

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

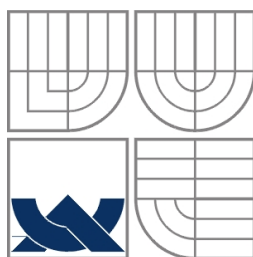
IDENTIFIKÁCIA POHYBU V PRIESTORE

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

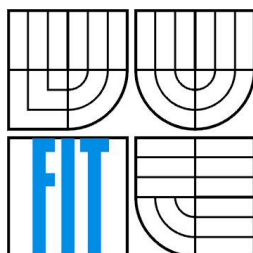
AUTOR PRÁCE
AUTHOR

MAREK ŠOLONY

BRNO 2007



VYSOKÉ UČENÍ TECHNICKÉ V
BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INFORMAČNÍCH SYSTÉMŮ
FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INFORMATION SYSTEMS

IDENTIFIKÁCIA POHYBU V PRIESTORE

MOVEMENT IDENTIFICATION IN THE SPACE

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE
AUTHOR

Marek Šolony

VEDOUCÍ PRÁCE
SUPERVISOR

Ing. Igor Potůček, Ph.D.

BRNO 2007

Abstrakt

Táto práca sa zaoberá identifikáciou pohybu v priestore a jeho rekonštrukciou do 3D priestoru. Navrhuje optický systém na zaznamenávanie pohybu špeciálnych značiek – markerov umiestnených na ľudskom tele pomocou páru kamier. Sú tu rozobrané jednotlivé kroky, od kalibrácie kamier, až po výpočet 3D pozície markeru. Na hľadanie korešpondencií objektov medzi snímkami je využitá epipolárna geometria. V závere sú uvedené a zhodnotené dosiahnuté výsledky.

Kľúčové slová

Záznam pohybu, detekcia pohybu, projektívna geometria, epipolárna geometria, fundamentálna matica, 3D triangulácia, kamerový model, kalibrácia kamery, rekonštrukcia pohybu.

Abstract

The aim of this paper is to suggest optical system capable of movement identification in space and its reconstruction. The motion capture system uses markers attached to parts of human body, and a camera pair to capture the movement. This paper describes step-by-step parts of this system. Epipolar geometry is used to deal with problem of object correspondence between two views.

Keywords

Motion capture, movement identification, projective geometry, epipolar geometry, fundamental matrix, 3D triangulation, marker, camera model, camera calibration.

Citácia

Marek Šolony: Identifikácia pohybu v priestore, bakalárska práca, Brno, FIT VUT v Brně, 2007

Identifikácia pohybu v priestore

Prehlásenie

Prehlasujem, že som túto bakalársku prácu vypracoval samostatne pod vedením Ing. Igora Potúčka, Ph.D. Uviedol som všetky literárne pramene a publikácie, z ktorých som čerpal.

.....
Jméno Příjmení
Datum

Pod'akovanie

Ďakujem vedúcemu bakalárskej práce Ing. Igorovi Potúčkovi, Ph.D. za pomoc pri získaní zdrojových dát pre prácu, a za cenné rady pri spracovaní bakalárskej práce.

© Marek Šolony, 2007.

Tato práce vznikla jako školní dílo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práce je chráněna autorským zákonem a její užití bez udělení oprávnění autorem je nezákonné, s výjimkou zákonem definovaných případů..

