

**PROTOTYPOWY SYSTEM STEREOWIZYJNY TYPU MACHINE VISION –  
DZIAŁANIE I DOKŁADNOŚĆ**

**PROTOTYPE STEREOVISION SYSTEM OF MACHINE VISION TYPE –  
OPERATION AND ACCURACY**

**Dariusz Rzeszotarski<sup>1</sup>, Piotr Sawicki<sup>2</sup>, Bartosz Ostrowski<sup>1</sup>**

<sup>1</sup>Instytut Elektroniki, Politechnika Łódzka

<sup>2</sup>Zakład Fotogrametrii i Teledetekcji, Uniwersytet Warmińsko-Mazurski w Olsztynie

SŁOWA KLUCZOWE: kamera cyfrowa, system stereowizyjny, dysparycja, dokładność, procesor sygnałowy

STRESZCZENIE: W pracy przedstawiono istotę działania i wstępną ocenę dokładności prototypowego cyfrowego systemu stereowizyjnego, który stanowi podstawę konstrukcji urządzenia wspomagającego osobę niewidomą w percepcji otoczenia (projekt badawczy realizowany w Instytucie Elektroniki Politechniki Łódzkiej). W prototypowej wersji systemu zastosowano 2 kolorowe kamery cyfrowe Flea® firmy Point Grey (rozdzielczość 1024×768, interfejs IEEE 1394), wyposażone w obiektywy Ernitec  $f = 3.5$  mm. Kalibracja kamer została przeprowadzona za pomocą programu wspomaganego biblioteką OpenCV (*Intel® Open Source Computer Vision Library*). Proces automatycznego wyznaczenia gęstej mapy dysparycji (mapy paralaks podłużnych) został wsparty algorytmem dopasowującym obszary (winxwiny - okno korelacji). Zastosowaną miarą podobieństwa jest suma wartości bezwzględnych różnic SAD (*Sum of Absolute Differences*), obliczana techniką „przesuwanych okien” (*sliding window*). Poszukiwanie wiarygodnych minimów miary SAD jest wykonywane w czterech krokach przez znalezienie minimum globalnego  $SAD_{min}$  oraz trzech najmniejszych wartości  $SAD_1 < SAD_2 < SAD_3$ . Za pomocą mapy dysparycji i parametrów orientacji zewnętrznej modelu (macierze przekształcenia  $M_R, M_L$ ) wyznaczane są współrzędne 3D sceny. Algorytm segmentacji sceny 3D realizuje w trybie iteracyjnym detekcję płaszczyzn oraz obiektów „przeszkód”. Procedura akwizycji obrazów cyfrowych, generowanie mapy dysparycji i segmentacji zostały zaimplementowane na platformie procesora sygnałowego DSP (*Digital Signal Processor*). Ocenę dokładności systemu wykonano na podstawie stereopar zdjęć cyfrowych (baza zdjęć  $B = 80.75$  mm) dwóch przestrzennych pól testowych. Współrzędne 3D punktów zostały wyznaczone na podstawie wygenerowanej gęstej mapy dysparycji (okno korelacji wyniosło 25×25 pikseli). Algorytm dysparycji wykrył średnio ca 72 % sygnalizowanych punktów kontrolnych. Średnia odchyłka kwadratowa dla współrzędnych 3D w układzie współrzędnych systemu stereowizyjnego wyniosła  $\Delta_{XYZ} = \pm 39.6$  mm (test 1) oraz  $\Delta_{XYZ} = \pm 124.1$  mm (test 2). Średnia odchyłka kwadratowa dla współrzędnych X, Y, Z po transformacji 3D do układu obiektu wyniosła  $\Delta_{XYZ} = \pm 72$  mm (test 1) oraz  $\Delta_{XYZ} = \pm 74$  mm (test 2).

