

Sławomir Jeżewski*, Sylwester Błaszczuk*

Porównywanie algorytmów wykrywania krawędzi do zadań rekonstrukcji trójwymiarowego kształtu obiektu

1. Wprowadzenie

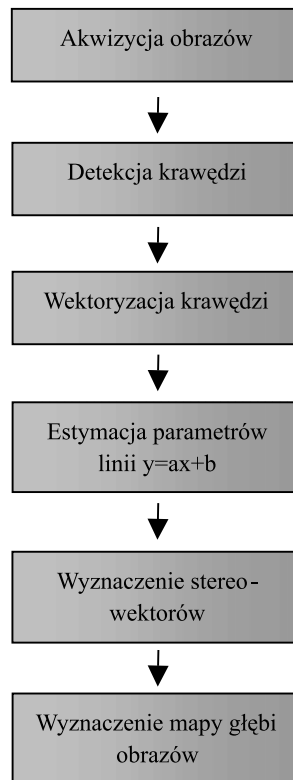
Rekonstrukcja i modelowanie obiektów w przestrzeni trójwymiarowej w oparciu o stereoskopowy układ kamer jest zagadnieniem badawczym rozpatrywanym od wielu lat. W ostatnich latach badania w tym zakresie nabrały tempa w związku z potencjalnymi obszarami wykorzystania takimi jak tworzenie map terenów zurbanizowanych [5] oraz w nawigacji robotów w terenie zurbanizowanym [6–9]. Zagadnienie rekonstrukcji i modelowania obiektów można rozpatrywać jako proces dwuetapowy składający się z: etapu tworzenia obrazu stereoskopowego (stereo rekonstrukcji), oraz etapu rozpoznawania i lokalizacji obiektów w stereobrazie. Na pierwszy etap składają się metody wyznaczające głębię obrazu na podstawie dwu lub więcej obrazów składowych, natomiast na drugi rozpoznawanie obiektów 3D i ich precyzyjna lokalizacja w przestrzeni. Metody stereorekonstrukcji można podzielić na metody krawędziowe [4] oraz metody korelacyjne [2, 3]. W metodach krawędziowych mapę dysparycji wyznacza się w oparciu o dopasowanie pikseli krawędziowych z poszczególnych obrazów układu stereo. Powstaje wtedy rzadka mapa dysparycji. Metody korelacyjne głębokość w obrazie obliczają na podstawie korelacji poszczególnych regionów obrazu. Efektem ich działania jest gęsta, choć obciążona dość sporym szumem mapa dysparycji. Choć obydwie metody stereorekonstrukcji dają koncepcyjnie tę samą wielkość – dysparycję, to cechy charakterystyczne tych map powodują, że algorytmy z etapu rozpoznawania i lokalizacji obiektów są specyficzne dla danego typu mapy.

W niniejszej publikacji została przedstawiona koncepcja algorytmu do rekonstrukcji trójwymiarowej obiektów oparta na wykrywaniu krawędzi. Łączy ona w sobie obydwie fazy rekonstrukcji. W proponowanym podejściu krawędzie rekonstruowane są do postaci wektorów i na podstawie informacji wektorowej wyznaczana jest głębia obrazu. Algorytm ten jest rozwinięciem idei stereowizji krawędziowej, ale przez swój charakter może stanowić uzupełnienie stereowizji korelacyjnej.

* Katedra Informatyki Stosowanej, Politechnika Łódzka

2. Koncepcja algorytmu

Zaproponowany algorytm składa się z 5 etapów: detekcji krawędzi, wektoryzacji krawędzi, estymacji parametrów linii, wyznaczenie par stereowektorów, rekonstrukcji położenia linii w przestrzeni (rys. 1).



Rys. 1. Schemat algorytmu stereowizyjnego opartego na stereowektorach

Do detekcji krawędzi autorzy stosowali różne filtry krawędziujące wraz z algorytmami usuwającym piksele niemaksymalne (zagadnienie opisane w rozdziale 5). Wektoryzację krawędzi oparto na uproszczonej transformacji Hougha [16–19], zaś do estymacji parametrów linii została użyta metoda najmniejszych kwadratów. W wyniku przeprowadzonych obliczeń otrzymano listę wektorów z wyznaczonymi współczynnikami kierunkowymi, początkami i końcami. W liście tej wyszukane zostały wektory korespondujące. Powstałe pary poddane zostały procesowi stereo rekonstrukcji według zależności opisanych rozdziale 4. Należy zauważyć, że w podanym schemacie obliczeń występują dwie czasochłonne operacje – wektoryzacja krawędzi oraz estymacja parametrów linii, których istnienie można postrzekać za wadę rozwiązania. Jednak wektoryzacja krawędzi jest nieodzownym etapem

analizy, przetwarzania i rozpoznawania obrazu. Informacja wektorowa jest więc dostępna w systemie przetwarzania obrazu i można jej użyć w procesie stereo rekonstrukcji nie zmieniając znacznie nakładu obliczeniowego w całym systemie. Informacja ta użyta w procesie stereo rekonstrukcji jak wykazano podniesie precyzję lokalizacji obiektów

