość 5184 x 3456 px) z obiektywem o ogniskowej  $f=18$  mm uzyskujemy dokładność pomiaru 0,4 mm (w osi  $Z$ ) dla punktów znajdujących się w odległości jednego metra, oraz 1,7 mm dla punktów znajdujących się w odległości dwóch metrów. Dalsze zwiększenie dokładności jest możliwe po zastosowaniu dłuższej ogniskowej.

Poprzez wyświetlenie odpowiednich deseni (rys. 4) możliwe jest zwiększenie liczby poprawnie wyznaczonych par odpowiadających punktów, co ma szczególnie istotne znaczenie w przypadku dużych, jednorodnych powierzchni. Ponieważ informacja o kolorze punktu pobierana jest tylko z środkowego obrazu, tylko dla tego obrazu konieczne jest wykonanie dwóch zdjęć: z deseniem oraz w normalnych warunkach oświetlenia.



**Rys. 2.** Aparat fotograficzny Canon EOS 550D zamontowany na suwnicy (z lewej) oraz układ siedmiu kamer internetowych na statywie (z prawej)

**Fig. 2.** Camera Canon EOS 550D mounted on a crane (left) and arrangement of seven web cameras on a tripod (right)



**Rys. 3.** Wyznaczenie położenia punktu P w trójwymiarowej przestrzeni opiera się na wykorzystaniu informacji o odpowiadających sobie punktach z kamery środkowej (c) i pierwszej oraz z kamery środkowej i ostatniej (n). Punkt  $P_L$  to efekt porównania obrazów c i 1, punkt  $P_R$  to efekt porównania obrazów c i n, natomiast największą precyzję wyznaczenia położenia punktu P to porównanie obrazów 1 i n ( $P = P_{LR}$ )

**Fig. 3.** Determination of the position of point P in 3D space of based on the use of information on corresponding points of the center (c), the first (1) and final (n) camera. The point  $P_L$  is the result of comparing images 1 and c, point  $P_R$  is a result of a comparison of images c and n. The greatest precision in position of point P is achieved as result of comparison of images 1 and n ( $P = P_{LR}$ )



**Rys. 4.** Przykłady deseni, których wyświetlenie zwiększa efektywność działania analizy stereo, desień stochastyczny (z góry) oraz desień składający się z pionowych pasów (z dołu)

**Fig. 4.** Examples of patterns, which allow to increase in efficiency of stereo analysis: stochastic pattern (up) and a pattern consisting of vertical bars (down)

Do wyznaczenia par odpowiadających sobie punktów w kolejnych obrazach ( $i$ ) oraz ( $i+1$ ) zastosowano metodę *graph-cut* [12], której implementacja znajduje się w bibliotece OpenCV [13]. Jest to globalna metoda analizy stereo o dużej złożoności obliczeniowej (nie nadaje się do zastosowania w systemach czasu rzeczywistego), ale charakteryzująca wysokim odsetkiem poprawnie rozpoznanych par odpowiadających sobie pikseli. Jednymi z ważniejszych zalet tej metody jest poprawne rozpoznawanie krawędzi, mała liczba parametrów kontrolnych oraz niewielka zależność uzyskiwanych wyników od wybranych wartości parametrów kontrolnych.

W przypadku metody *graph-cut* problem analizy stereo dla dwóch obrazów jest formułowany jako problem przypisania odpowiednich etykiet dla poszczególnych pikseli (przykład przypisania etykiet pikselom na podstawie ich jasności przedstawiono na rys.5). Wartość etykiety to przesunięcie  $d$  pomiędzy parą odpowiadających sobie punktów (wzór (1)). W celu obliczenia etykiet  $d_p$  dla wszystkich punktów  $p$ , przyjmuje się dwa założenia:

- (a) odpowiadające sobie punkty w lewym i prawym obrazie mają podobną jasność,
- (b) mapa głębi  $d$  jest gładka, tzn. sąsiednie piksele  $p$  i  $q$  powinny mieć podobne wartości przesunięć  $d_p$  i  $d_q$ .

