

# Wprowadzenie do pomiaru głębi obrazu za pomocą stereoskopowego układu kamer

An introduction to scene depth measurement with stereoscopic set of cameras

## Streszczenie

W artykule zaprezentowano podstawy widzenia stereoskopowego zarówno w układzie wzrokowym człowieka, jak również w systemie komputerowym wyposażonym w parę kamer. Przedstawione zostały matematyczne podstawy widzenia stereoskopowego, które z wykorzystaniem dwóch przetworników wizyjnych umożliwiają ocenę nie tylko przesunięcia obiektów, ale również ich oddalenia od kamer. Zaprezentowane systemy znajdują zastosowanie również w spawalnictwie do pomiaru parametrów powierzchni jeziorka spawalniczego oraz wyznaczenia stopnia przetopu spoiny.

## Abstract

The paper presents an introduction to stereoscopic vision in the human visual system as well as in computer system equipped in two cameras. Mathematical foundations of the stereoscopic vision are presented which thanks to two cameras allow not only measurement of objects shifts but also their distance to the cameras. Presented systems find application in welding for measurement of parameters of the weld pool and weld penetration.

## Wstęp

W wielu dziedzinach techniki niezbędny jest pomiar pełnej geometrii obserwowanych zjawisk i obiektów. W zależności od warunków oraz wymagań pomiar taki można wykonać za pomocą metod aktywnych, takich jak np. skaner laserowy, bądź też pasywnych, takich jak stereoskopowy układ kamer. W artykule zaprezentowano wstęp do pomiaru głębi obrazów za pomocą stereoskopowego układu kamer, począwszy od omówienia układu widzenia człowieka, aż do krótkiej charakterystyki metod komputerowych. Wykazano, że również w przypadku zastosowań spawalniczych stereoskopowy układ kamer pozwala na m.in. pomiar parametrów powierzchni jeziorka spawalniczego, co z kolei umożliwia wyznaczenie stopnia przetopu spoiny [5, 10]. Rozwój nowoczesnych technik mikroprocesorowych, jak również metod obliczeniowych, umożliwia zastosowanie komputerów do automatycznej analizy obrazów, również w systemach

spawalniczych [6, 7]. Artykuł rozpoczyna rozdział poświęcony układowi widzenia człowieka ze szczególnym uwzględnieniem jego zdolności postrzegania otoczenia. Następnie omówiony został matematyczny model pojedynczej kamery. Zastosowanie dwóch kamer zmienia sytuację jakościowo, umożliwiając percepcję głębi obserwowanej sytuacji. Innymi słowy, wykorzystując dwie kamery, można określić usytuowanie obiektów. W artykule przedstawiono matematyczne podstawy tego procesu. Następnie omówiono komputerowe metody pomiaru mapy głębi powstałej na podstawie obrazów dostarczonych ze stereoskopowego systemu kamer.

## Widzenie stereoskopowe – układ wzrokowy człowieka

Występująca u człowieka oraz większości zwierząt para oczu nie służy wyłącznie do zdublowania tego cennego receptora sygnałów wizyjnych, ale zapewnia również zdolność do głębszego postrzegania obserwowanych obrazów. Efekt ten zwany jest *widzeniem stereoskopowym*.

Dr hab. inż. Bogusław Cyganek – Akademia Górniczo-Hutnicza w Krakowie.

## Odbiór wzrokowy sygnałów wizyjnych

Promienie świetlne przenikają przez soczewkę oka na fotoczulą tkankę nazywaną siatkówką oka, na której wywołują obraz pomniejszony i obrócony [8]. Siatkówka zbudowana jest z kilku warstw komórek nerwowych: czopków oraz pręcików, jak również komórek dwubiegunowych oraz komórek zwojowych. Pręciki znajdują się w części zewnętrznej siatkówki i umożliwiają odbiór natężenia światła, czyli zapewniają jej widzenie monochromatyczne. Czopki z kolei są obecne w centralnej części siatkówki, najwięcej jest ich na tzw. plamce żółtej, i dzięki zawartym w nich barwnikom wrażliwym na różne długości fali umożliwiają widzenie barwne. Największe zagęszczenie czopków występuje w plamce żółtej oka (łac. *macula lutea*). Powoduje to największą rozdzielczość widzenia właśnie w tej części oka. Siatkówka połączona jest z mózgiem nerwem wzrokowym.