3. Teoretyczne podstawy opracowanego algorytmu

W niniejszym rozdziale wyprowadzono podstawowe zależności wektorowej stereorekonstrukcji, na ich podstawie w rozdziale 4 oszacowano niepewność pomiaru. W rozważaniach przyjęto, że zespół dwu kamer perspektywicznych [11] obserwuje obiekt liniowy (wektor) położony dowolnie w przestrzeni 3D. Dla prostoty wywodu przyjęto najprostszy model kamery – kamerę perspektywiczną [11], mimo że istnieją bardziej rozbudowane modele, zawarte w pracach Toskani [12], Fauergas i Toskani [13], Tsai [14, 15].

Przy rozważaniach przyjęto, że układ współrzędnych świata pokrywa się z układem współrzędnym kamery. Punkt centralny kamery A położony jest w środku układu współrzędnych. Zaś jej oś główna pokrywa się z osią Z układu współrzędnych. Kamera B jest umieszczona równolegle do niej i oddalona o odległość T_x w kierunku osi X. W sensie matematycznym kamera perspektywiczna transformuje trójwymiarowy punkt $P = [p_x, p_y, p_z]^T$ do jego dwuwymiarowego obrazu $Q = [u, v]$.

Podstawowe zależności rzutowe wiążą wewnętrzne parametry kamery z obiektem i jego obrazem:

$$\frac{u}{p_x} = \frac{f}{Z}, \quad \frac{v}{p_y} = \frac{f}{Z} \quad (1)$$

gdzie f – jest ogniskową kamery.

Zapisem macierzowym tej operacji jest:

$$Q' = \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot P = F' \cdot P \quad (2)$$

$$u = \frac{u'}{w}, \quad v = \frac{v'}{w} \quad (3)$$

Obserwowany odcinek w przestrzeni 3D oznaczono jako $R(t)$:

$$R(t) = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} = \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} \cdot t + \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} = {}_3\bar{V} \cdot t + {}_3\bar{A} \quad (4)$$

gdzie $t \in (0, 1)$. Jego transformacja rzutowa wyraża się wzorem:

$$Q' = F' \cdot R(t) \quad (5)$$

W przypadku kanonicznego układu kamer stereowizyjnych otrzymujemy dwa różne obrazy linii prostej:

$$Q_1' = \begin{bmatrix} f & 0 & 0 & T_x \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot R(t) \quad (6)$$

$$Q_2' = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot R(t)$$

Po podstawowych przekształceniach równań (1) oraz (2) (4) (6) otrzymujemy:

- wyrażenia określające współczynnik kierunkowy prostej a w obrazie:

$$a = \frac{A_x V_z - A_z V_x}{A_y V_z - A_z V_y} - \frac{T_x}{f} \cdot \frac{V_z}{A_y V_z - A_z V_y} \quad (7)$$

- wyrażenie określające różnicę w ukierunkowaniu prostej w kamerach $\Delta a = a_A - a_B$:

$$\Delta a = \frac{T_x}{f} \cdot \frac{V_z}{A_y V_z - A_z V_y} \quad (8)$$

- wyrażenie na składowe wektora kierunkowego prostej $R(t)$ wyliczone na podstawie wielkości pikselowych:

$$V_x = \frac{T_x}{fK} \left[u_A(1) - u_A(0) + \frac{u_A(1)K}{\Delta a \cdot v_A(0) - K} \right]$$

$$V_z = \frac{T_x}{fK} \frac{K}{\Delta a \cdot v_A(0) - K} \quad (9)$$

$$V_y = \frac{T_x}{f} \frac{v_A(0)}{K}$$

$$K = u_A(0) - u_B(0)$$

gdzie:

- $u_A(1), u_A(0)$ – współrzędne x początku i końca wektora na kamerze A,
- $v_A(1), v_A(0)$ – współrzędne y początku i końca wektora na kamerze A.