## 1. WSTĘP

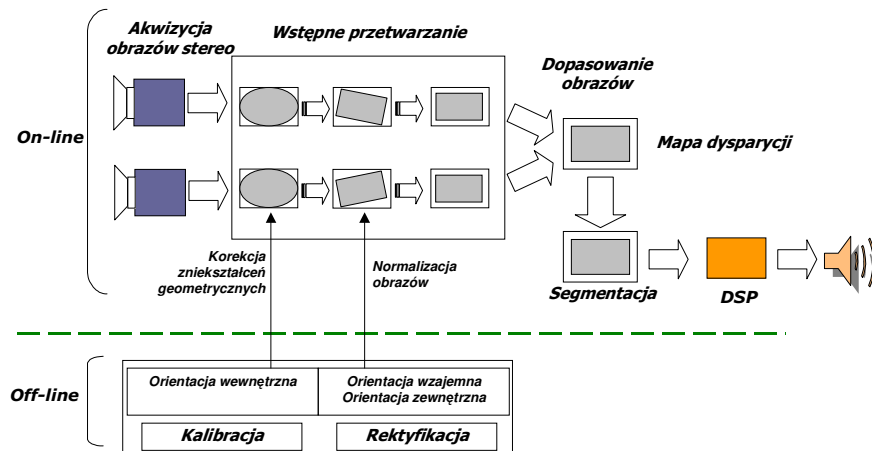
Stereowizja cyfrowa jest współczesną techniką obrazową umożliwiającą rekonstrukcję scen 3D na podstawie obrazów pozyskanych z co najmniej dwóch sensorów optyczno-elektronicznych. Zastosowanie komputerów i procesów sygnałowych umożliwiło w ostatnich latach opracowanie metod szybkiej, działającej w czasie rzeczywistym, automatycznej rekonstrukcji i pozycjonowania obiektów scen 3D. W konsekwencji powstały cyfrowe systemy stereowizyjne pracujące z szybkością do 30 ramek na sekundę dla rozdzielczości zbliżonych do obrazów typu CIF (*Common Intermediate Format*), tj. 352×288 pikseli obrazu (Brown *et al.*, 2003).

Spektrum zastosowań stereowizji cyfrowej obejmuje diagnostykę techniczną i medyczną, wizyjną kontrolę jakości i procesu produkcji w wielu gałęziach przemysłu, autonomiczne poruszanie się robotów, detekcję i śledzenie obiektów w ruchu, kreację rzeczywistości wirtualnej, itd. Stereowizja cyfrowa jest jedną z technik maszynowego widzenia (*machine vision*).

W pracy przedstawiono istotę działania i wstępną ocenę dokładności prototypowego systemu stereowizyjnego, który stanowi moduł akwizycji obrazów cyfrowych w urządzeniu wspomagającym osobę niewidomą w percepcji otoczenia (projekt badawczy MNiSW T11B03827 realizowany w Instytucie Elektroniki Politechniki Łódzkiej).

## 2. BUDOWA I DZIAŁANIE SYSTEMU STEREOWIZYJNEGO

Przenośny elektroniczny system wspomagania osobę niewidomą w samodzielnym poruszaniu się wykorzystuje do rekonstrukcji sceny 3D cyfrowy układ stereowizyjny jako źródło informacji obrazowych, które następnie po przetworzeniu, selekcji i segmentacji tworzą system dźwiękowej reprezentacji obiektów oraz ich otoczenia.



Rys. 1. Schemat blokowy działania prototypowego systemu stereowizyjnego

Prototypowy cyfrowy system stereowizyjny składa się z części sprzętowej (sensory wizyjne i głowica mechaniczna, służąca do zamocowania kamer) oraz opracowanego oprogramowania. Kalibracja i rektyfikacja wzajemnego układu kamer wykonywana jest w trybie *off-line*, natomiast procedura akwizycji i przetwarzania cyfrowych obrazów stereo następuje w trybie *on-line* (Rys. 1).

Moduł wizyjny elektronicznego systemu wspomagania osobę niewidomą realizuje sekwencję 6 procedur obrazowych (opisane zostały w rozdziałach 2.1÷2.7), dla których opracowano odpowiednie oprogramowanie.

## 2.1. Akwizycja cyfrowych obrazów stereo

W prototypowej wersji systemu stereowizyjnego (Rzeszotarski *et al.*, 2006) zastosowano 2 kolorowe kamery cyfrowe Flea® firmy Point Grey (Rys. 2) wyposażone w przetwornik CCD typu 1/3" (*Inteline Transfer*, 12 bit, rozdzielczość 1024×768, szybkość 30 FPS), interfejs IEEE 1394 (*FireWire*) oraz obiektywy Ernitec o stałej ogniskowej  $f = 3.5$  mm. Synchronizację akwizycji obrazów z dwóch kamer zapewnił producent przez konstrukcję samych kamer oraz dostarczone funkcje API (*Application Programming Interface*). Synchronizacja migawek obu kamer podłączonych do magistrali *FireWire* jest zapewniona sprzętowo z dokładnością 125  $\mu$ s.



Rys. 2. Kamery cyfrowe Flea® umieszczone w głowicy mechanicznej (baza)

## 2.2. Kalibracja kamer

Kalibracja fotogrametryczna kamer została przeprowadzona za pomocą opracowanego programu wykorzystującego funkcje i klasy kompilatora C++, które są dostępne w bibliotece OpenCV (*Intel® Open Source Computer Vision Library*) (Open SCVL, 2007).

