

Stanisław Fuksa*, Witold Byrski*

Czteropunktowa metoda identyfikacji transformacji stereowizyjnej**

1. Wprowadzenie

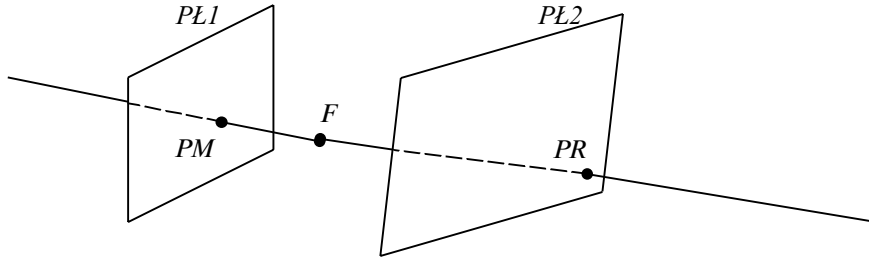
Problem wykorzystywania kamer do oceny położenia obiektów w przestrzeni trójwymiarowej jest jednym z podstawowych, klasycznych w robotyce. Istnieje bogata literatura tego przedmiotu. Każdy sposób uzyskiwania położenia przestrzennego punktu zawiera w sobie element wzorcowania kamer i wykorzystuje prosty zbiór relacji. Proponowana przez nas transformacja płaska i jej identyfikacja cechuje się pewną prostotą operacji matematycznych wymaganych przy identyfikacji i minimalną ilością założeń co do jakości sprzętu optycznego oraz precyzji pomiarów. Wykorzystywana czteropunktowa metoda kalibracji prowadzi do równań liniowych na nieznane współczynniki transformacji i nie wymaga rygorystycznych założeń odnośnie do ortogonalności lub jednakowości skal osi układu współrzędnych. Transformacja płaska będzie wykorzystana do konstrukcji transformacji stereowizyjnych, to jest takich, które w oparciu o obrazy punktu na monitorach dwóch kamer, pozwalają określić jego rzeczywiste położenie w zewnętrznym układzie współrzędnych.

2. Transformacja „płaska”

Załóżmy, że w przestrzeni trójwymiarowej wyróżnione są dwie dwuwymiarowe płaszczyzny $PL1$ i $PL2$. Płaszczyznę $PL1$ utożsamiamy z płaszczyzną elementu światłoczułego CCD kamery, a płaszczyzna $PL2$ jest dowolną, leżącą w tej przestrzeni. Niech punkt $F \in R^3$ pełni rolę ogniska układu optycznego kamery. Przez transformację płaską będziemy rozumieć odwzorowanie przyporządkowujące punktowi $PM \in PL1$ punkt $PR \in PL2$ przez jednokładność o środku w punkcie F . Inaczej, punkt PR jest punktem przecięcia płaszczyzny $PL2$ przez prostą (promień światła) przechodzącą przez ognisko F i punkt PM z $PL1$ (rys. 1).

* Katedra Automatyki, Wydz.EAiE AGH, Kraków.

** Badania i publikacja finansowane z umowy 10.10.120.40 AGH/MNiI



Rys. 1. Geometria układu: kamera–płaszczyzna rzeczywista $PŁ2$

Oznaczmy transformację płaską przez $TRPŁ$:

$$PR = TRPŁ(PM), \quad PM \in R^2, \quad PR \in R^2 \quad (1)$$

Rozumiemy to tak, że na płaszczyźnie $PŁ1$ wprowadzono pewien układ współrzędnych, w którym wektor PM ma dwie składowe

$$PM = \begin{bmatrix} XM \\ YM \end{bmatrix} \in R^2 \quad (2)$$

Powyższy układ współrzędnych XM i YM wiążemy w sposób naturalny z osiami x i y płaszczyzny przetwornika CCD, czyli liniami i kolumnami pikseli (piksel jednostka długości). Podobnie na płaszczyźnie $PŁ2$ wprowadzamy układ współrzędnych, w którym punkt PR wyraża się jako

$$PR = \begin{bmatrix} XR \\ YR \end{bmatrix} \in R^2 \quad (3)$$

Podkreślmy tutaj następujące założenia, precyzujące sytuację geometryczną:

- układ współrzędnych na płaszczyźnie $PŁ1$ jest dowolny, tzn. nie wymaga się ortogonalności osi, ani ich jednakowych skal;
- układ współrzędnych na płaszczyźnie $PŁ2$ jest także dowolny w sensie jak wyżej;
- oś optyczna soczewki nie musi być prostopadła do płaszczyzny światłoczułej;
- nie znamy *a priori* parametrów optycznych kamery.

