

Epipolárna geometria

Bc. Ivan Blanárik

Slovenská technická univerzita v Bratislave
Fakulta informatiky a informačných technológií
ivan_blanarik@yahoo.com

Abstrakt

Epipolárna geometria ako súčasť počítačovej grafiky sa zaoberá rekonštrukciou súradníc objektov na základe dvoch alebo viacerých snímok. Cieľom článku je oboznámiť čitateľa s princípmi epipolárnej geometrie a popísať jednotlivé kroky, ktoré treba vykonať pri výpočte 3D súradníc objektu alebo scény z dvoch jej 2D obrázkov.

1. Úvod

Epipolárna geometria je prirodzenou projektívnou geometriou medzi dvomi a viacerými pohľadmi na scénu. Pohľadom na scénu sa tu myslí projekcia objektov scény do roviny. V praxi sa väčšinou jedná o snímky z fotoaparátu alebo kamery. Druh a vlastnosti projekcie sú závislé na konštrukcii a nastavení kamery. Epipolárna geometria je nezávislá na štruktúre scény. Odvodzuje sa iba od vnútorných parametrov kamery a vzájomnej polohy kamier.

Hlavným využitím epipolárnej geometrie je identifikácia korešpondujúcich si bodov zachytených na snímkach a rekonštrukcia ich 3D súradníc. Aby sme však vedeli určiť 3D súradnice zachytených bodov, je nutné vedieť správne nastaviť hodnoty interných a externých parametrov popisujúcich danú

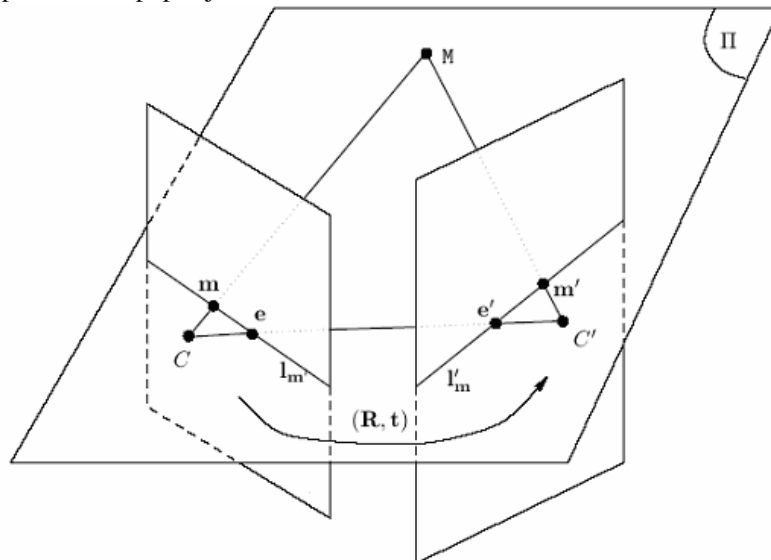
scénu. Medzi vnútorné parametre patrí ohnisková vzdialenosť kamier, pozícia riadiacich bodov, deformácia a zakrivenie... Za vonkajšie parametre sa označuje posunutie a natočenie snímok voči sebe.

Pokiaľ predpokladáme, že parametre sa počas nahrávania chôdze nemenia, nie je nutné použiť na ich určenie špeciálne kalibračné prístroje – dajú sa priamo vypočítať zo snímok scény (tzv. samokalibrácia).

V kapitolách 2., 3. a 4. sú vysvetlené základné pojmy z oblasti epipolárnej geometrie. V 5. kapitole je rozpísaná postupnosť jednotlivých krokov, ktoré sa musia vykonať na získanie 3D súradníc objektu. Kapitola 6 popisuje experiment zameraný na výber vhodného algoritmu na výpočet fundamentálnej matice a výpočet súradníc a v kapitole 7 je zhodnotenie dosiahnutých výsledkov

2. Parametre kamery

Kamera sprostredkúva mapovanie 3D scény do 2D obrazovej roviny, ktoré je reprezentované P , maticou kamery, ktorá transformuje homogénne súradnice bodu trojrozmernej scény $M=[x,y,z,1]$ na homogénne (projektívne) súradnice bodu v obraze $m=[x', y', 1]$. Platí rovnica $m = P.M$. Bod m je obrazom bodu M v rovine premietania (v obraze scény, ako je zachytený kamerou).