Na podstawie wielokrotnych zdjęć pola testowego (tablica w formie czarno-białej szachownicy o wymiarach 0.6 m×0.6 m) program wyznacza elementy orientacji wewnętrznej (Fusiello *et al.*, 2000, 2007):  $c_x, c_y, x'_0, y'_0$  zawarte w macierzy A (1), współczynniki wielomianu korygującego dystorsję radialną:  $k_1, k_2$  (2), parametry modelujące dystorsję tangencjalną (3):  $p_1, p_2$  oraz współczynnik  $\gamma$  (1) określający kąt między osiami  $x$  i  $y$  sensora cyfrowego (*shear*).

$$(1) \quad A = \begin{bmatrix} c_x & \gamma & x_0 \\ 0 & c_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2) \quad \begin{bmatrix} x_d \\ y_d \end{bmatrix} = (1 + k_1 r^2 + k_2 r^4) \begin{bmatrix} x_n \\ y_n \end{bmatrix} + \Delta$$

$$(3) \quad \Delta = \begin{bmatrix} 2p_1 x_n y_n + p_2 (r^2 + 2x_n^2) \\ 2p_2 x_n y_n + p_1 (r^2 + 2y_n^2) \end{bmatrix}$$

Współrzędne obrazowe, po uwzględnieniu parametrów korekcji liczone są na podstawie wzoru (4).

$$(4) \quad \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = A \times \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix}$$

### 2.3. Wyznaczenie elementów orientacji zewnętrznej i wzajemnej

Rejestracja pola testowego pozwala na wyznaczenie elementów orientacji zewnętrznej kamery prawej  $R_R$ ,  $T_R$  oraz lewej  $R_L$ ,  $T_L$ , które określają położenie ich lokalnych układów współrzędnych względem tego samego układu zewnętrznego. Na ich podstawie zostają wyznaczone macierz obrotu  $R_{LR}$  (5) oraz wektor translacji  $T_{LR}$  (6), które definiują wzajemne położenie kamery lewej względem prawej:

$$(5) \quad R_{LR} = R_R R_L^T \quad (6) \quad T_{LR} = T_L - R_{LR}^T T_R$$

W środowisku naukowym asocjacji IEEE (*Institute of Electrical and Electronics Engineers*) zajmującym się stereowizją cyfrową macierz  $R_{LR}$  oraz wektor  $T_{LR}$  określane są mianem zewnętrznych parametrów układu kamer stereo (Cyganek, 2002). Dla danego punktu P w przestrzeni sceny, związek między współrzędnymi tego punktu w układzie kamery lewej względem układu kamery prawej jest następujący (7):

$$(7) \quad P_R = R_{LR} (P_L - T_{LR})$$

### 2.4. Normalizacja obrazów cyfrowych

Podstawowym problemem w stereowizji jest utworzenie układu kanonicznego dwóch kamer, w którym linie epipolarne są równoległe do wierszy obrazów cyfrowych. Utworzenie układu kanonicznego wymaga wcześniejszej kalibracji (wyznaczenie podstawowych elementów orientacji wewnętrznej:  $c_K$ ,  $x'_0$ ,  $y'_0$ , parametrów dystorsji obiektywów:  $k_i$ ,  $p_i$  i sensorów cyfrowych  $c_i$ ) oraz normalizacji pary obrazów. Dzięki temu algorytm realizujący *matching* może ograniczyć szukanie punktów reprezentujących ten sam punkt w przestrzeni do tego samego wiersza odpowiadających sobie obrazów kamer.

Na podstawie znanych elementów orientacji zewnętrznej układu stereowizyjnego  $R_{LR}$ ,  $T_{LR}$  oraz elementów orientacji wewnętrznej obu kamer, zostaje wyznaczona para macierzy przekształcenia (rektyfikacji)  $M_R$ ,  $M_L$  (Fusiello *et al.*, 2000, 2007) za pomocą, których można dokonać normalizacji obrazów pierwotnych do układu kanonicznego. Współrzędne punktów po normalizacji dla obrazów kamery lewej ( $x_{rectL}$ ,  $y_{rectL}$ ) oraz prawej ( $x_{rectR}$ ,  $y_{rectR}$ ) określone są następującymi wzorami (8):

$$(8) \quad \begin{bmatrix} a_L \\ b_L \\ c_L \end{bmatrix} = M_L \times \begin{bmatrix} x_{pL} \\ y_{pL} \\ 1 \end{bmatrix} \quad \begin{bmatrix} a_R \\ b_R \\ c_R \end{bmatrix} = M_R \times \begin{bmatrix} x_{pR} \\ y_{pR} \\ 1 \end{bmatrix}$$

$$\begin{aligned} x_{rectL} &= a_L / c_L & x_{rectR} &= a_R / c_R \\ y_{rectL} &= b_L / c_L & y_{rectR} &= b_R / c_R \end{aligned}$$

## 2.5. Algorytm generowania gęstej mapy dysparycji

Rozwiązanie zadania odpowiedności (*correspondence problem*) homologicznych punktów obrazowych w układzie kanonicznym sprowadza się do wyznaczenia paralaks podłużnych, tzw. dysparycji (*disparity*)  $d_{RL}$ . Obliczona dysparycja dla każdego punktu sceny 3D tworzy tzw. gęstą mapę dysparycji.