Powyższe założenia pozwalają na sformułowanie następującej funkcji energii (dzięki czemu problem analizy stereo sprowadza się do problemu minimalizacji energii):

$$E(f) = \sum_p D_p(d_p) + \sum_{p,a} V(p,q)(d_p,d_a) \quad (3)$$

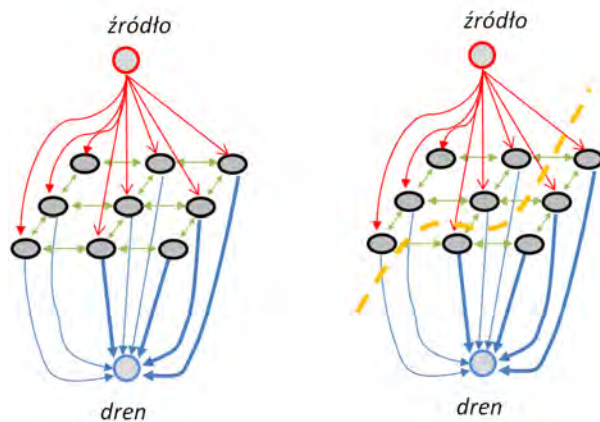
gdzie sumowanie odbywa się po wszystkich pikselach, wartość funkcji  $D_p$  to kwadrat różnicy jasności odpowiadających sobie pikseli, a wartość funkcji  $V$  rośnie wraz ze wzrostem modułu różnicy przesunięć dla pikseli  $p$  i  $q$ ,  $V \sim |d_p - d_q|$ . Dzięki takiej postaci funkcji  $V$  możliwe jest wyznaczenie globalnego minimum funkcji energii  $E(f)$ . Jednakże przed wyznaczeniem globalnego minimum funkcji energii  $E(f)$  konieczne jest zbudowanie odpowiedniego grafu skierowanego łączącego sąsiadujące piksele.



**Rys. 5.** Przykład przypisania etykiet pikselom na podstawie ich jasności. Grubą linią oddzielono grupy pikseli o takiej samej etykiecie

**Fig. 5.** Example of assigning labels to pixels based on their brightness. Thick line separated group of pixels with the same label

Graf skierowany z wagami  $G = (V, E)$  składa się ze zbioru węzłów  $V$  oraz zbioru skierowanych połączeń pomiędzy nimi  $E$ . Węzły grafu odpowiadają pikselom obrazu. Dodatkowo graf zawiera również dwa węzły terminalne, są to węzły specjalnego przeznaczenia, z którymi połączone są wszystkie pozostałe węzły (piksele). Pierwszy z węzłów terminalnych to źródło  $s$ , z którego wyprowadzone są połączenia do wszystkich pikseli. Drugi to dren  $t$ , do którego prowadzą połączenia od wszystkich pikseli. Wartość tych węzłów to etykieta pikseli, natomiast waga połączenia pikseli do węzła terminalnego związana jest z kosztem przypisania pikselowi określonej etykiety (jest to wartość funkcji  $D_p$  ze wzoru (3)). W grafie występują też połączenia pomiędzy sąsiednimi pikselami, których waga odpowiada funkcji  $V$  z równania (3). Przykład grafu dla obrazu składającego się z dziewięciu pikseli ( $3 \times 3$ ) pokazany jest na rys. 6. Dalsza procedura polega na znalezieniu minimalnej liczby cięć, tak aby zminimalizować wartość funkcji energii  $E(f)$ , a tym samym wyznaczyć również grupy pikseli o takiej samej wartości etykiety (porównaj rys. 5).



**Rys. 6.** Graf dla obrazu składającego się z  $3 \times 3$  pikseli oraz ten sam graf wraz cięciami, które minimalizuje funkcję energii  $E(f)$ . Koszt (waga) połączeń jest oznaczona przez ich grubość

**Fig. 6.** Graph for an image consisting of  $3 \times 3$  pixels (left) and the same graph with cutting (right), which minimizes the energy function  $E(f)$ . The cost (weight) of connections is indicated by their thickness

Wykorzystanie serii zdjęć pochodzących tylko z jednego punktu jest niewystarczające do odtworzenia informacji o całej scenie. Konieczne jest zatem wykonanie serii pomiarów głębi tak, aby zebrać informacje o przestrzennym rozmieszczeniu wszystkich obiektów. Z każdego pomiaru uzyskujemy jedną chmurę punktów. Uzyskane chmury należy następnie połączyć po nałożeniu na punkty

odpowiednich transformacji wynikających z położenia i orientacji kamery.