Obr.1: Scéna pozostávajúca z dvoch kamier, ktoré zachytávajú bod M.

Body C a C' sú stredy premietania prvej a druhej kamery. Ich vzdialenosť od roviny zobrazenia sa potom nazýva ohnisková vzdialenosť f . Prieniky spojnice ohnísk C a C' s rovinami zobrazenia sa nazývajú epipóly e a e' . Označujú miesta, kde na snímke vidíme druhú kameru. Bod m na prvej rovine zobrazenia môže ležať kdekoľvek na priamke l'_m pri pohľade z druhej kamery a naopak.

Tieto priamky určujú epipolárnu rovinu Π , ktorá je vlastne množinou všetkých potenciálnych pozícií bodu M v priestore.

3. Projekčná matica

Projekčná matica P má rozmery 3×4 a popisuje vzťah medzi reálnym bodom a bodom scény. V prípade, že súradnicová sústava má začiatok v C a jej osi sú rovnobežné s osami roviny zobrazenia, môže byť P rozložená do tvaru $P = K \cdot [I, 0]$, kde K je kalibračná matica 3×3 mapuje súradnice prostredia do súradníc kamery, I je štvorcová jednotková matica 3×3 je posunutie a 0 je nulový stĺpec.

$$K = \begin{bmatrix} a_x & c & t_x \\ 0 & a_y & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

a – škálovanie v smeroch osí x , y , súvisiace so zobrazovacím systémom
 t – súradnice optického centra – prienik zobrazovacej roviny s optickou osou
 c – parameter skreslenia (obvykle 0)

Matica projekcie P sa vo všeobecnosti líši od vyššie uvedenej reprezentácie orientáciou a posunutím voči začiatku súradnicovej sústavy. Pridáva sa preto matica transformácie súradnicovej sústavy kamery do sústavy prostredia:

$$P = K[I, 0] * \begin{bmatrix} R & -RC' \\ 0 & 1 \end{bmatrix}$$

kde R je matica 3×3 popisujúca rotáciu a C' je vektor posunutia vzhľadom na súradnicovú sústavu prostredia.

4. Fundamentálna matica

Epipolárna geometria vychádza z dvoch obrazov - projekcií scény. Jej cieľom je nájsť rovnicu, ktorá popíše vzťah medzi obrázkami. Potrebujeme teda nájsť vzťah, ktorý bodu m z prvého obrazu priradí bod m' z druhej snímky.

Pre ľubovoľnú dvojicu korešpondujúcich bodov existuje fundamentálna matica F rozmerov 3×3 , pre ktorú platí $m'^T \cdot F \cdot m = 0$. Jej hodnoty

závisia iba na vzájomnej pozícii kamier a ich kalibrácii. V prípade, že sa na jej výpočet použijú normalizované súradnice (upravené kalibračnou maticou), hovoríme o esenciálnej matici E .

5. Postup pri výpočte 3D súradníc

Pri výpočte 3D súradníc na základe 2D súradníc získaných z dvoch snímok je nutné vykonať nasledujúce kroky.

Kalibračné matice

Treba určiť kalibračné matice K , K' obidvoch kamier použitých pre získanie snímok. Kalibračné matice sú zložené z vnútorných parametrov kamery ako je ohnisková vzdialenosť a umiestnenie stredu premietania. Kalibračné matice budú slúžiť na normalizáciu 2D súradníc podľa vzťahu $x_{\text{norm}} = K^{-1} \cdot x$.

Ak ich nevieme určiť, môžu sa namiesto nich použiť jednotkové matice. V takom prípade je však nutné určiť maticu homografie (uvedené nižšie) a upravovať ňou vypočítané 3D súradnice.