Proces automatycznego wyznaczenia w czasie rzeczywistym gęstej mapy dysparycji dla pary obrazów został wsparty algorytmem dopasowującym obszary. W algorytmie tym dopasowaniu podlegają obszary o zadanych wymiarach ( $winx \times winy$  - okno korelacji). Najczęściej zastosowaną miarą podobieństwa pomiędzy pikselem obrazu odniesienia a pikselem drugiego obrazu jest suma wartości bezwzględnych różnic *SAD* (*Sum of Absolute Differences*) (Brown *et al.*, 2003; Cyganek, 2002; Di Stefano *et al.*, 2004; Mao *et al.*, 2004; Muhlmann *et al.*, 2002). Dla obrazów kolorowych miara *SAD* opisana jest formułą (9):

$$(9) \quad \begin{aligned} SAD(x_R, y_R, d) &= \\ &= \sum_{i=-\frac{1}{2}(winx-1)}^{\frac{1}{2}(winx-1)} \sum_{j=-\frac{1}{2}(winy-1)}^{\frac{1}{2}(winy-1)} [ |R_R(x_R+i, y_R+j) - R_L(x_L+i+d, y_L+j)| + \\ & \quad |G_R(x_R+i, y_R+j) - G_L(x_L+i+d, y_L+j)| + \\ & \quad |B_R(x_R+i, y_R+j) - B_L(x_L+i+d, y_L+j)| ] \end{aligned}$$

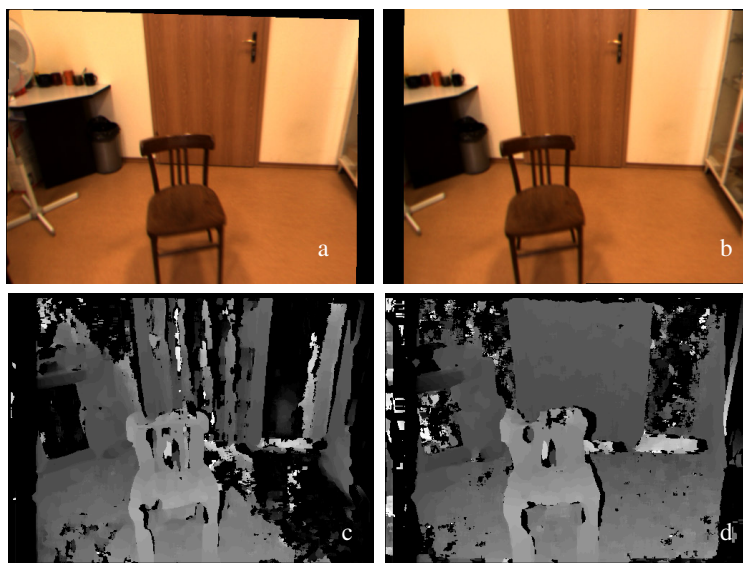
Miara *SAD* jest obliczana dla każdego punktu obrazu odniesienia w zadanym zakresie dysparycji (przesunięcia)  $d$ , tj.  $d_{min} - d_{max}$ . Wartości *SAD* ( $x_R, y, d$ ) przechowywane są w trójwymiarowej tablicy, tzw. mapie przestrzeni dysparycji (*Disparity Space Map*) o wymiarach  $(w+winx+1) \times (d_{max} - d_{min} + 1) \times (h+1)$ . Dla obliczania kolejnych sum *SAD* zastosowano w algorytmie technikę "przesuwanych okien" (*sliding window*).

Poszukiwanie wiarygodnych minimów miary *SAD* jest wykonywane w czterech krokach przez znalezienie minimum globalnego  $SAD_{min}$  oraz trzech najmniejszych wartości  $SAD_1 < SAD_2 < SAD_3$  (Brown, *et al.*, 2003). Wizualizację mapy dysparycji wyznaczonej metodą "trzech minimów", przed i po normalizacji obrazów, przedstawiono na rysunku 3. Odpowiadająca znalezionej minimum  $SAD_{min}$  wartości  $d$  uznawana jest za prawidłowo wyznaczoną dysparycję, jeżeli trzecia najmniejsza wartość  $SAD_3$  jest większa od zadanej

wartości progowej  $Th$  wyznaczonej na podstawie  $SAD_{min}:Th=SAD_{min}+p \times SAD_{min}$ , gdzie zalecaną wartością parametru  $p$  jest  $0.05 \div 0.1$  (Muhlmann *et al.*, 2002).