Istotnym elementem oka jest tzw. *dołek środkowy oka* (łac. *fovea centralis*), widoczny na rysunku 1. Jest to niewielkie zagłębienie w centrum plamki żółtej siatkówki, które znajduje się dokładnie w osi poziomej oka. U człowieka jest to ok. 3,5 mm od brzoju tarczy nerwu wzrokowego. Obraz obiektu, na którym skupiamy wzrok, jest rzutowany właśnie w to miejsce. Jest to obszar najostrożniejszego widzenia, pokrywający jednak niewielki kąt widzenia (ok. dwa stopnie).

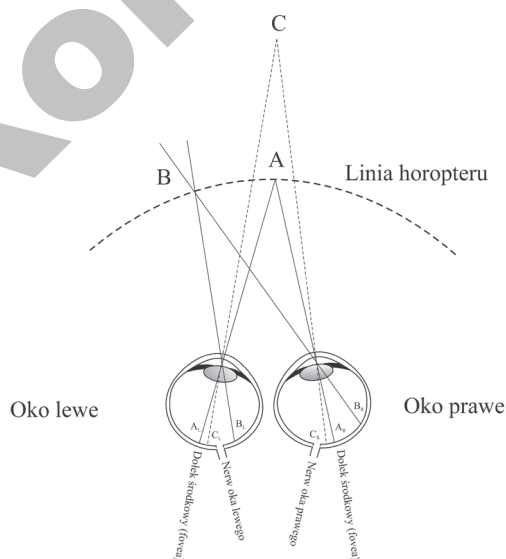
Tak jak wspomniano, para oczu powoduje zmianę jakościową procesu widzenia, gdyż umożliwia detekcję głębi obserwowanej sceny. Rysunek 1 jest ilustracją procesu widzenia stereoskopowego u człowieka.

Świat obserwowany wyłącznie jednym okiem jest płaski. Co prawda jesteśmy zdolni do postrzegania względnego położenia obiektów w przestrzeni dzięki innemu zjawiskom takim jak ich wzajemne przysłanianie – obiekt, który jest przysłonięty przez inny, musi

znajdować się dalej od obserwatora. Niemniej jednak pełnię odbioru zapewnia dopiero para oczu. Podstawowym zjawiskiem, które umożliwia ocenę odległości obserwowanego obiektu, jest różnica w położeniu jego obrazu na ich siatkówkach.

W przypadku skupienia obydwu oczu na punkcie A, jego obraz powstaje w siatkówkach dokładnie w miejscach występowania plamki żółtej (rys. 1). W przypadku innych punktów, położenia ich obrazów powstałych na siatkówkach określamy w stosunku do umiejscowienia odpowiednich żółtych plamek. Przy tym założeniu, dwa punkty obrazowe na siatkówkach odpowiadają sobie, jeżeli są oddalone od odpowiadających im żółtych plamek o jednakową odległość. Jest to w sposób oczywisty spełnione dla punktu A. Z kolei, wyznaczając położenie punktów przestrzennych, których obrazy odpowiadają sobie, otrzymujemy obszar przestrzeni zwany *horopterem*. Wszystkie inne punkty przestrzenne, oprócz punktów należących do horoptera, dają na siatkówkach punkty obrazowe, które nie są wzajemnie odpowiadające. Różnicę ich położenia na siatkówkach nazywamy *dysparycją siatkówkową* (ang. *retinal disparity*) [3]. Stąd też punkty horoptera mają zawsze zerową wartość dysparycji siatkówkowej. Dysparycja siatkówkowa jest wielkością umożliwiającą ocenę odległości od otaczających punktów przestrzeni. I tak, im większa różnica położenia tych obrazów, tym bliżej znajduje się obserwowany obiekt. Z kolei obrazy obiektów będących w znacznym oddaleniu nie wykazują przesunięcia

Na rysunku 2 przedstawiono przykładową parę obrazów stereoskopowych otrzymaną w wyniku wykonania w tym samym momencie zdjęcia dwoma kamerami będącymi w niewielkiej odległości od siebie. Aby zaobserwować efekt głębi, należy pomiędzy obrazami umieścić kartkę papieru, która umożliwi obserwację obrazu lewego wyłącznie lewym okiem i odpowiednio obrazu prawego prawym. Po pewnym czasie wpatrywania następuje „złanie” się dwóch obrazów w jeden z wyraźnie widocznymi obszarami bliższymi obserwatorowi. W dalszej części artykułu opisany jest proces wyznaczania mapy głębi obrazu za pomocą komputera.