3.1. Estymacja niepewności lokalizacji wektora w przestrzeni 3D

Rekonstrukcja linii prostej w przestrzeni 3D na podstawie danych stereowizyjnych jest silnie zależna od sumy w obrazie i niedokładności lokalizacji pojedynczych pikseli. Analiza niepewności pomiaru ma istotne znaczenie dla zakresu stosowności tej metody.

Poniżej przedstawiono oszacowanie niepewności pomiaru współczynników kierunkowych rekonstruowanych prostych. Do estymacji przyjęto, że prosta w obrazie reprezentowana jest przez zbiór punktów (x_i, y_i) gdzie y_i jest zmienną stochastyczną o rozkładzie normalnym: $N(\hat{y}_i, \sigma) = \hat{y}_i + N(0, \sigma) = \hat{y}_i + e_i$.

Współczynnik kierunkowy prostej wyznaczony metodą najmniejszych kwadratów przyjmuje postać:

$$a = \frac{N \cdot \sum_{i=0}^N x_i \hat{y}_i - \sum_{i=0}^N \hat{y}_i \sum_{i=0}^N x_i}{N \sum_{i=0}^N x_i^2 - \left(\sum_{i=0}^N x_i \right)^2} + \frac{N \cdot \sum_{i=0}^N x_i e_i - \sum_{i=0}^N e_i \sum_{i=0}^N x_i}{N \sum_{i=0}^N x_i^2 - \left(\sum_{i=0}^N x_i \right)^2} = a_0 + a_m \quad (10)$$

Wartość oczekiwana tego rozkładu wynosi a_0 , zaś wariancja parametru wynosi:

$$\text{var}\{a\} = E(a - a_0)(a - a_0) = E(a_m)^2 \quad (11)$$

$$\text{var}\{a\} = E \left(\frac{N^2 \sum_{i=0}^N (x_i e_i)^2 - 2N \sum_{i=0}^N e_i x_i \sum_{i=0}^N e_i \sum_{i=0}^N x_i + \left(\sum_{i=0}^N e_i \right)^2 \sum_{i=0}^N (x_i)^2}{\left[N \sum_{i=0}^N x_i^2 - \left(\sum_{i=0}^N x_i \right)^2 \right]^2} \right) \quad (12)$$

Powyższy wzór można uprościć, przyjmując, że prosta składa się z sąsiadujących pikseli oraz e_i, e_j są realizacjami niezależnych zmiennych losowych:

$$T = N \sum_{i=0}^N x_i^2 - \left(\sum_{i=0}^N x_i \right)^2 \quad (13)$$

$$\text{var}\{a\} = \frac{N\sigma^2}{T^2} \left[N \sum_{i=0}^N x_i^2 - \left(\sum_{i=0}^N x_i \right)^2 \right] = \frac{N\sigma^2}{T} \quad (14)$$

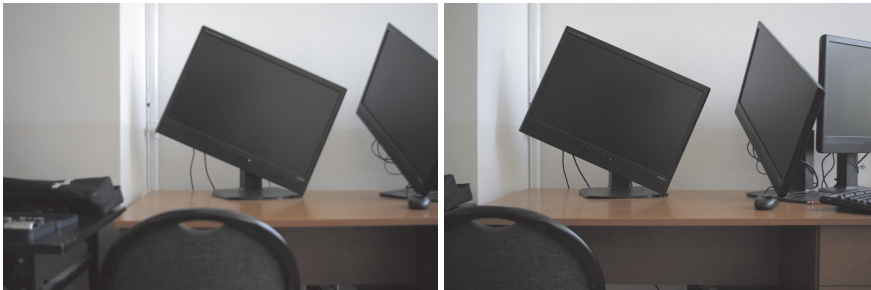
$$\text{var}\{a\} = \frac{12}{N(N^2 - 1)} \sigma^2 \quad (15)$$

gdzie N – liczba punktów w odcinku.