Zaimplementowany algorytm dopasowujący obszarami przebiega w czterech krokach:

1. Wyznacz  $SAD(x_R, y, d)$  dla danego punktu obrazu prawego  $p_R(x_R, y)$  i dla danego punktu obrazu lewego  $p_L(x_L, y)$  przesuniętego o dysparycję  $d$ .
2. Powtórz pkt. 1 dla wszystkich wartości  $d$  dla zadanego zakresu  $d_{min}, d_{max}$ .
3. Z uzyskanego zbioru miar  $SAD$  dla różnych przesunięć  $d$ , tj.  $SAD(d=d_{min}), SAD(d=d_{min}+1), SAD(d=d_{max})$  wybierz najmniejszą wartość  $SAD_{min}$ ; wartość dysparycji  $d$ , dla której została wyznaczona minimalna wartość  $SAD_{min}$  jest poszukiwaną dysparycją pomiędzy punktem obrazu odniesienia a punktem obrazu lewego:  $d_{RL} = d(SAD_{min}), x_L = x_R + d_{RL}$ .
4. Weryfikuj wybór przesunięcia  $d$ , dla którego  $SAD_{min}$  osiąga minimum przez badanie kolejnych dwóch najmniejszych wartości i odpowiadających im wartości przesunięć.



Rys. 3. Obrazy zarejestrowane przez kamerę lewą (a), prawą (b), obraz dysparycji wyznaczony wg algorytmu "trzech minimów" (c) oraz obraz dysparycji wyznaczony wg algorytmu "trzech minimów" po wcześniejszej normalizacji obrazów (d)

W zastosowanym algorytmie dopasowania obrazów (*matching*) zaproponowano dodatkową miarę pewności, z jaką wyznaczane są minima wartości  $SAD$  (Rzeszotarski *et al.*, 2007). Jest to stosunek nachylenia funkcji  $SAD$  do minimalnej wartości  $SAD$ . Ponieważ są dwa nachylenia funkcji  $SAD$  z lewej i prawej strony wartości  $SAD_{min}$ , to jako ostateczna miara pewności została przyjęta mniejsza wartość. Zatem przyjęto, że im mniejsza wartość minimum  $SAD$  i im większe nachylenie, to tym bardziej prawdopodobne jest, że znaleziona wartość minimum  $SAD$  odpowiada właściwej wartości wyznaczonej dysparycji.

## 2.6. Rekonstrukcja sceny 3D

Za pomocą mapy dysparycji oraz elementów orientacji wewnętrznej zdjęć po normalizacji oraz zewnętrznej modelu, zapisanych w macierzach przekształcenia  $M_R, M_L$ , wyznaczane są współrzędne przestrzenne obiektów sceny 3D.

W celu wyznaczenia współrzędnych przestrzennych rekonstruowanego punktu trzeba rozwiązać liniowe równanie (10):

$$(10) \quad ap_R - bRp_L + c(p_R \times Rp_L) = T$$

Oznaczając rozwiązanie powyższego układu równań jako  $[a_o, b_o, c_o]$ , to współrzędne punktu w przestrzeni można wówczas wyznaczyć na podstawie wzoru (11):

$$(11) \quad P = a_o p_R + 0.5 \cdot (b_o (T + Rp_L) - a_o p_R)$$

## 2.7. Segmentacja obrazu

Zastosowany algorytm segmentacji sceny 3D realizuje w trybie iteracyjnym detekcję płaszczyzn oraz „przeszkód”, tj. obiektów, które nie spełniają równania płaszczyzny. Proces poszukiwania jest wykonywany w siatce punktów o strukturze TIN, w interwale 16 pikseli. Segmentacja obrazu polega na klasyfikacji punktów na podstawie dwóch składowych barwy obrazu  $R_N, B_N$ , normalizowanych w celu eliminacji wpływu lokalnych zmian jasności oświetlenia zewnętrznego.

Wynikiem segmentacji jest tablica (mapa segmentacji), której elementy przechowują indeksy klas barwy, do których zostały przyporządkowane piksele analizowanego obrazu. Na podstawie mapy dysparycji i mapy segmentacji wyznaczone zostają: pole powierzchni, współrzędne przestrzenne środka ciężkości wydzielonych obiektów oraz odległość między środkiem ciężkości a początkiem układu współrzędnych systemu stereowizyjnego. Na podstawie danych opisujących scenę 3D zostaje wygenerowana dźwiękowa reprezentacja otoczenia (Skulimowski, Strumiłło, 2006).