Rys. 1. Widzenie stereoskopowe u człowieka  
Fig. 1. Human stereoscopic vision



Rys. 2. Przykład pary obrazów stereoskopowych (część Renault, INRIA C). Aby zaobserwować efekt głębi, należy ustawić pionową kartkę papieru i obserwować tak, aby oko lewe patrzyło wyłącznie na obraz lewy, a prawe na prawy

Fig. 2. A pair of stereoscopic images (part of Renault, INRIA C). The depth effect observed with vertical position of paper sheet and observing the left eye image on the left and right one the right

# Model kamery punktowej

Na rysunku 3 przedstawiono model kamery punktowej (ang. *pin-hole camera*). Charakterystyczne jest występowanie trzech układów współrzędnych: układu zewnętrznego oznaczonego w, układu kamery oznaczonego c oraz układu płaszczyzny obrazowej kamery II z dwoma wymiarami x-y zgodnymi z liniami skanowania obrazu, tak jak w przypadku sygnału TV.

Punkt  $O_c$  nazywany jest *punktem ogniskowym* lub *centralnym* kamery. Elementy płaszczyzny P kamery o całkowitych współrzędnych nazywamy *pixelsami obrazu*. Punkt główny  $o = (o_x, o_y)$  (ang. *principal point*) jest wynikiem rzutu punktu środkowego  $O_c$  układu współrzędnych kamery na płaszczyznę II w kierunku zgodnym z osią  $Z_c$ . Odległość f płaszczyzny obrazowej II od punktu  $O_c$  nazywana jest *długością ogniskową*, natomiast oś optyczną wyznacza prosta  $O_c o$ . Fizyczne wymiary piksela określają parametry  $s_x$  oraz  $s_y$ . Załóżmy, że w układzie współrzędnych kamery punkty P oraz jego obraz p mają następujące współrzędne

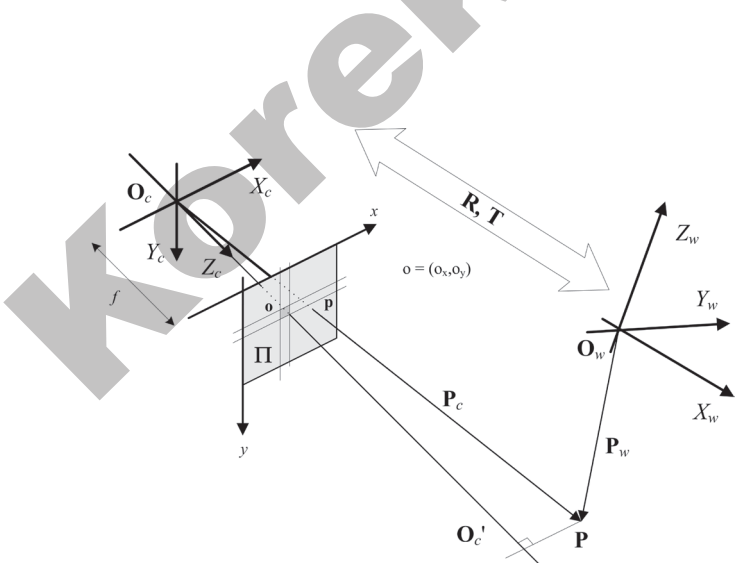
$$\begin{aligned} P &= (X, Y, Z) \\ p &= (x, y, z) \end{aligned} \tag{1}$$

Korzystając z podobieństwa trójkątów  $\Delta O_c p o$  oraz  $\Delta O_c P O_c'$  oraz przyjmując  $z = f$ , otrzymujemy następującą zależność

$$x = f \frac{X}{Z}, \quad y = f \frac{Y}{Z}, \quad z = f \tag{2}$$

Równanie (2) opisuje model kamery perspektywicznej z obiektywem punktowym.

Położenie punktu P w układzie kamery opisuje wektor  $P_c$ , natomiast w układzie zewnętrznym wektor  $P_w$  (rys. 3). Z kolei punkt p jest rzutem perspektywicznym, z centrum w punkcie  $O_c$ , punktu P na płaszczyznę obrazową II. Wzajemną zależność współrzędnych punktu



Rys. 3. Punktowy model pojedynczej kamery  
Fig. 3. Spot single camera model

P w tych dwóch układach współrzędnych opisuje poniższe równanie [1, 2]

$$P_c = R(P_w - T) \tag{3}$$

gdzie  $R$  jest macierzą obrotu, natomiast  $T = O_w - O_c$  macierzą przesunięcia pomiędzy układami współrzędnych w oraz c. Z kolei współrzędne punktu p na płaszczyźnie obrazu  $x_u$  oraz  $y_u$  oraz współrzędne x i y w lokalnym układzie kamery łączy zależność

$$\begin{aligned} x &= (x_u - o_x) s_x \\ y &= (y_u - o_y) s_y \end{aligned} \tag{4}$$