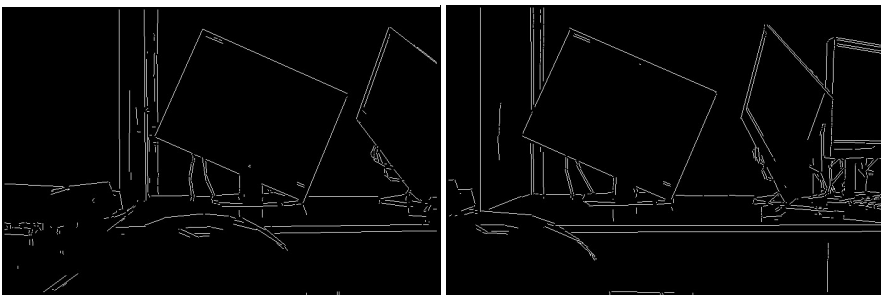
Zależność (15) wykazuje, że przy odcinku o długości już pięciu punktów wariancja oznaczenia współczynnika kierunkowego prostej jest 10-krotnie mniejsza niż przy oznaczeniu pojedynczego punktu i szybko maleje wraz z długością wektora. Wielkość wariancji σ^2 występująca w zależności jest zbiorczym wynikiem wielu procesów począwszy od akwizycji aż do procedury wyznaczania krawędzi i jest trudna do oszacowania teoretycznego. Przy danym rozwiązaniu sprzętowym systemu akwizycji obrazu można ją uważać za stałą tak długo, jak długo nie ulegają zmianie procedury wykrywania krawędzi.

4. Doświadczalne określenie szumu rekonstrukcji

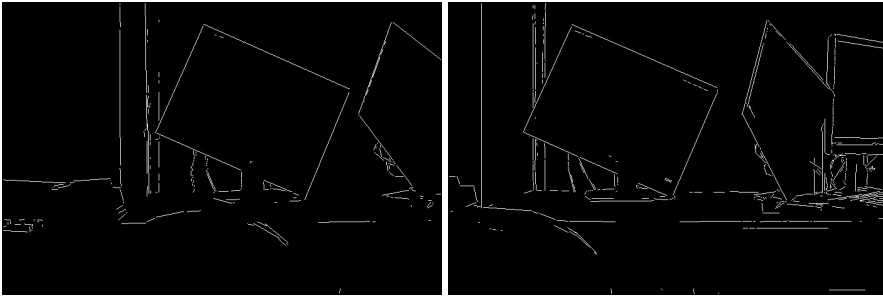
Na rysunkach 2–5 zostały zaprezentowano pomiary wpływu szumu na proces wektoryzacji w obrazie. Wykonano akwizycję sekwencji 64 obrazów tej samej sceny za pomocą dwu kamer WAT 202D tworzących kanoniczny układ stereowizyjny. Uzyskane sekwencje zdjęć poddano procesowi wykrywania krawędzi kolejno przy wykorzystaniu algorytmów: Canny’ego, Fuzzy, oraz LoG w celu ich porównania. Wykryte krawędzie poddano wektoryzacji z użyciem dwuwymiarowej transformaty Hougha. Wybrane zespoły krawędzi poddano analizie statystycznej.



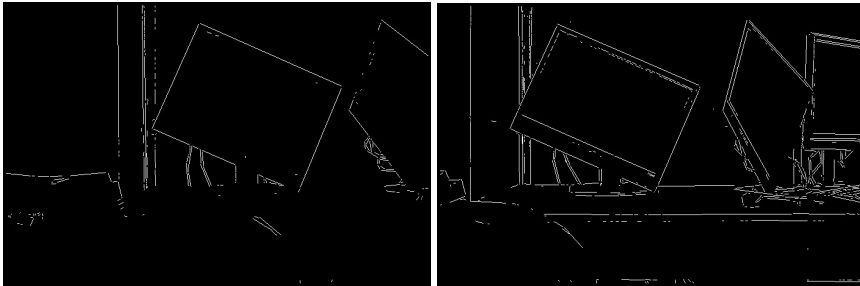
Rys. 2. Przykład zastosowanej stereopary