Fundamentálna matica

Na základe odpovedajúcich si bodov na snímkach sa vypočíta fundamentálna matica F , ktorá bude popisovať vzťah medzi týmito bodmi. Pre súradnice x , x' toho istého objektu by malo platiť $x \cdot F \cdot x' = 0$.

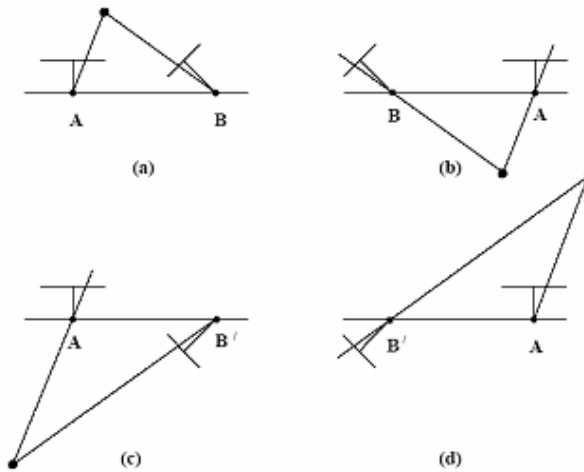
Existuje viacero algoritmov ako vypočítať F a treba zvoliť ten, ktorý bude dávať najlepšie výsledky. Ak sa pri počítaní používajú normalizované súradnice (pôvodné súradnice sú prenášobené kalibračnou maticou), získame namiesto F esenciálnu maticu E .

Projekčné matice

Predpokladá sa, že jedna projekčná matica bude umiestnená v začiatku súradnicovej sústavy, čiže nemá voči nej žiadne posunutie ani otočenie. V takom prípade stačí určiť iba projekčnú maticu druhej kamery.

Spraví sa singulárny rozklad (SVD) fundamentálnej matice F (príp. E), ktorého výsledkom sú tri matice rozmerov 3×3 $F = UDV^T$. Matice U , $D = \text{diag}[1 \ 1 \ 0]$ a V majú rozmery 3×3 . Z nich sa dajú zostrojiť 4 projekčné matice: $P' = [UWV^T | u_3]$ alebo $[UW^T V^T | u_3]$ alebo $[UWV^T | -u_3]$ alebo $[UW^T V^T | -u_3]$, pričom u_3 tretí stĺpec matice U a matica W má tvar

$$W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



Obr. 2: 4 možnosti orientácie a polohy kamier.

Takto získané projekčné matice sa líšia iba posunutím a orientáciou ich pohľadu na scénu. Projekčné matice treba otestovať na ľubovoľnom bode a vybrať tú správnu, resp. zistiť, či bod leží pred obidvoma kamerami.

Na ich vyriešenie použijeme metódu najmenších štvorcov a tým dostaneme 3D súradnice bodu X . Premenné x, y predstavujú 2D súradnice, p^i je i -ty riadok matice P .

Výpočet 3D súradníc

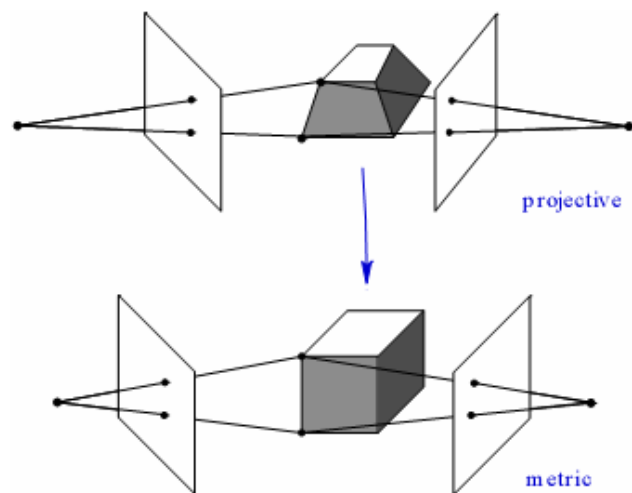
Ak už máme projekčné matice P (používa sa jednotková matica) a P' , môžeme pomocou metódy lineárnej triangulácie zrekonštruovať súradnice bodov. Získame sústavu lineárnych rovníc:

$$\begin{bmatrix} xp^{3T} - p^{1T} \\ yp^{3T} - p^{2T} \\ x'p'^{3T} - p'^{1T} \\ y'p'^{3T} - p'^{2T} \end{bmatrix} \cdot X = 0$$