Po podstawieniu związków (3) i (4) do zależności (2) otrzymujemy liniowe równanie opisujące model kamery perspektywicznej z punktowym obiektywem [1, 2]

$$x = MX \tag{5}$$

gdzie  $x$  jest punktem obrazu o współrzędnych jednorodnych  $[x_1, x_2, x_3]^T$ ,  $X$  punktem przestrzeni trójwymiarowej o współrzędnych jednorodnych  $[X_1, X_2, X_3, X_4]^T$ . Użycie współrzędnych jednorodnych powoduje zastąpienie dzielenia występującego w równaniu (2) dodaniem dodatkowego składnika, gdyż związane są one następującą zależnością

$$x \equiv \frac{x_1}{x_3}, \quad y \equiv \frac{x_2}{x_3} \tag{6}$$

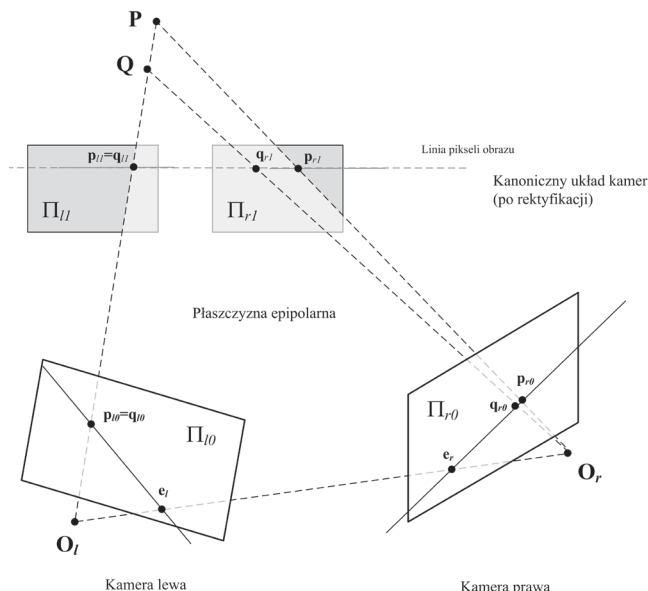
Równanie (5) opisuje transformację z przestrzeni rzutowej  $\mathcal{P}^3$  na płaszczyznę rzutową  $\mathcal{P}^2$ . Transformacja ta przekształca wszystkie punkty leżące na prostej w dokładnie jeden punkt-obraz  $x$  na płaszczyźnie rzutowej  $\mathcal{P}^2$ . Prosta ta jest jednoznacznie określona przez punkt centralny  $O_c$  oraz punkt przestrzeni  $\mathcal{P}^3$ . Stąd też transformacja taka jest jednoznaczna w stosunku do przekształcenia prostej w punkt. Dlatego równanie (5) można przedstawić ogólnie w postaci:

$$cx = MX \tag{7}$$

gdzie dodatkowy w stosunku do (5) parametr c jest współczynnikiem skalującym. Bez utraty ogólności, tam gdzie jest to jednoznaczne, przyjmuje się  $s = 1$ , a macierz  $M$  traktuje jako zdefiniowaną z dokładnością do pewnego współczynnika skalującego, efektywnie przyjmując z powrotem postać równania (5). Wprowadzenie do elementów geometrii rzutowej można znaleźć m.in. w publikacjach [1, 2].

Macierz  $M$  opisuje przekształcenie punktu przestrzeni na jego obraz dokonywane przez kamerę. Macierz tę można przedstawić w postaci iloczynu:

$$M = M_i M_e \tag{8}$$



**Rys. 4.** Model stereoskopowego układu kamer w postaci ogólnej oraz kanonicznej. Zewnętrzny punkt  $P$  zostaje zobrazowany w postaci lokalnych punktów  $p_{l0}$  oraz  $p_{r0}$ . Punkty  $O_l$  oraz  $O_r$  wraz z punktem przestrzeni  $P$  definiują płaszczyznę epipolarną