Ponieważ rekonstrukcja 3D jest oparta na systemie wizyjnym możliwe jest stosunkowo proste wyznaczenie macierzy transformacji dla chmury punktów korzystając ze zbioru odpowiadających sobie punktów charakterystycznych (rys. 7). Do wyszukiwania i dopasowania punktów charakterystycznych wykorzystano metodę SURF (*Speeded Up Robust Feature*) [14], która jest częściowo oparta na metodzie SIFT (*Scale-Invariant Feature Transform*), ale jest od niej wielokrotnie szybsza i bardziej odporna na transformacje obrazu. Metoda SURF wykorzystuje w swoim działaniu analizę falkową przeprowadzaną w okolicy punktu charakterystycznego.

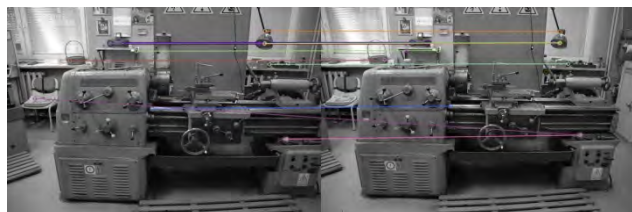
Dysponując zbiorem odpowiadających sobie punktów charakterystycznych dla dwóch obrazów oraz położeniem tych punktów w trójwymiarowej przestrzeni (każdemu pikselowi obrazu przypisane są współrzędne 3D uzyskane dzięki analizie stereo) możemy przystąpić do wyznaczenia macierzy transformacji pomiędzy lokalnymi układami współrzędnych dwóch chmur punktów. Po przekształceniu punktów z jednej z chmur wszystkie punkty z obu chmur będą znajdowały się w tym samym układzie współrzędnych.

Macierz transformacji posiada sześć stopni swobody, trzy związane z translacją oraz trzy związane z rotacją. Korzystając ze symetrii problemu możemy zadanie wyznaczenia optymalnej macierzy transformacji znacznie uprościć przyjmując za postać funkcji, której wartość chcemy zminimalizować, sumę różnic odległości pomiędzy odpowiadającymi sobie punktami z dwóch różnych chmur i średnią odległością:  $f = \sum_i |d_i - \langle d \rangle|$ . Wówczas konieczne jest znalezienie wartości tylko trzech parametrów opisujących rotację, bowiem dla dwóch zbiorów punktów, które są tylko przesunięte względem siebie odległości pomiędzy odpowiadającymi sobie punktami powinny być takie same. Po wyznaczeniu macierzy rotacji problem wyznaczenia wartości parametrów związanych z translacją sprowadza się do prostego obliczenia średniego przesunięcia niezależnie dla wszystkich trzech osi.

Mając dwie chmury znajdujące się w tym samym układzie współrzędnych możemy dodać je do siebie, pamiętając przy tym, że możemy mieć do czynienia z dużą liczbą punktów, które po połączeniu chmur punktów będą znajdowały się bardzo blisko siebie (żeby dobrze połączyć dwie chmury punktów powinny one się częściowo pokrywać). W celu przyspieszenia operacji łączenia chmur, jak również filtrowania nadmiarowych punktów (ostateczna

chmura punktów nie powinna być zbyt duża) wykorzystano drzewo ósemkowe z minimalnym rozmiarem komórki 1 mm, tzn. że w jednym milimetrze sześciennym znajdować się może co najwyżej jeden punkt. Drzewo ósemkowe (ang. *octree*) to stosowana w grafice komputerowej struktura danych będąca drzewem, używana do przestrzennego podziału trójwymiarowej przestrzeni na mniejsze, regularne części. Konstrukcja drzewa ósemkowego polega na otoczeniu całości sceny trójwymiarowej sześcianem, który następnie dzielony jest na osiem mniejszych, a następnie każdy z nich na osiem kolejnych itd. - proces podziału ma charakter rekurencyjny.