Matica homografie

Pokiaľ sa pri výpočte nepoužívali normalizované súradnice, je nutné vypočítané 3D súradnice ešte upraviť. Ak sa nepoužívajú kalibračné matice, pozície bodov získaných zo snímok môže odpovedať viacerým objektom.

Na úpravu súradníc je potrebné vypočítať maticu homografie H , ktorá bude transformovať projektívne súradnice do normalizovanej súradnicovej sústavy. K jej vypočítaniu je treba mať k dispozícii 5 a viac 3D súradníc, ktorých polohu na scéne presne poznáme.



Obr. 3: Matica homografie pretransformuje vypočítané súradnice do normalizovanej súradnicovej sústavy.

6. Implementácia

Aby sme overili správnosť krokov pre výpočet 3D súradníc a zároveň zistili, s akou presnosťou ich dokážeme vypočítať, rozhodli sme sa vytvoriť testovací program realizujúci celý výpočet. Pri jeho písaní bolo výhodné použiť voľne dostupné knižnice OpenCV a GSL, ktoré nám uľahčili prácu pri rekonštrukcii súradníc zo snímok.

OpenCV je knižnica, ktorá je priamo určená na spracovanie obrazu a poskytuje množstvo funkcií, ktoré by sa dali využiť v rámci nášho projektu. Konkrétne pri výpočte 3D súradníc sa uplatnia funkcie na určenie kalibračných matíc a na výpočet fundamentálnej matice.

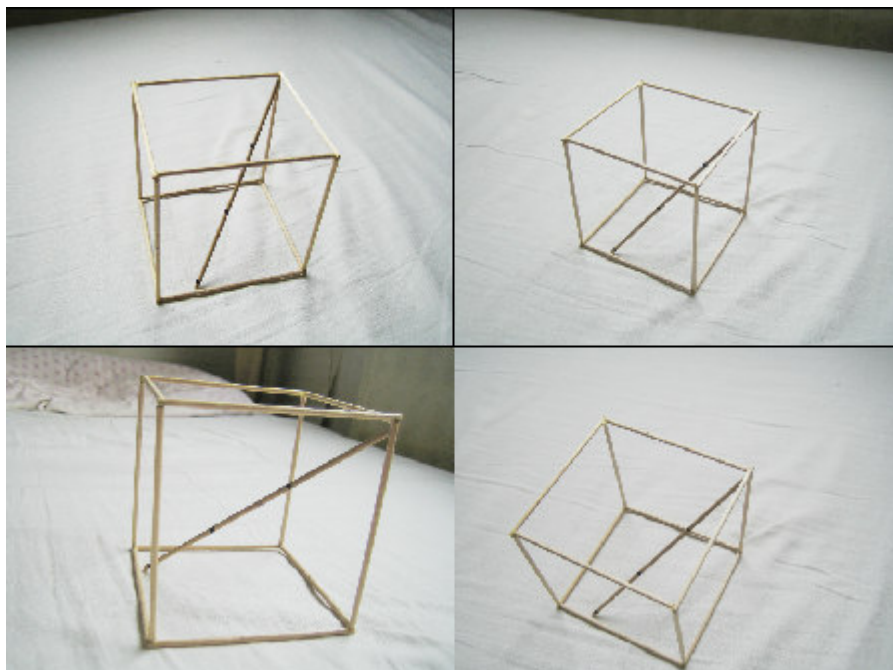
Knižnica GNU Scientific Library (GSL) je rozsiahla matematická knižnica pre C a C++

programátorov. Z nej je pre nás zaujímavá hlavne časť venujúca sa práci s maticami, funkcie na výpočet SVD a časť venujúca sa minimalizačným metódam.