Rys. 3. Wynik procesu wektoryzacji przy wykorzystaniu algorytmu Canny’ego



Rys. 4. Wynik procesu wektoryzacji przy wykorzystaniu algorytmu Fuzzy



Rys. 5. Wynik procesu wektoryzacji przy wykorzystaniu algorytmu LoG

Dla wyselekcjonowanych wektorów określono statystyczny rozrzut parametrów a , b

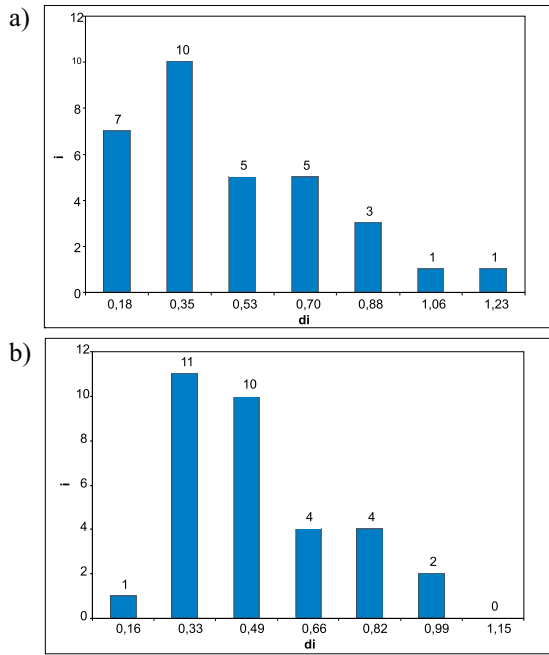
$$y = ax + b \quad (16)$$

Rozrzut współczynników a oraz b został obliczony zgodnie z wzorem:

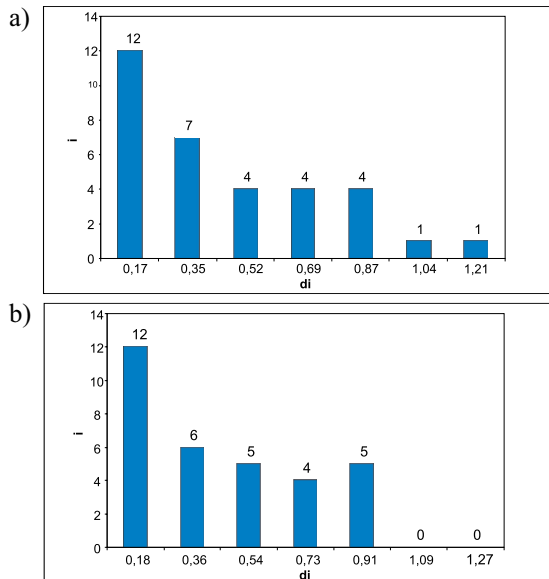
$$d_i = \sqrt{(a_i - \bar{a})^2 + (b_i - \bar{b})^2} \quad i = 0 \dots N \quad (17)$$

Wyniki przeprowadzonych doświadczeń zostały zamieszczone na histogramach przedstawionych na rysunkach 6–8.

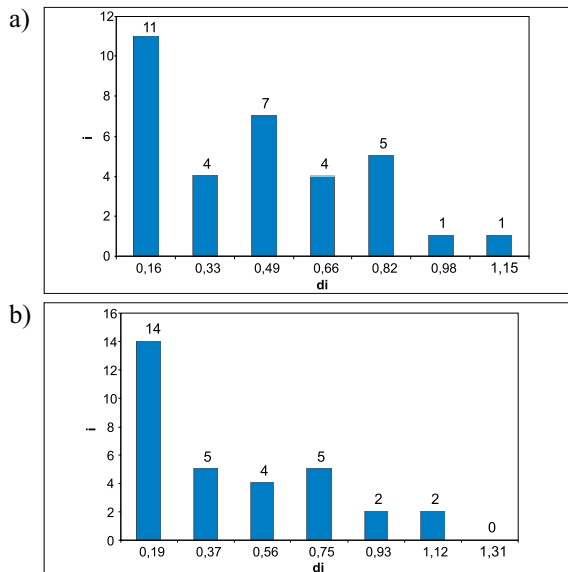
Przedstawione wyniki wskazują, że detektor krawędzi LoG posiada najlepsze właściwości ze względu na wielkość błędu, jednostajność i powtarzalność rozkładu. Detektor Fuzzy posiada większą wariancję mimo to również nadaje się do rekonstrukcji stereowizyjnej. Detektor Canny'ego plasuje się na trzecim miejscu ze względu na stosunkowo dużą wariancję oraz niejednorodność rozkładu z kilkoma minimami.