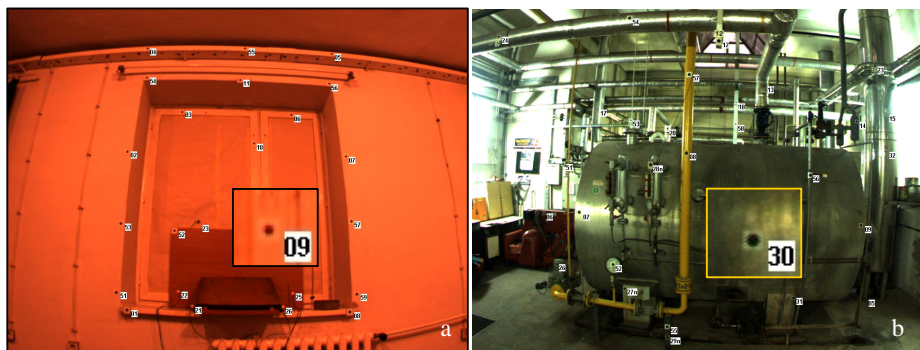
## 3. BADANIE DOKŁADNOŚCI SYSTEMU STEREOWIZYJNEGO

Praktyczną ocenę dokładności systemu stereowizyjnego wykonano na podstawie zdjęć cyfrowych dwóch przestrzennych pól testowych o wymiarach ca 1.5 m×2 m×2.5 m (test 1) oraz 5 m×3 m×7 m (test 2). Pole testowe 1 (Rys. 4a) posiadało 26 sygnalizowanych punktów kontrolnych o dokładności  $s_{XYZ} = \pm 0.5$  mm, pole testowe 2 (Rys. 4b) wyposażone było w 30 punktów określonych z dokładnością  $s_{XY} = \pm 0.2$  mm,  $s_z = \pm 0.3$  mm.

Zdjęcia o rozdzielczości 1024×768 pikseli rejestrowano z bazą fotografowania  $B = 80.75$  mm (stosunek bazowy wynosił  $v = 1/30 \div 1/75$ ) w warunkach bardzo niekorzystnego sztucznego oświetlenia. Wykonana dodatkowo cyfrowa normalizacja nie wpłynęła w widoczny sposób na poprawę radiometrycznej jakości obrazów.

Współrzędne punktów zostały wyznaczone w 3 seriach pomiarowych na podstawie wygenerowanej gęstej mapy dysparycji (okno korelacji wynosiło 25×25 pikseli) oraz na podstawie pomiarów manualnych na obrazach cyfrowych. Zastosowany algorytm dysparycji wykrył ca 70 % punktów dla pola testowego 1, natomiast dla pola testowego 2

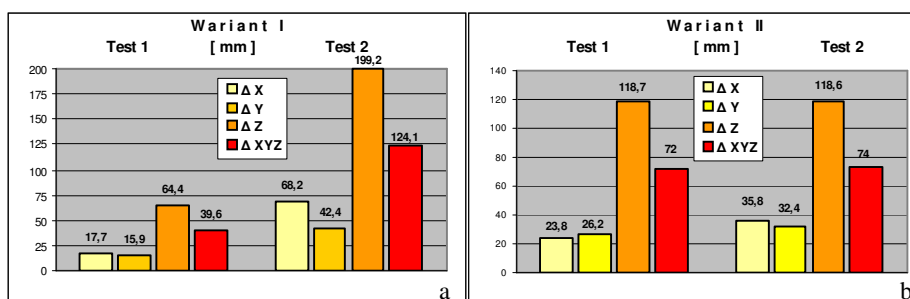
rozpoznał ca 75 % punktów. Przy rozdzielczości obrazów cyfrowych 1024x768 pikseli uzyskano prędkość przetwarzania 0.5 klatki/s.



Rys. 4. Obiekty testowe – test 1 (a) oraz test 2 (b)

Ocenę dokładności pozycjonowania punktów wykonano w dwóch wariantach. Porównano współrzędne przestrzenne obliczone na podstawie wygenerowanej gęstej mapy dysparycji w odniesieniu do współrzędnych wyznaczonych na podstawie bezpośrednich pomiarów na obrazach cyfrowych (wariant I). Średnia odchyłka kwadratowa w układzie lokalnym systemu stereowizyjnego (Rys. 5a) wyniosła odpowiednio  $\Delta_{XYZ} = \pm 39.6$  mm (test 1) oraz  $\Delta_{XYZ} = \pm 124.1$  mm (test 2).