Fundamentálna matica

Pri výpočte fundamentálnej matice je potrebné mať k dispozícii niekoľko (minimálne 7) dvojíc súradníc bodov. Pre tento účel sme zostrojili model kocky (na obrázku), pričom jeho rohy sú považované za značky, ktoré budú použité v algoritme.

Súradnice značiek sú určené z pozície pixlov na obrázku a sú uložené do súborov v rovnakom poradí (najprv rohy podstavy, potom horného štvorca a nakoniec značky na paličke). Každý takýto súbor predstavuje množinu súradníc získaných z jednej snímky.



Obr.5: Snímky modelu kocky. Súradnice rohov a značiek na paličke umiestnenej v kocke sú vstupnými údajmi pre výpočet fundamentálnej matice.

Pri určovaní fundamentálnej matice sa používa funkcia z OpenCV `cvFindFundamentalMat`, ktorá umožňuje použiť na jej výpočet viaceré metódy, konkrétne:

1. základný sedem-bodový algoritmus
2. základný osem-bodový algoritmus
3. algoritmus RANSAC
4. „least mean square“ algoritmus

Aby sme zistili, ktorá z metód nám poskytne najpresnejšiu fundamentálnu maticu, boli vykonané testy pre každú z nich. Zároveň bolo testované aký má vplyv počet poskytnutých značiek na výsledok.

Na posúdenie presnosti riešenia bol využitý fakt, že pre F a dvojicu zodpovedajúcich si bodov

m platí $x \cdot F \cdot x' = 0$. To však platí iba v ideálnom prípade – v praxi býva výsledkom nejaké nenulové číslo. Pre každú dvojicu snímok sa teda určila priemerná absolútna odchýlka od nuly

$$E = \frac{1}{n} \sum_{i=1}^n \text{abs}(x_n \cdot F \cdot x'_n)$$
 V nasledujúcej tabuľke sú uvedené dosiahnuté výsledky.

Metóda	1-2	1-3	1-4	2-3	2-4	3-4	Priemer
7POINT	0,6883	0,7019	0,1743	2,2489	0,1954	0,4179	0,7378
8POINT - 8	0,3100	5,3621	0,3744	0,1924	0,1951	0,2676	1,1169
8POINT - 11	0,2496	0,5471	1,6932	0,2168	0,2955	0,4437	0,5743
RANSAC - 8	1,5556	0,7827	14,9959	0,5202	0,1774	0,4546	3,0811
RANSAC - 11	2,3461	0,3595	10,6615	0,5612	0,3990	0,1302	2,4096
LMEDS - 8	9,5506	1,0217	10,1508	1,6107	0,6301	0,4546	3,9031
LMEDS - 11	6,1261	20,2759	0,8315	1,7515	2,5225	1,9511	5,5764

Testy boli uskutočnené iba na súboroch čísla použitím štyroch snímok. Sedembodový algoritmus dokáže pracovať iba so siedmimi bodmi, zatiaľ čo ostatné metódy (potrebujú 8 a viac) boli vyskúšané na množine 8 a 11 bodov.

Z dosiahnutých výsledkov vyplýva, že najpresnejšiu fundamentálnu maticu dostaneme použitím základného osembodového algoritmu. Zároveň je vidieť, že pridaním ďalších bodov sa zlepšia dosahované výsledky - v prípade 8-bodového algoritmu už pridanie 3 bodov umožnilo dosiahnuť 50% zlepšenie.