Po sprowadzeniu pierwszych dwóch chmur punktów do jednego układu współrzędnych w analogiczny sposób postępujemy z pozostałymi chmurami punktów uzyskując w ten sposób pełną informację o geometrii rekonstruowanego miejsca pracy.



**Rys. 7.** Punkty charakterystyczne (zaznaczone kolorowymi okręgami) na dwóch różnych obrazach. Tylko niektóre odpowiadające sobie punkty charakterystyczne zostały ze sobą połączone liniami, ze względu na zachowanie przejrzystości ilustracji

**Fig. 7.** The same key points (marked with colored circles) on two different images

Zbiór wszystkich punktów jest podstawą do przygotowania siatki (*mesh*) komputerowego modelu 3D. Najbliższe punkty (wierzchołki) łączone są ze sobą w trójkąty – figury płaskie posiadające najmniejszą liczbę wierzchołków niezbędną do zdefiniowania powierzchni. Do przygotowania siatki połączonych wierzchołków oraz przeprowadzenia procedury usuwania szumów i wygładzania powierzchni można użyć biblioteki PCL (Points Cloud Library).

Ponieważ mamy informację o położeniu w przestrzeni wszystkich wierzchołków, jak również informację o ich kolorze, możliwe jest przygotowanie fotorealistycznego komputerowego modelu 3D. Model ten może być dowolnie zmodyfikowany (np. możliwe jest nałożenie na podłogę dowolnego desenia pasywnego tła odniesienia). Możliwe jest też dzięki metodzie śledzenia promieni wykonanie syntetycznych obrazów z różnych ujęć (np. zależnych od rozmieszczenia kamer wizyjnego systemu bezpieczeństwa)

oraz realistyczne uwzględnienie różnych źródeł oraz pojawiania się cieni.

### 3. Wnioski

Zastosowanie bibliotek oprogramowania o otwartym źródle, takich jak OpenCV i PCL (Points Cloud Library) pozwala znacząco skrócić czas na przygotowanie oprogramowania do rekonstrukcji 3D. Wiąże się to jednakże również z pewnymi trudnościami, takimi jak brak pełnej dokumentacji oraz przykładów pokazujących możliwości zastosowania kodu. Często konieczne jest przeprowadzenie analizy działania poszczególnych funkcji. Dobrym przykładem jest funkcja umożliwiająca przeprowadzenie analizy stereo w oparciu o metodę GraphCut znajdującą się w bibliotece OpenCV – funkcja ta w ogóle nie znajduje się w dostępnej dokumentacji a jej implementacja jest dostępna tylko w C. To, że część funkcji ma swoje implementacje tylko w C lub C++ (brak jest odpowiedniej fasady umożliwiającej wywołanie tych funkcji zarówno w C jak i C++) również jest dodatkowym utrudnieniem ze względu na różne struktury danych stosowanych w OpenCV w zależności od języka programowania.

Rozważając wszystkie za i przeciw stosowania otwartego oprogramowania oraz biorąc pod uwagę bardzo niski koszt wysokiej jakości aparatów fotograficznych i kamer, wykorzystanie dostępnego oprogramowania (np. OpenCV i PCL) pozwala na drastyczne zmniejszenie nakładu czasu i środków niezbędnych na przygotowanie własnego systemu do rekonstrukcji 3D.

Opracowana metoda rekonstrukcji 3D zostanie wykorzystana do testowania optoelektronicznych urządzeń ochronnych wyposażonych w układ wizyjny [15]. Zrekonstruowane w wirtualnej rzeczywistości stanowisko pracy, np. wyposażone w robota przemysłowego [16], zostanie wzbogacone o fotorealistyczny komputerowy model pracownika oraz komputerowe modele obiektów pobierczych. Syntetyczne obrazy wygenerowane w odpowiedni sposób za pomocą bezpłatnego oprogramowania do grafiki 3D (Blender) zostaną przesłane do wizyjnego systemu bezpieczeństwa w celu przeprowadzenia m.in. testów funkcjonalnych lub wyznaczenia rozmiaru strefy tolerancji. Takie podejście znacząco może ułatwić i przyspieszyć proces testowania algorytmów analizy obrazu wykorzystywanych w wizyjnych systemach bezpieczeństwa.