Przy porównaniu wartości współrzędnych X, Y, Z (odległość) obliczonych na podstawie wygenerowanej gęstej mapy dysparycji, po wykonaniu transformacji 3D do układu obiektu, ze współrzędnymi punktów kontrolnych pola testowego (wariant II), średnia odchyłka kwadratowa (Rys. 5b) wyniosła  $\Delta_{XYZ} = \pm 72$  mm (test 1) i  $\Delta_{XYZ} = \pm 74$  mm (test 2).



Rys. 5. Dokładność pozycjonowania systemem stereowizyjnym dla punktów pola testowego 1 oraz 2 (odchyłki średniokwadratowe dla wyznaczonych współrzędnych X,Y,Z i położenia punktu  $\Delta_{XYZ}$ )



Zmniejszenie dokładności pozycjonowania obiektów za pomocą cyfrowego systemu stereowizyjnego może być spowodowane wpływem przede wszystkim następujących czynników:

- niska rozdzielczość sensorów wizyjnych
- bardzo mały, niekorzystny stosunek bazowy  $v = B/Z_F$ , który ma głównie wpływ na zmniejszenie dokładności wyznaczenia współrzędnej  $Z$
- błąd synchronizacji pozyskania obrazów stereo podczas rejestracji obiektów będących w ruchu
- wykonanie pomiaru punktów oraz przetwarzanie cyfrowe na całkowitych współrzędnych pikselowych
- niedostateczna korekcja błędów geometrycznych obrazu wynikająca z techniki kalibracji, opartej na algorytmie OpenCV
- błędy generowania gęstej mapy dysparycji przez zastosowany algorytm dopasowujący obrazy stereo (*stereo matching algorithm*)
- błędy segmentacji obrazu

#### 4. WNIOSKI, *OUTLOOK*

Prototypowy cyfrowy system stereowizyjny oparty na sensorach o rozdzielczości 1024×768 pikseli, przy odległości obrazowania  $Z_F$  kilku metrów, pozwala na automatyczne pozycjonowanie obiektów sceny 3D ze średnią dokładnością rzędu  $\pm 0.1$ m. Otrzymana dokładność wyznaczenia współrzędnych 3D jest zgodna z rozkładem błędów dla stereogramu zdjęć normalnych, tzn. błąd wyznaczenia współrzędnej  $Z$  rośnie w kwadracie, natomiast współrzędnych  $X$  i  $Y$  wprost proporcjonalnie do odległości fotografowania  $Z_F$ .

W elektronicznym systemie wspomaganie ruch osoby niewidomej, ze względu na wielkość obiektów klasyfikowanych do rozpoznania, ich liczba i dokładność pozycjonowania w obszarze sceny 3D odgrywa mniejszą rolę. Akwizycja obrazów o mniejszej rozdzielczości spowoduje przede wszystkim znaczące skrócenie czasu przetwarzania oraz zwiększenie częstotliwości rejestracji, docelowo do trybu *real-time*.

Konstruowany system stereowizyjny pracować będzie w czasie rzeczywistym, tj. 25 obrazów/s (obecnie 3.5 ramki/s), na obrazach o rozdzielczości 360×288 pikseli i w modelu barw YCrCb. Zastosowanie modelu barw YCrYCb spowodowane jest trybem pracy sensora wizyjnego i portów wyjścia procesora wideo. Poza tym, w stosunku do tradycyjnego modelu barw sRGB, obraz ma znacznie mniejszy ciężar bajtowy a mapa dysparycji liczona jest wprost ze składowej jasności  $Y$ , bez konieczności jej wyznaczenia na drodze interpolacji.

Rejestracja obrazów cyfrowych za pomocą systemu stereowizyjnego, algorytm wyznaczania dysparycji przez dopasowanie obszarów za pomocą miary *SAD* oraz algorytm segmentacji zostały obecnie zaimplementowane na platformie procesora sygnałowego DSP (*Digital Signal Processor*). Na podstawie danych przesłanych z procesora DSP zostaje wygenerowana w technice 3D audio przez układ FPGA (*Field Programmable Gate Array*) firmy Xilinx procedura dźwiękowej reprezentacji otoczenia.