2.1. Postać transformacji „płaskiej”. Oznaczenia współczynników

Proste rozważania geometryczne określają ogólną formę transformacji płaskiej.

$$TRPŁ(PM) = PR = \begin{bmatrix} XR \\ YR \end{bmatrix} = \frac{1}{\alpha \cdot XM + \beta \cdot YM + 1} \cdot \begin{bmatrix} a_1 \cdot XM + b_1 \cdot YM + c_1 \\ a_2 \cdot XM + b_2 \cdot YM + c_2 \end{bmatrix} \quad (4)$$

Widać, że 8 współczynników: $a_1, b_1, c_1, a_2, b_2, c_2, \alpha, \beta$ definiuje konkretną transformację płaską. Wprowadźmy także oznaczenie wektorowe:

$$WSP \in R^8, \quad WSP^T = [a_1, b_1, c_1, a_2, b_2, c_2, \alpha, \beta] \quad (5)$$

Zwróćmy uwagę, że współczynniki WSP odzwierciedlają własności optyczno-geometryczne kamery (montaż soczewki i przetwornika do obudowy, ogniskowa). Odzwierciedlają także sytuację zewnętrzną, tzn. położenie płaszczyzny $PL2$, wraz ze swoim układem współrzędnych względem kamery.

2.2. Kalibracja kamery (wzorcowanie, identyfikacja $TRPL$) – metoda czteropunktowa

Dla identyfikacji 8 współczynników transformacji wybierzmy na płaszczyźnie $PL2$ cztery punkty, z których żadne trzy nie leżą na jednej prostej: PR_1, PR_2, PR_3, PR_4 .

Dokonyjemy pomiarów położenia tych punktów na płaszczyźnie $PL1$ (chipu CCD, monitora kamery) i oznaczamy je odpowiednio PM_1, \dots, PM_4 .

Można stworzyć układ 8 równań o 8 niewiadomych współczynnikach WSP :

$$PR_i = TRPL(PM_i), \quad i = 1, 2, 3, 4 \quad (6)$$

Powyższy układ równań wydaje się na pierwszy rzut oka nieliniowy. Jednak po wymnożeniu we wzorze (4) obydwu stron przez czynnik liczbowy $(\alpha \cdot XM_i + \beta \cdot YM_i + 1)$

$$(\alpha \cdot XM_i + \beta \cdot YM_i + 1) \cdot PR_i = (\alpha \cdot XM_i + \beta \cdot YM_i + 1) \cdot TRPL(PM_i)$$

otrzymuje się układ równań liniowych ze względu na współczynniki WSP .

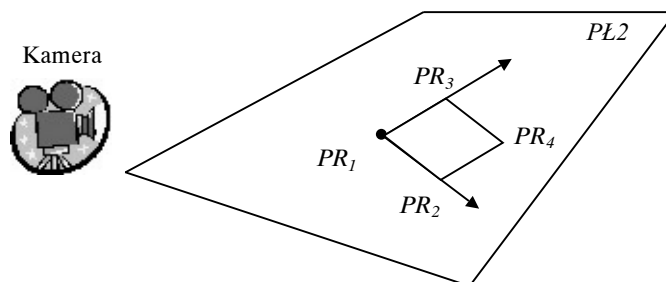
Ma on następującą postać macierzową z macierzą kwadratową 8×8