Základný 8-bodový algoritmus vychádza z rovnice $x \cdot F \cdot x' = 0$. Po jej rozpísaní získame rovnicu $x'x'f_{11} + x'y'f_{12} + x'f_{13} + y'x'f_{21} + y'y'f_{22} + y'f_{23} + x'f_{31} + y'f_{32} + f_{33} = 0$

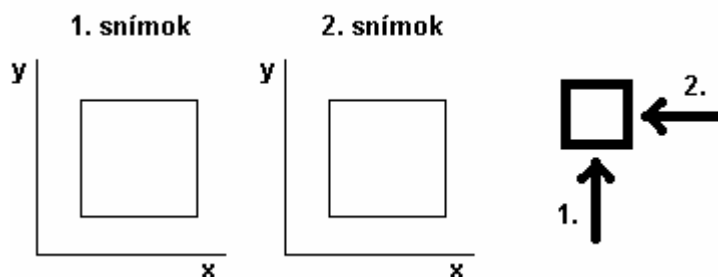
. Pokiaľ máme k dispozícii 8 a viac bodov, dostávame sústavu rovníc, z ktorých dokážeme minimalizáciou vypočítať prvky fundamentálnej matice.

Projekčné matice a výpočet súradníc

Po určení fundamentálnej matice F základným osembodovým algoritmom, sa táto matica rozloží a zostroja sa 4 alternatívy projekčnej matice. Súradnice vypočítané týmito maticami sa medzi sebou líšia znamienkom a vzájomnou polohou, vzdialenosti medzi nimi však zostávajú rovnaké.

V tejto fáze testovania nás skôr zaujímalo, ako presne sú 3D súradnice vypočítané, resp. či výsledný útvar nie príliš zdeformovaný, a nie či je výsledný útvar nejako natočený. Výber správnej projekčnej matice sme odložili na neskôr.

Pri prvom meraní sme použili nami pripravené dáta. Neboli odčítané zo snímok, ale boli určené spôsobom, ako je zobrazené na obr. 6. Chceli sme tak otestovať algoritmus na jednoduchom prípade ešte pred tým, ako prejdeme na reálne snímky.



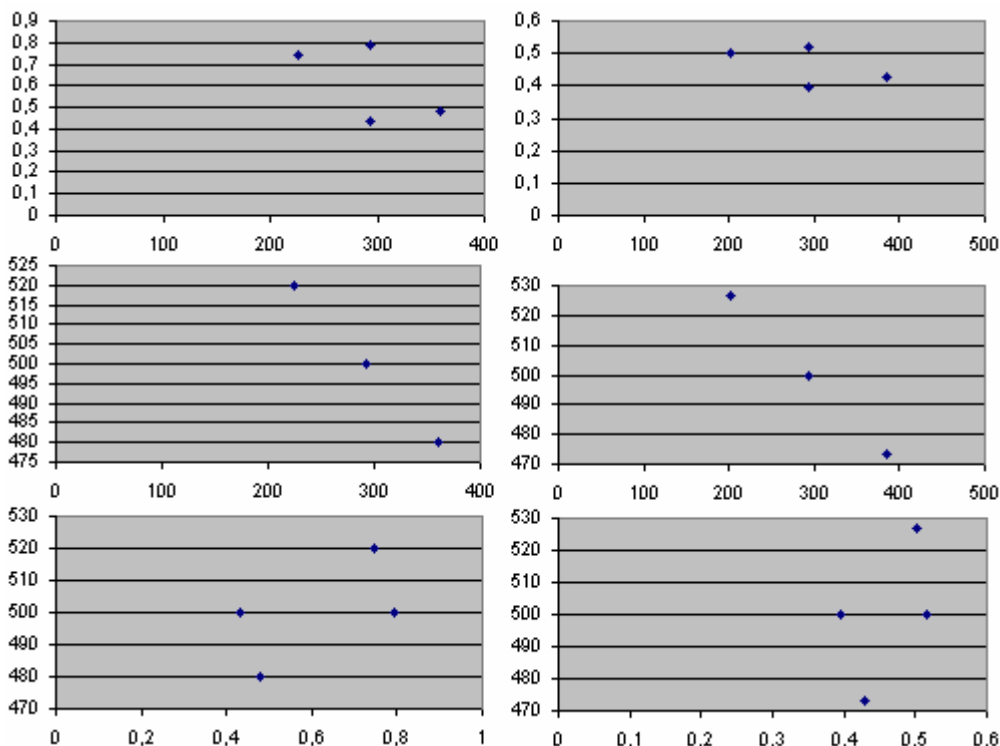
Obr. 6: Ideálna situácia. Predné body kocky zakrývajú zadné. Kamery sú umiestnené v rovnakej vzdialenosti a kolmo na seba.