## 5. LITERATURA

- Brown M. Z., Burschka D., Hager G. D., 2003. Advances in computational stereo. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 8, s. 993-1008.
- Cyganek B., 2002. *Komputerowe przetwarzanie obrazów trójwymiarowych*. Akademicka Oficyna Wydawnicza EXIT, Warszawa.
- Di Stefano L., Marchionni M., Mattoccia S., 2004. A PC-based Real-Time Stereo Vision System. *Machine Graphics & Vision*, Vol. 13, No. 3, s. 197-220.
- Fusiello A., Trucco E., Verri A., 2000. A compact algorithm for rectification of stereo pairs. *Machine Vision and Applications*, Vol. 12, s. 16-22
- Fusiello A., Trucco E., Verri A., 2007. Rectification of Stereo Pairs. <http://profs.sci.univr.it/~fusiello/demo/rect/>
- Mao Y., Soatto S., Kosecka J., Sastry S.S., 2004. An invitation to 3-D vision: from images to geometric transformations, *Springer-Verlag*, New York, Inc.
- Muhlmann K., Maier D., Hesser J., Manner R., 2002. Calculating dense disparity maps from color stereo images, an efficient implementation. *International Journal on Computer Vision*, Vol. 47, No. 1-3, s. 79-88.
- Open Source Computer Vision Library, 2007. <http://www.intel.com/technology/computing/opencv/index.htm>
- Rzeszotarski D., Strumiłło P., Pełczyński P., Więcek B., Lorenc A., 2006. System obrazowania stereoskopowego sekwencji scen trójwymiarowych. *Zeszyty Naukowe Elektronika Politechniki Łódzkiej*, Nr 10, s. 165-184.
- Rzeszotarski D., Skulimowski P., Strumiłło P., 2007. A method for verification of dense disparity maps computed from the matching algorithm implemented in the stereovision system. Signal Processing Symposium, Jachranka.
- Skulimowski P., Strumiłło P., 2006. Detekcja płaszczyzn w sekwencji obrazów stereowizyjnych. V Sympozjum Naukowe Techniki Przetwarzania Obrazu, Serock, s. 282-287.

### PROTOTYPE STEREOVISION SYSTEM OF MACHINE VISION TYPE – OPERATION AND ACCURACY

KEY WORDS: digital camera, stereovision system, disparity, accuracy, signal processor

SUMMARY: The prototype of digital stereovision system is presented along with its procedures, main parts and its accuracy evaluation. The presented system is a prototype for vision module in electronic travel aid device for the blind, being designed and developed in Institute of Electronics in Technical University of Lodz. The system consists of two Flea type Point Grey digital cameras (resolution of 1024×768, IEEE 1394 interface) and Ernitec f=3.5 mm objective). The cameras calibration was performed by means of the software using OpenCV library (Intel® Open Source Computer Vision Library). The process of automatic computation of dense disparity maps was performed by the stereo matching local algorithm (with correlation window of a given size). The measure of similarity is defined as a sum of absolute differences (SAD) between pixels form windows surrounding the compared pixels. The SAD values are calculated with the use of sliding window method in order to obtain real time speed of dense disparity maps computation. The procedure of finding robust

minimum SAD values consists of four steps: searching for global minimum value  $SAD_{\min}$  and three other smallest values  $SAD_1 < SAD_2 < SAD_3$ . The 3D scene coordinates are calculated by means of disparity map values and model exterior orientation parameters (transformation matrices  $M_R, M_L$ ). 3D scene segmentation algorithm detects planes and surrounding objects, obstacles. The digital images acquisition, disparity and segmentation map generation are implemented on the DSP (Digital Signal Processor) platform. The accuracy of the system is evaluated by analysing stereo pairs of digital pictures of two different 3D test fields. The base line of stereoscopic set of cameras used in the system was  $B = 80.75$  mm. 3D coordinates were computed by means of dense disparity map values (search window of the matching algorithm was of the size  $25 \times 25$  pixels) The average effectiveness of signalised check points detection by disparity algorithm was 72 %. RMSE for 3D object coordinates in the stereovision system was  $\Delta_{XYZ} = \pm 39.6$  mm (test 1) and  $\Delta_{XYZ} = \pm 124.1$  mm (test 2). RMSE for X,Y,Z coordinates after 3D transformation into the object coordinates system was  $\Delta_{XYZ} = \pm 72$  mm (test 1) and  $\Delta_{XYZ} = \pm 74$  mm (test 2).

Mgr inż. Dariusz Rzeszotarski  
e-mail: [dariusz.rzeszotarski@p.lodz.pl](mailto:dariusz.rzeszotarski@p.lodz.pl)  
tel. +42 6312656

Dr inż. Piotr Sawicki  
e-mail: [piotr.sawicki@geodezja.pl](mailto:piotr.sawicki@geodezja.pl)  
tel. +89 5233282

Mgr inż. Bartosz Ostrowski  
e-mail: [bostrow@p.lodz.pl](mailto:bostrow@p.lodz.pl)  
tel. +42 6312642