$$\begin{bmatrix} XR_1 \\ YR_1 \\ \vdots \\ XR_4 \\ YR_4 \end{bmatrix} = \begin{bmatrix} XM_1 & YM_1 & 1 & 0 & 0 & 0 & -XM_1 \cdot XR_1 & -YM_1 \cdot XR_1 \\ 0 & 0 & 0 & XM_1 & YM_1 & 1 & -XM_1 \cdot YR_1 & -YM_1 \cdot YR_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ XM_4 & YM_4 & 1 & 0 & 0 & 0 & -XM_4 \cdot XR_4 & -YM_4 \cdot XR_4 \\ 0 & 0 & 0 & XM_4 & YM_4 & 1 & -XM_4 \cdot YR_4 & -YM_4 \cdot YR_4 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ b_1 \\ c_1 \\ a_2 \\ b_2 \\ c_2 \\ \alpha \\ \beta \end{bmatrix} \quad (7)$$

Opiszmy bliżej wybór tych 4 punktów pomiarowych. Leżą one wszystkie na płaszczyźnie $PL2$ i żadne trzy z nich nie leżą na jednej prostej. To ostatnie gwarantuje, że układ równań (7) jest jednoznacznie rozwiązywalny. Można dodatkowo wybrać w specjalny sposób układ tych punktów – jako narożniki kwadratu.

$$PR_1 = [0 \ 0]^T, \quad PR_2 = [1 \ 0]^T, \quad PR_3 = [0 \ 1]^T, \quad PR_4 = [1 \ 1]^T \quad (8)$$

Wybór taki pozwala nieco uprościć postać macierzy w (7), a przede wszystkim łatwo ukierunkować osie i środek lokalnego układu współrzędnych na płaszczyźnie $P\mathcal{L}2$, (jak na rys. 2.)



Rys. 2. Układ: kamera – 4 punkty kalibrujące

W celu określenia 4 punktów wzorcujących można wykorzystać przygotowany płaski wzorec kwadratu z zaznaczonymi narożami. W takiej sytuacji kalibrowanie transformacji płaskiej można wykonać w pięciu etapach:

1. kładziemy wzorec kwadratu na afinicznej na razie płaszczyźnie $P\mathcal{L}2$, tam gdzie chcemy zdefiniować lokalny na $P\mathcal{L}2$ układ współrzędnych, wiążąc pewne punkty płaszczyzny z wektorami naroży PR_i ;
2. wykonujemy zdjęcie wzorca w takim ustawieniu kamery, w jakim będzie ona pracowała w przyszłości;
3. odzyskujemy, poprzez obróbkę obrazu, położenia monitorowe naroży kwadratu PM_i , $i = 1, \dots, 4$;
4. przeprowadzamy identyfikację współczynników WSP (kalibrację) według formuły (8);
5. usuwamy teraz wzorec kwadratu z pola widzenia kamery (o ile jest to konieczne i o ile pozycja kamery w przyszłości nie zostanie zmieniona).

Znalezione współczynniki reprezentują informację o wzajemnym położeniu kamery i płaszczyzny $P\mathcal{L}2$.

2.3. Wielopunktowa kalibracja transformacji „płaskiej”

W poprzednim podrozdziale opisano czteropunktową metodę kalibracji transformacji płaskiej. Można ją bez trudu rozbudować na większą liczbę punktów pomiarowych $n > 4$, tzn. na ciągu punktów PR_i , $i = 1, \dots, n$, i ich odpowiednich obrazów monitorowych PM_i , $i = 1, \dots, n$. Ma to na celu zwiększenie precyzji określenia współczynników transformacji „płaskiej”. Stwarza to również możliwość określenia rodziny transformacji płaskich w przypadku, gdy płaszczyzna $P\mathcal{L}2$ jest zastąpiona powierzchnią utworzoną przez rodzinę figur płaskich. Nawet wtedy, gdy powierzchnia jest dość dowolna, można zawsze aproksymować ją za pomocą rodziny figur płaskich, co umożliwi opis przestrzenny tej powierzchni w oparciu o rodzinę stosownych transformacji „płaskich” i zdjęcie monitorowe układu punktów.