Rys. 6. Wpływ szumu kamery na położenie wybranego wektora przy zastosowaniu algorytmu Canny'ego dla obrazu: a) lewego; b) prawego



Rys. 7. Wpływ szumu kamery na położenie wybranego wektora przy zastosowaniu algorytmu *Fuzzy edge detector* dla obrazu: a) lewego; b) prawego



Rys. 8. Wpływ szumu kamery na położenie wybranego wektora przy zastosowaniu algorytmu LoG dla obrazu: a) lewego; b) prawego

5. Podsumowanie

Przedstawiona w niniejszym artykule wektorowa stereorekonstrukcja pozwala na wyznaczenie mapy głębokości oraz rekonstrukcję wektorów w przestrzeni 3D. Zrekonstruowane wektory charakteryzują się tym większą dokładnością, im więcej pikseli na nie się składa. Załączone wyniki pokazują, że na jakość rekonstrukcji stereowizyjnej wpływa dobór algorytmu krawędziowania. Przebadano algorytmy Fuzzy edge detection, LoG, oraz Canny'ego. Pierwsze dwa algorytmy cechują się mniejszym rozrzutem położenia wektorów w stosunku do wektorów uzyskanych przy wykorzystaniu algorytmu Canny'ego. Jednak rekonstrukcja wektorowa obrazu wizualnie najlepiej przebiegła dla algorytmu Canny'ego.

Literatura

- [1] Jeżewski S., Sankowski D., Dadan W., *Koncepcja autonomicznego robota pola walki przeznaczanego do zadań zwiadu i wykrywania min*. Automatyka (półrocznik AGH), t. 13, z. 3, 2009.
- [2] Chambon S., Crouzil A., *Towards correlation-based matching algorithms that are robust near occlusion*. Pattern Recognition, 2004, ICPR 2004, Proceedings of the 17th International Conference on.
- [3] Konolige K., *Small Vision Systems: Hardware and Implementation*. Eighth International Symposium on Robotics Research, 1997.
- [4] Eric W., Grimson L., *Computational Experiment with a Feature Based Stereo Algorithm*. 1984.

- [5] Borrmann D., Elseberg J., *Global Konsistente 3D Kartierung am Beispiel des Botanischen Gartens in Osnabrück*. Universität Osnabrück, October 2006 (Bachelor's thesis).
- [6] Cantzler H., Fisher R.B., Devy M., *Improving architectural 3D reconstruction by plane and edge constraining*. Proceedings of the British Machine Vision Conference (BMVC 2002), Cardiff, U.K., September 2002, 43–52.
- [7] Cantzler H., Fisher R.B., Devy M., *Quality enhancement of reconstructed 3D models using coplanarity and constraints*. Van Gool L. (ed.) DAGM 2002. LNCS, vol. 2449, Springer, Heidelberg, 2002, 34–41.
- [8] Grau O., *A Scene Analysis System for the Generation of 3-D Models*. Proceedings IEEE International Conference on Recent Advances in 3D Digital Imaging and Modeling (3DIM 1997), Ottawa, Canada, May 1997, 221–228.
- [9] Früh C., Zakhor A., *3D Model Generation for Cities Using Aerial Photograph and Ground Level Laser Scans*. Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR 2001), Kauai, Hawaii, USA, December 2001.
- [10] Georgiev A., Allen P.K., *Localization Methods for a Mobile Robot in Urban Environments*. IEEE Transaction on Robotics and Automation (TRO) (2005), 2004, 851–864.
- [11] Hall E.L., Tio J.B.K., McPherson C.A., Sadjadi, *Measuring curver supfaces for robot vision*. Comput., J. 15, 1982, 42–54.
- [12] Toscani G., *Systèmes de calibration des perception du mouvement en vision artificielle*. Université Paris Sud, 1987 (Ph.D Thèses).
- [13] Faergas O.D., Toscani G., *The calibration problem for stereo*. Proceedings of the IEEE Computer Vision and Patter Recognition, 1986, 15–20.
- [14] Tsai R.Y., *A versatile camera calibration technique for high accuracy machine vision metrology using off the shelf TV cameras and lenses*. IEEE Int. J. Robot Atuomat., RA-3 (1987) 323–344.
- [15] Weng J., Cohen P., Herniou M., *Camera calibration with distortion models and accuracy evaluation*. IEEE Trans Pattern Anal. Mach. Intell., 14, 1992, 965–980.
- [16] Duda R.O., Hart P.E., *Use of the Hough Transformation to Detect Lines and Curves in Pictures*. Comm. ACM, vol. 15, January 1972, 11–15.
- [17] Leandro A.F., Fernandes., Oliveira M.M., *Real-time line detection through an improved Hough transform voting scheme*. Pattern Recognition, 41, 2008, 299–314.
- [18] Ballard D.H., *Generalizing the Hough transform to detect arbitrary shapes*. Pattern Recognition, 13 (2), 1981, 111–122.
- [19] Li Z.N., Yao B., Tong F., *Linear generalized Hough transform and its parallelization*. Image Vision Comput., 11, 1993, 11–24.