**Fig. 4.** Stereoscopic camera system model in the general and canonical form. External point  $P$  depicted as local points  $p_{l0}$  and  $p_{r0}$ . Points  $O_l$  i  $O_r$  with the  $P$  point defines plane

gdzie

$$M_i = \begin{bmatrix} \frac{f}{s_x} & 0 & o_x \\ 0 & \frac{f}{s_y} & o_y \\ 0 & 0 & 1 \end{bmatrix}_{3 \times 3} \quad (9)$$

jest macierzą parametrów wewnętrznych, natomiast

$$M_e = \begin{bmatrix} R_1 & -R_1 T \\ R_2 & -R_2 T \\ R_3 & -R_3 T \end{bmatrix}_{3 \times 4}, \quad R_i = [r_{i1} \ r_{i2} \ r_{i3}]_{1 \times 3} \quad (10)$$

jest macierzą parametrów zewnętrznych kamery. Są one zdefiniowane jako zbiór parametrów geometrycznych, jednoznacznie określających przekształcenie układu współrzędnych kamery w układ współrzędnych zewnętrznych, tj. za pomocą wektora translacji  $T$  oraz macierzy rotacji  $R$ . Proces znajdowania parametrów macierzy  $M_i$  oraz  $M_e$  nazywamy kalibracją kamery [3, 9].

## Układ dwóch kamer – stereoskopia

Na rysunku 4 przedstawiono proces obserwacji zjawiska zachodzącego w otoczeniu za pomocą dwóch kamer, czyli układu stereoskopowego. Kamery te, jednocześnie obserwując daną scenę, umożliwiają wnioskowanie o oddaleniu punktów przestrzeni.

Każda z widocznych płaszczyzn  $\Pi_i$  wraz z punktem centralnym rzutowania  $O_i$  tworzy kamerę z obiektywem punktowym omawianą w poprzednim rozdziale.



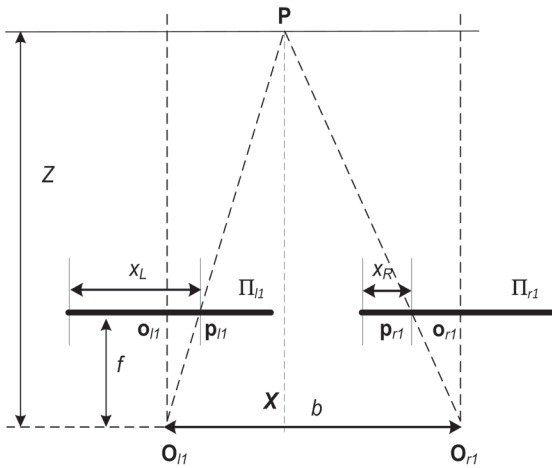
**Rys. 5.** Przykładowy system kamer przemysłowych w układzie stereoskopowym

**Fig. 5.** Example for industrial camera in the stereoscopic configuration system

Prosta  $O_l O_r$  łącząca centra rzutowania, wyznacza linię bazową. Punkty jej przecięcia z płaszczyznami kamer nazywane są *punktami epipolarnymi*. W szczególnym przypadku, gdy linia bazowa jest równoległa do płaszczyzn kamer, odpowiadający punkt epipolarny jest tzw. *punktem w nieskończoności* [3]. Na rysunku 5 przedstawiono przykładowy system kamer przemysłowych w układzie stereoskopowym. Możliwe jest również zbudowanie układu stereoskopowego z wykorzystaniem jednej kamery wyposażonej w odpowiedni zestaw soczewek i lusterek (układ katadioptryczny) [1, 2]. Zauważmy, że punkty  $P$  oraz  $Q$  są rzutowane do jednego obrazu na płaszczyźnie kamery lewej  $\Pi_{l0}$ . Stąd też wyznaczenie ich oddalenia jest możliwe wyłącznie dzięki zastosowaniu drugiej kamery, oznaczonej jako  $\Pi_{r0}$ . Obrazy punktów  $P$  oraz  $Q$  na płaszczyźnie  $\Pi_{r0}$  są różne. Widzimy też, że wraz z przybliżaniem się punktu  $Q$  do centrum kamery  $O_l$  jego obraz na drugiej płaszczyźnie zbliża się do położenia punktu  $e_r$ . Zauważamy również, że punkty obrazu znajdują się wyłącznie na prostych powstałych w wyniku przecięcia płaszczyzn kamer  $\Pi_{l0}$  oraz  $\Pi_{r0}$  z płaszczyzną określoną przez centra rzutowania  $O_l$  i  $O_r$  oraz rzutowany punkt przestrzeni  $P$ , zwaną *płaszczyzną epipolarną*. Jest to istotny rezultat, gdyż ogranicza położenie odpowiadających sobie punktów obrazu wyłącznie do linii epipolarnych. Widzimy więc, że położenie punktu przestrzeni oraz jego obrazów na płaszczyznach kamer jest ściśle określone przez geometrię epipolarną danego układu kamer. Dla wszystkich rzutowanych punktów wzajemne położenie odpowiadających sobie ich punktów obrazu określa jedna macierz zwana *macierzą fundamentalną*  $F$ , dla której zachodzi zależność [1÷3]:

$$\bar{p}_{r0}^T F \bar{p}_{l0} = 0 \quad (11)$$

gdzie  $p_{l0}$  oraz  $p_{r0}$  są punktami obrazowymi punktu  $P$ , odpowiednio na płaszczyźnie obrazowej kamery lewej oraz prawej, których współrzędne wyrażone są w lokalnym układzie współrzędnych związanym z odpowiednimi płaszczyznami kamer  $p_{l0}$  oraz  $p_{r0}$ .



Rys. 6. Kanoniczny układ stereoskopowy  
Fig. 6. Canonical stereoscopic configuration

Jak już wspomniano, w celu znalezienia położenia punktu  $P$  należy odnaleźć jego obrazy na płaszczyznach  $\Pi_{l0}$  oraz  $\Pi_{r0}$ . Jednak w praktyce proces ten jest trudny i często niejednoznaczny. Znaczącym ułatwieniem jest fakt, że poszukiwania odpowiadających sobie punktów można ograniczyć wyłącznie do przeszukania punktów linii epipolarnych. Niemniej w przypadku dowolnego układu kamer położenie tych linii nie zawsze jest znane *a priori*. Sytuacja jest znacznie łatwiejsza, jeżeli płaszczyzny kamer ułożone są w ten sposób, że osie kamer są równoległe, a krawędzie płaszczyzn współliniowe, tak jak w przypadku kamer  $\Pi_{l1}$  oraz  $\Pi_{r1}$  (rys. 4). Tego typu układ nazywamy *kanonicznym układem kamer*, którego przekrój przedstawiono na rysunku 6. Może on być wynikiem wstępnego ułożenia kamer lub też kanoniczny obraz może być otrzymany z obrazu ogólnego w wyniku przeprowadzenia tzw. *procesu rektyfikacji obrazów*.

Korzystając z właściwości podobieństwa trójkątów  $\Delta p_{l1} o_{l1} O_{l1}$  oraz  $\Delta P X O_{l1}$ , jak również  $\Delta p_{r1} o_{r1} O_{r1}$  oraz  $\Delta P X O_{r1}$  (rys. 6), otrzymujemy następujące zależności definiujące dysparycję horyzontalną  $D_x(p_{l1}, p_{r1})$  pomiędzy parą punktów  $p_{l1}$  oraz  $p_{r1}$

$$D_x(p_{l1}, p_{r1}) = p_{r1} - p_{l1} = x_L - x_R = \frac{bf}{Z} \quad (12)$$

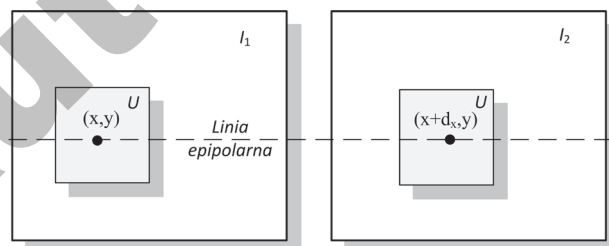
gdzie punkty  $p_{l1}$  oraz  $p_{r1}$  są obrazami pewnego punktu  $P$ ,  $p_{l1}$  oraz  $p_{r1}$  są ich współrzędnymi w kierunku osi  $x$  (w kierunku  $y$  są równe 0),  $b$  oznacza odległość bazową między kamerami,  $f$  jest ogniskową kamer, natomiast  $Z$  jest szukaną odległością punktu  $P$  od linii bazowej, czyli głębokością sceny. Szukaną wartość  $Z$  w układzie kanonicznym obliczamy na podstawie znalezionej przesunięcia horyzontalnego  $D_x(p_{l1}, p_{r1})$  pomiędzy punktami obrazowymi  $p_{l1}$  oraz  $p_{r1}$  nieznanego punktu  $P$ , jak poniżej:

$$Z = \frac{bf}{D_x(p_{l1}, p_{r1})} \quad (13)$$

W przypadku gdy  $D_x \rightarrow 0$ , przyjmujemy że obserwowany punkt  $P$  znajduje się w nieskończoności, czyli  $Z \rightarrow \infty$ .