V nasledujúcej tabuľke sú uvedené vstupné hodnoty a hodnoty vypočítané našim algoritmom. Ako je z výsledkov vidieť (Obr. 7), vypočítané súradnice sú oproti pôvodným posunuté a pootočené, takže výsledný útvar je skosený. Je to spôsobené tým, že sme nepoužívali kalibračné kamery a chýba aj matica homografie, ktorá by

súradnice pretransformovala do pôvodných súradníc. Súradnice z-ovej osi sú navyše o tri rády nižšie, nie je však problém ich upraviť, aby rádovalo sedeli s ostatnými.

Bod	Vstupný súbor		Vypočítané súradnice		
	x	y	x	y	z
1	500	500	292,584	499,761	0,792431
2	1000	500	359,757	480,008	0,479817
3	1000	500	292,634	499,975	0,43421
4	500	500	225,465	519,728	0,745532
5	500	1000	292,64	499,92	0,516989
6	1000	1000	384,377	473,067	0,42872

7	1000	1000	292,652	499,993	0,396344
8	500	1000	200,914	526,836	0,501857



Obr. 7: Súradnice vypočítaného útvaru. Vľavo je zobrazená spodná stena kocky, vpravo vrchná, pričom prvá dvojica predstavuje pohľad zhora, nasleduje pohľad spredu a z boku.

Výsledný útvar je síce deformovaný, ale stále pripomína skosenú kocku. Pri použití reálnych snímok však dochádza ešte k väčšiemu skresleniu a navyše sa niektoré body nachádzajú v značnej vzdialenosti od ostatných.

7. Zhodnotenie

Na základe experimentov sa ukázal ako najlepší na výpočet fundamentálnej matice základný osem-bodový algoritmus. Z testovaných algoritmov dosahoval najmenšiu chybu a tá sa s pribúdajúcim počtom bodov ešte ďalej znižovala.

Do akej miery je deformácia výsledných 3D súradníc spôsobená používaním nenormalizovaných súradníc v tejto chvíli nedokážeme povedať. Z časových dôvodov sa nám nepodarilo implementovať výpočet kalibračných matíc (poprípade matice homografie), čo by bol posledný zostávajúci krok pri výpočte súradníc. Zároveň by nám to umožnilo porovnať skutočné a vypočítané súradnice a určiť celkovú presnosť.

Výsledky získané pomocou nenormalizovaných súradníc sú však porovnateľné s ukázkami uvedenými v študijných materiáloch. Tie sa od našich nejakým výrazne nelíšia a môžeme predpokladať, že doteraz vykonané kroky nevášajú do výpočtu súradníc žiadnu veľkú chybu.

Aj napriek tomu, že momentálne nevieme určiť presnosť, s akou budú 3D súradnice

zodpovedať skutočnosti, bolo by výhodné v budúcnosti používať epipolárnu geometriu. Kvalita dosahovaných výsledkov by mala závisieť čisto iba na presnosti vstupných údajov, takže vypočítané 3D súradnice by mali byť zaťažené rovnakou chybou, aká bude aj pri použití iných metód výpočtu.

Použitá literatúra

1. Borghese, A., Cerveri, P.: Calibrating a video camera pair with a rigid bar. Laboratory of Human Motion Study and Virtual Reality 1999.
2. Zimányi, M.: Parametre kamery a kalibrácia. Prednáška o 3D syntéze obrazu 2004.
3. Zimányi, M.: Princípy trojrozmernej analýzy obrazu. Prednáška o 3D syntéze obrazu, 2004.
4. Pospíšilová, R.: Kritické konfigurácie pro výpočet geometrie kamery. 25. konferencie o geometrii a počítačové grafice.
5. Dařílková, K.: Modelovanie 3D objektov z fotografií. Katedra algebry, geometrie a didaktiky matematiky UK v Bratislave 2004.
6. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge 2000