### Acknowledgment

This paper has been based on the results of a research task carried out within the scope of the second stage of the National Programme “Improvement of safety and

working conditions” partly supported in 2011–2013 — within the scope of research and development — by the Ministry of Science and Higher Education/National Centre for Research and Development. The Central Institute for Labour Protection – National Research Institute is the Programme’s main co-ordinator

## Bibliografia

1. Y.T.Young: *Handbook of Pattern Recognition and Image Processing: Computer Vision*. Vol. 2. Academic Press (1994).
2. E.J.Lerner: *Laser microscopy opens a new dimension*, Laser Focus World 36, 141-144 (2000).
3. B. Siemiątkowska, A. Zychewicz, M. Gnatowski: *Zastosowanie metod przetwarzania obrazów w segmentacji danych pochodzących z dalmierza laserowego 3D*, Pomiary Automatyka Robotyka, 02.2010 (2009).
4. E.R.Davies: *Machine Vision. Theory, Algorithms, Practicalities* 2nd Edition, Academic Press (1997).
5. N.Yokoya, T.Shakunaga, M.Kanbara: *Passive Range Sensing Techniques: Depth from Images*, IEICE Transactions on Information and Systems, E82-D, 523-533, (1999).
6. R.Klette, K.Schliins, A.Koschan: *Computer Vision. Three-Dimensional Data from Images*, Springer-Verlag (1998)
7. S.Chaudhuri, A.N.Rajagopalan, *Depth From Defocus: A Real Aperture Imaging Approach*, Springer Verlag (1999).
8. L.L.Grewe, A.C.Kak: *Stereo Vision. Handbook of Pattern Recognition and Image Processing: Computer Vision*, Academic Press (1994).
9. D.Scharstein: *View Synthesis Using Stereo Vision*, Springer-Verlag (1999).
10. R.Szeliski, R.Zabih: *An Experimental Comparison of Stereo Algorithms. Microsoft Technical Report*, www.research.microsoft.com/szeliski (2000).
11. E.Truccho, A.Verr: *Introductory Techniques for 3-D Computer Vision*, Prentice-Hall (1998).
12. V. Kolmogorov: *Graph Based Algorithms for Scene Reconstruction from Two or More Views*, PhD thesis, Cornell University, September 2003.
13. G. Bradski, A. Kaehler: *Learning OpenCV: Computer Vision with the OpenCV Library*, O’Reilly, 2008.
14. H.Bay, A.Ess, T.Tuytelaars, L. Gool: *"SURF: Speeded Up Robust Features"*, Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, 346--359, 2008
15. A.Grabowski, R.Kosiński, M.Dźwiarek, *Artificial neural network for worker presence identification in dangerous zones using stereovision*, Pomiary Automatyka Robotyka, 02/2009, 142-151.
16. J. Domińczuk: *Wpływ technik wirtualnych na rozwój automatyzacji i robotyzacji*, Pomiary Automatyka Robotyka, 02/2010. ■

## The use of open source software for the 3D reconstruction of work places

**Abstract:** Vision systems are used for many years to reconstruct the 3D objects of various types. Usually, however, their use is associated with certain limitations, e.g. size of the scanned object or lighting conditions. This paper presents a universal and low-cost method of 3D reconstruction of work places based on open source software such as OpenCV and PCL (Points Cloud Library). The developed method gives high accuracy of reconstruction, can be used for both small and large objects, and is largely independent of ambient light - works indoors and outdoors. Presented method will be used for 3D reconstruction of the work places in order to test the video based safety systems with the help of mixed reality techniques.

**Keywords:** vision systems, 3D reconstruction, mixed reality

### dr inż. Andrzej Grabowski

Zatrudniony w Centralnym Instytucie Ochrony Pracy – Państwowym Instytucie Badawczym na stanowisku kierownika Pracowni Techniki Rzeczywistości Wirtualnej. W swojej pracy zajmuje się m.in. wizyjnymi systemami bezpieczeństwa i tworzeniem aplikacji szkoleniowych wykorzystujących rzeczywistość wirtualną.  
e-mail: angra@ciop.pl