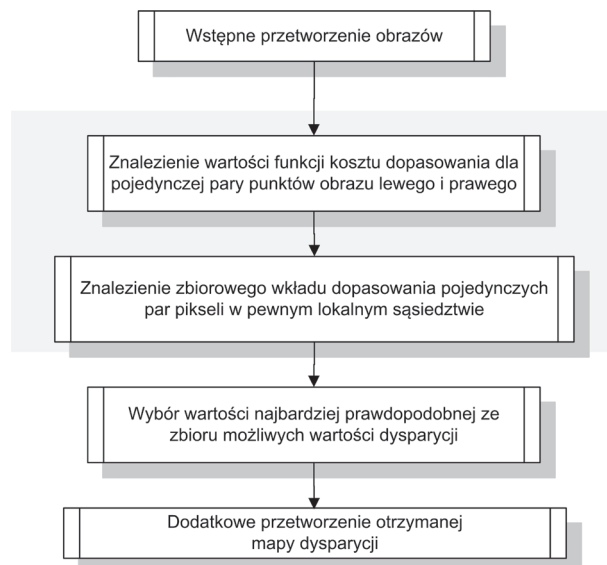
## Komputerowe obliczanie głębi obrazu

Podstawy widzenia stereoskopowego zostały zarysowane w poprzednim rozdziale. W przypadku systemu komputerowego wyposażonego w dwie kamery obliczanie mapy głębi polega głównie na znalezieniu wszystkich odpowiadających sobie punktów obrazu dla nieznanego punktu przestrzeni, a następnie wyznaczeniu ich wzajemnego przesunięcia horyzontalnego  $D_x$  występującego we wzorach (12) oraz (13). Jak już wspomniano, proces ten nie jest ani łatwy, ani też jednoznaczny. To znaczy, mając pewien punkt na płaszczyźnie kamery, przeważnie nie wiemy, jakiego obiektu jest on obrazem, ani też nie wiemy, gdzie znajduje się odpowiadający mu obraz z drugiej kamery. Pewnym ułatwieniem jest fakt, że obraz taki znajduje się wyłącznie na odpowiedniej linii epipolarnej, która w przypadku kanonicznego układu kamer jest zgodna z linią skanowania pikseli w kamerze. Zdarza się często, że dany punkt przestrzeni widoczny jest wyłącznie w płaszczyźnie jednej kamery, podczas gdy w drugiej nie występuje ze względu na ułożenie geometryczne lub też



Rys. 7. Poszukiwanie odpowiadających sobie punktów na linii epipolarnej z wykorzystaniem wartości pikseli w pewnym obszarze  $U$  otaczającym punkt  $(x, y)$

Fig. 7. The search for corresponding points on the line using pixel values in an  $U$  area surrounding the point  $(x, y)$



Rys. 8. Algorytm obliczania mapy głębi z pary obrazów stereoskopowych

Fig. 8. The algorithm for calculation of depth maps from a pair of stereoscopic images

przysłonięcia. Innym częstym przypadkiem jest znalezienie fałszywego dopasowania punktów obrazowych. Powstaje tutaj bardzo istotne pytanie – w jaki sposób ocenić, czy obserwowane punkty na płaszczyznach kamer faktycznie są obrazami jakiegoś rzeczywistego punktu przestrzeni? Jedną z możliwości jest zastosowanie pewnej miary porównania danych dwóch punktów obrazu. Jednakże same wartości sygnału intensywności lub też koloru pojedynczego piksela mogą być niewystarczające do jednoznacznego określenia, czy punkty te są obrazami tego samego punktu przestrzeni. W celu poprawy jakości odpowiedzi można wykorzystać również wartości pikseli sąsiadujących, jak również dodatkowe parametry geometryczne obserwowanej sceny. Proces ten zilustrowano na rysunku 7.

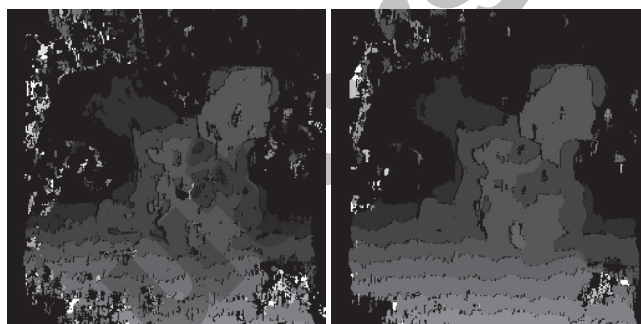
Obrazy można porównać za pomocą wielu miar. Jedną z takich miar, która jest odporna na lokalne wariacje sygnału intensywności, jest  $S_{CoVar}$  dana poniższym wzorem