W wypadku kalibracji jednej transformacji płaskiej i danych w postaci n par punktów PR_i i PM_i , można zaproponować standardową metodę najmniejszych kwadratów dla estymacji wektora współczynników WSP . Rozpatrujemy mianowicie rodzinę równań liniowych w ilości $2n$ związanych z postacią (6) i ośmiowymiarowym wektorem WSP .

$$PR_i = TRP\mathcal{L}(PM_i), \quad i = 1, \dots, n \quad (9)$$

System tych równań liniowych opisuje się za pomocą macierzy prostokątnej C o $2n$ wierszach i 8 kolumnach

$$\begin{bmatrix} XR_1 \\ YR_1 \\ \vdots \\ XR_n \\ YR_n \end{bmatrix} = \begin{bmatrix} XM_1 & YM_1 & 1 & 0 & 0 & 0 & -XM_1 \cdot XR_1 & -YM_1 \cdot XR_1 \\ 0 & 0 & 0 & XM_1 & YM_1 & 1 & -XM_1 \cdot YR_1 & -YM_1 \cdot YR_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ XM_n & YM_n & 1 & 0 & 0 & 0 & -XM_n \cdot XR_n & -YM_n \cdot XR_n \\ 0 & 0 & 0 & XM_n & YM_n & 1 & -XM_n \cdot YR_n & -YM_n \cdot YR_n \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ b_1 \\ c_1 \\ a_2 \\ b_2 \\ c_2 \\ \alpha \\ \beta \end{bmatrix} = C \cdot WSP \quad (10)$$

Znana postać rozwiązania minimalno-normowego dana jest wzorem

$$WSP = [C^T C]^{-1} C^T \cdot H,$$

gdzie wektor H jest określony

$$H = \begin{bmatrix} XR_1 \\ YR_1 \\ \vdots \\ XR_n \\ YR_n \end{bmatrix}.$$

Przejdźmy teraz do przypadku, gdy zamiast jednej płaszczyzny mamy powierzchnię aproksymowaną rodziną płaszczyzn (figur płaskich). Oznaczmy przez PL_i , $i = 1, \dots, k$, owe figury płaskie. Załóżmy, że na każdej z figur płaskich mamy co najmniej 4 punkty wzorcujące. Można więc utworzyć rodzinę transformacji płaskich $TRP\mathcal{L}_i$ wiążących położenia monitorowe PM punktu z jego rzeczywistym położeniem PR_i na fragmencie PL_i powierzchni. Sklejanie poszczególnych lokalnych układów współrzędnych na fragmentach PL_i tak, by móc odzyskiwać ich rzeczywiste położenia, wymaga zastosowania odrębnego algorytmu.

3. Transformacja typu „stereo” wykorzystująca transformacje „płaskie”

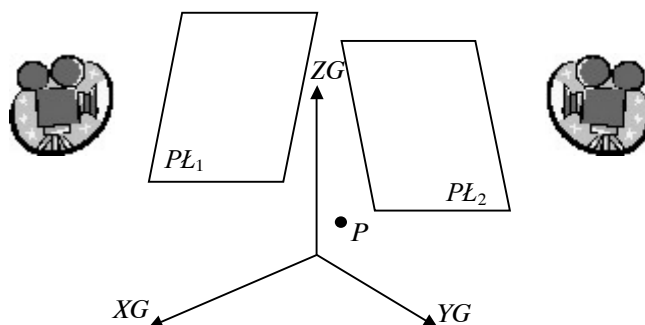
Przez transformacje typu „stereo” rozumiemy odwzorowania z położeń monitorowych punktu widzianego na dwóch monitorach kamer w jego położenie przestrzenne w pewnym zewnętrznym układzie współrzędnych.

Przyjmijmy oznaczenia prostej przechodzącej przez dwa punkty $W_1, W_2 \in R^3$ jako zbioru

$$\text{prosta}(W_1, W_2) = \{x \in R^3: x = W_1 + t \cdot (W_2 - W_1), \quad t \in R^1\} \quad (11)$$

Stereotransformacja wymaga sprecyzowania sytuacji geometrycznej i sprzętowej, z którą jest związana (rys. 3). W tym wypadku odnosi się to do elementów:

- dwie kamery Kam1, Kam2;
- dwie płaszczyzny nierównoległe, zewnętrzne $PŁ_1, PŁ_2$, obie obserwowane przez każdą kamerę i wyposażone we własne lokalne układy współrzędnych;
- globalny zewnętrzny trójwymiarowy układ współrzędnych XG, YG, ZG ;
- punkt P , którego położenie będzie odtwarzane przez transformację stereowizyjną.



Rys. 3. Sytuacja stereowizyjna

Niech

$$TRPŁ_k^i, \quad i, k = 1, 2 \quad (12)$$

oznacza transformację płaską związaną z kamerą o numerze k i lokalnym układem współrzędnych na płaszczyźnie o numerze i .

Sposób postępowania prowadzący do uzyskania stereotransformacji, scharakteryzujemy poniżej.

- Identyfikujemy cztery transformacje płaskie za pomocą stosownego układu co najmniej czterech punktów na każdej płaszczyźnie.

- Ustalamy związki między lokalnymi płaskimi układami współrzędnych na PL_1 i PL_2 , a ich położeniami w układzie globalnym. Oznaczmy przez $B_i: R^2 \rightarrow R^3$ odwzorowania, które wektorowi dwuwymiarowemu $PR_i \in PL_i$ przyporządkowuje jego położenie PG_i w układzie globalnym

$$PG_i = B_i(PR_i) \quad (13)$$

Zakładamy, że odwzorowania B_i związane z położeniem płaszczyzn PL_i względem globalnego układu współrzędnych są znane.

- Punkt P jest widoczny na monitorze kamery 1 jako PM_1 , a na monitorze kamery 2 jako PM_2 . Z położenia monitorowego PM_1 kamery 1, możemy wydedukować w oparciu o transformację płaską, $TRPL_1^1$ położenie punktu PR_1^1 na płaszczyźnie PL_1 . Z tym samym punktem monitorowym PM_1 , możemy stowarzyszyć przez transformację płaską $TRPL_1^2$ punkt PR_1^2 leżący na płaszczyźnie PL_2 . Analogicznie z obrazem PM_2 punktu P stowarzyszamy poprzez stosowane transformacje punkty PR_2^1 i PR_2^2 . Przypomnijmy, że zgodnie z (12) dolny indeks oznacza numer kamery, a górny numer płaszczyzny.
- Zauważmy, że punkty PR_1^1 i PR_1^2 są widziane na ekranie monitora kamery 1 jako jeden punkt PM_1 . Tak samo punkty PR_2^1 i PR_2^2 są widziane na ekranie monitora kamery 2 jako jeden punkt PM_2 . Oznacza to z jednej strony, że punkt P leży na prostej przechodzącej przez punkty PR_1^1 i PR_1^2 (na tej samej prostej leży ognisko kamery 1). Z drugiej strony punkt P leży też na prostej przechodzącej przez PR_2^1 i PR_2^2 (na tej samej prostej leży ognisko kamery 2).
- Rezultat końcowy w postaci odtworzonego położenia punktu P w układzie globalnym przedstawiamy następująco

$$P = \text{prosta}(B1[TRPL_1^1(PM_1)], B1[TRPL_1^2(PM_1)]) \cap \\ \cap \text{prosta}(B2[TRPL_2^1(PM_2)], B2[TRPL_2^2(PM_2)]).$$

Powyższe wymaga rozwiązania stosownego układu równań liniowych na punkt wspólny prostych.

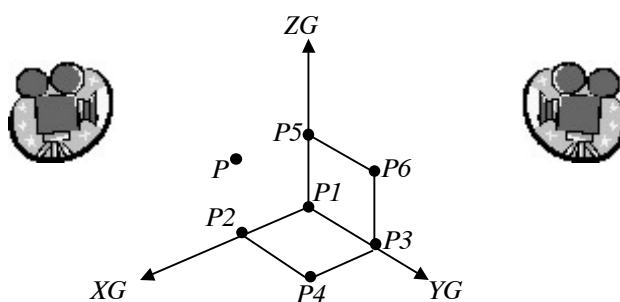
Przyjęcie specjalnego położenia płaszczyzn i układu globalnego stwarza możliwość znacznego uproszczenia. Pokazano to na rysunku 4. Dodatkowo przeprowadzamy kalibrację kamer.