$$S_{CoVar}(I_1, I_2, i, j, U, d_x, d_y) = \frac{\hat{\alpha}_{(i,j) \in U} (I_1(x+i, y+j) - \bar{I}_1(x, y)) \times (I_2(x+d_x+i, y+d_y+j) - \bar{I}_2(x+d_x, y+d_y))}{\sqrt{\hat{\alpha}_{(i,j) \in U} (I_1(x+i, y+j) - \bar{I}_1(x, y))^2 \times \hat{\alpha}_{(i,j) \in U} (I_2(x+d_x+i, y+d_y+j) - \bar{I}_2(x+d_x, y+d_y))^2}} \quad (14)$$

gdzie  $I_k(x, y)$  jest wartością intensywności piksela w punkcie  $(x, y)$  k-tego obrazu,  $\bar{I}_k(x, y)$  jest średnią

wartością intensywności dla k-tego obrazu w otoczeniu  $U$  wokół punktu  $(x, y)$ ,  $d_x$  oraz  $d_y$  oznaczają wartości względnego przesunięcia (dysparycji), natomiast  $i, j$  są indeksami. Na rysunku 8 przedstawiono prosty algorytm znajdowania mapy głębi na podstawie pary obrazów stereoskopowych. Dalsze informacje na temat procesu korelacji stereoskopowej zawierają liczne publikacje, np. książki [1, 2]. Z kolei, na rysunku 9 przedstawiono wyniki działania tego algorytmu dla pary pokazanej na rysunku 2, dla różnych wielkości otoczenia  $U$  we wzorze (14).

Jaśniejsze miejsca na rysunku 9 oznaczają punkty znajdujące się bliżej obserwatora, ale poprawne wartości znajdują się w miejscach o silnej teksturze.



Rys. 9. Komputerowo obliczona mapa głębi w wyniku korelacji punktów pomiędzy parą obrazów przedstawionych na rysunku 2

Fig. 9. Computer calculated depth map as the points correlation in the pair of images shown on the fig. 2

## Wnioski

W artykule zaprezentowane zostały podstawy widzenia stereoskopowego, zarówno w ludzkim układzie wzrokowym człowieka, jak również w komputerowym systemie wyposażonym w dwie kamery. Przedstawione zostały matematyczne modele kamery, jak również podstawy geometrii epipolarnej.

Zaprezentowany został przegląd algorytmów obliczania mapy głębi, jak również wspomniane zostały wybrane aplikacje. Między innymi, stereoskopowe metody pomiaru głębi mogą być wykorzystane w spawalnictwie do monitorowania jeziora spawalniczego celem wyznaczenia stopnia wtopienia spoiny.

## Literatura

- [1] Cyganek B.: Komputerowe przetwarzanie obrazów trójwymiarowych (in Polish) EXIT, 2002.
- [2] Cyganek B., Siebert J.P.: An Introduction to 3D Computer Vision Techniques and Algorithms. Wiley, 2009.
- [3] Hartley R.I., Zisserman A.: Multiple View Geometry in Computer Vision, 2nd edn, Cambridge University Press, 2003.
- [4] Howard I.P., Rogers B.J.R.: Binocular Vision and Stereopsis. Oxford University Press, 1995.
- [5] Kim Y. S. and Eagar T. W.: Analysis of metal transfer in gas metal arc welding. Welding Journal 1993, 72, s. 269-278.
- [6] Tadeusiewicz R.: New Trends in Neurocybernetics. Computer Methods in Materials Science, Vol. 10, 2010, No. 1, s. 1-7.
- [7] Tadeusiewicz R.: Place and Role of Intelligent Systems in Computer Science. Computer Methods in Materials Science, Vol. 10, 2010, No. 4, s. 193-206.
- [8] Wandell B.A.: Foundations of Vision. Sinauer Associates Publishers Inc. 1995.
- [9] Zhang Z.: A Flexible New Technique for Camera Calibration. Technical Report MSR-TR-98-71. Microsoft Research, Microsoft Corporation, 1999.
- [10] Zhao C. X., Richardson I. M., Kenjeres S., Kleijn C. R., Saldi Z.: A stereo vision method for tracking particle flow on the weld pool surface. Journal of Applied Physics 105, 2009.

## Podziękowanie

Artykuł ten powstał dzięki wsparciu finansowemu w roku 2012 projektu badawczego Narodowego Centrum Nauki nr DEC-2011/01/B/ST6/01994.