Procedura odzyskiwania transformacji stereowizyjnej jest następująca:

1. Podkładamy w pole widzenia dwóch kamer atrapę sześcianu z widocznymi i wyróżnionymi sześcioma punktami narożnymi $P1-P6$ (jak na rys. 4). Wszystkie punkty muszą być widziane przez obie kamery, np. sześcian jest przezroczysty. Definiujemy w ten sposób globalny układ współrzędnych, ze środkiem w P_1 i stosownych osiach.
2. Definiujemy płaszczyznę pierwszą PL_1 jako płaszczyznę (XG, YG) z układem czterech punktów wzorcujących $P1-P4$. Płaszczyzna PL_2 to płaszczyzna (YG, ZG) z układem czterech punktów wzorcujących $P1, P3, P5, P6$.

3. Dokonujemy identyfikacji transformacji płaskich $TRPL_k^i$, $i, k = 1, 2$, k – numer kamery, i – numer płaszczyzny. Stosuje się czteropunktową metodę kalibracji.
4. Dokonujemy zdjęć punktu P z dwóch kamer, uzyskując w ten sposób położenia monitorowe PM_1, PM_2 punktu P . Ze względu na sposób określenia płaszczyzn, odwzorowania B_i są trywialne. Ostatecznie proponuje się odzyskanie położenia punktu P , z relacji

$$P = \text{prosta} \left(\begin{bmatrix} TRPL_1^1(PM_1) \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ TRPL_1^2(PM_1) \end{bmatrix} \right) \cap \text{prosta} \left(\begin{bmatrix} TRPL_2^1(PM_2) \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ TRPL_2^2(PM_2) \end{bmatrix} \right)$$



Rys. 4. Sytuacja stereowizyjna uproszczona

4. Podsumowanie

Literatura dotycząca problemów stereowizji i transformacji 3D jest bardzo bogata. Wymieńmy tutaj monografię Fougerasa [1] i prace Kwona [2] dotyczące metody DLT (*Direct Linear Transformation*). Praca niniejsza prezentuje pewną modyfikację metod typu DLT polegającą na istotnym zmniejszeniu liczby niezbędnych punktów kalibrujących, z 11 lub 12 do 6. Podana tutaj metoda konstrukcji transformacji stereo cechuje się wielokrotnym wykorzystaniem transformacji płaskiej jako podstawowego elementu. Transformację płaską charakteryzujemy poprzez 8 parametrów, co umożliwia ujęcie różnego typu błędów, np. złego montażu kamery, nieprostokątności osi układów lokalnych i globalnego, skal na poszczególnych osiach układów.

Dodatkowo przedstawiono dwa pomysły ułatwiające kalibrację transformacji:

- 1) definiowanie układu lokalnego na płaszczyźnie poprzez kwadrat z wyróżnionymi narożami,
- 2) definiowanie układu globalnego w przestrzeni poprzez wyróżnione naroża przezroczystego sześcianu.

W laboratorium Katedry Automatyki wykonano testowy układ kalibracji i przeprowadzono syntezę stereotransformacji, uzyskując bardzo zadowalające rezultaty (prostowanie zdjęć robionych pod dużym kątem i detekcja trajektorii ruchomego punktu).

Literatura

- [1] Faugeras O.: *Three-Dimensional Computer Vision*. Cambridge, USA, MIT Press 1993
- [2] Kwon Y.H., 1998: <http://kwon3d.com/theory/dlt/mdlt.html>
- [3] Zhang Z.: *A flexible new technique for camera calibration*. IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 22, No. 11, 2000, 1330
- [4] Tsai R.Y.: *An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision*. Proc. IEEE Conference On Computer Vision and Pattern recognition, Miami Beach, USA, 1986, 364