Obsah

Obsah	1
Úvod.....	3
1 Úvod do projektívnej geometrie a kamerových systémov	6
1.1 Projektívna geometria.....	6
1.1.1 Homogénne koordináty	6
1.1.2 Projekčná plocha	7
1.1.3 Prienik dvoch priamok	7
1.1.4 Dvoj pomer.....	7
1.1.5 Kužeľosečky.....	8
1.1.6 Projekčný priestor	9
1.2 Jednoduchý kamerový model a epipolárna geometria.....	9
1.2.1 „Pinhole“ kamerový model.....	9
1.2.2 Vnútorne a vonkajšie parametre kamery	10
1.2.3 Radiálna distordia.....	11
1.2.4 Epipolárna geometria	12
1.2.5 Fundamentálna matica.....	12
1.3 Rekonštrukcia objektu do 3D priestoru	13
1.3.1 Mid-point triangulačný algoritmus.....	14
2 Špecifikácia vlastného MC systému.....	16
2.1 Kalibrácia kamier a určenie počiatku svetovej súradnicovej sústavy.....	17
2.1.1 Určenie počiatku svetovej súradnicovej sústavy.....	18
2.2 Extrakcia pozícií markerov z obrazu	19
2.2.1 Problémy pri detekcii objektov	20
2.3 Korešpondencia objektov medzi snímkami z kamier	21
2.3.1 Určenie korešpondencie objektov medzi snímkami.....	21
2.3.2 Problémy nejednoznačnosti priradenia objektu	22
2.3.3 Problémy pri rozdielnych počtoch objektov v snímkach	23
2.4 Výpočet 3D pozície markerov v priestore	23
2.5 Korešpondencia objektov v čase	24
2.5.1 Kalmanov filter	25
2.5.2 Využitie Kalmanovho filtru za zdokonalenie algoritmu priradovania korešpondencií objektov medzi snímkami z rôznych kamier	26
3 Dosaiahnuté výsledky	27
3.1 Test detekcie objektov v snímkach.....	27

3.2	Test správneho priradenia korešpondencie medzi snímkami z kamier	29
3.3	Test presnosti výpočtu triangulačného algoritmu.....	30
3.4	Test korešpondencie markerov v čase	31
3.5	Test rekonštrukcie pohybu.....	33
Záver	34
Literatúra	36
Zoznam príloh	37

Úvod

Vytvoriť dôveryhodnú animáciu pohybu objektu tradičným spôsobom animácie je veľmi náročné, pretože človek veľmi ľahko spozoruje pohyb, ktorý je neprirodzený alebo nepravdepodobný. Alternatívny spôsob k získaniu realisticky vyzerajúcej animácie je možný pomocou systémov, ktoré využívajú rôzne spôsoby *identifikácie pohybu* reálneho objektu v priestore. Tieto systémy sa nazývajú Motion Capture (MC) systémy. Dáta získané pomocou nich môžu byť následne využité k rekonštrukcii pohybu pri animácii.

Pojem Motion Capture je definovaný ako „vytvorenie 3D reprezentácie živého vystúpenia“. Je to kontrast k animácii vytvorenej „ručne“. Motion Capture zahŕňa identifikáciu pohybu, meranie pozície a orientácie objektu v priestore, a zaznamenávanie tejto informácie. Objekty záujmu môžu byť humanoidného alebo nehumanoidného charakteru, výrazy tváre, pozície kamier, svetiel alebo iných prvkov v priestore. Vo väčšine prípadov sú snímané živé objekty, ľudské postavy ktoré slúžia ako zdrojové dáta, a tie sa následne transformujú do digitálnej podoby.

Prvé pokusy o analýzu a záznam ľudského pohybu prebehli už v okolo roku 1800, keď sa niektorí vedci snažili analyzovať ľudský pohyb z fotografií na lekárske účely. Moderná technológia Motion Capture začína ako vojenský projekt a bola určená na sledovanie cieľov pomocou magnetických senzorov. Po roku 1986 sa táto technológia dostala do súkromného sektoru. Vtedajšie MC systémy patrili ku kontroverzným spôsobom získavania animácie, pretože úprava a vyhladzovanie získaných dát zaberala rovnaké množstvo času, ako by trvalo vytvorenie animácie tradičnými spôsobmi. Odvtedy sa však táto technológia vyvíjala, vznikali nové druhy zachytenia pohybu a našla uplatnenie v biomechanickom výskume, ale najmä v počítačovej grafike, keďže sa v súčasnej dobe s jej vývojom kladú čoraz väčšie nároky na kvalitu a realističnosť animácie. Ak sa chceme priblížiť realistickej animácii, musíme analyzovať pohyb animovaného objektu v skutočnom svete, následne syntetizovať získané dáta a použiť ich k dosiahnutiu kvalitného výsledku.

Pri snímaní pohybu pomocou MC systémov je pohyb jedného alebo viacerých hercov zaznamenávaný mnoho krát za sekundu. Keďže je nahrávaný len pohyb herca, a nie jeho reálny výzor, možno získané pohybové dáta namapovať na 3D model, a prinútiť ho vykonávať rovnaký pohyb. Výhodou MC systémov oproti tradičným animačným metódam spočíva v tom, že môžu zachytiť komplexné pohyby, realistické interakcie a fyzikálne vlastnosti ako hmotnosť, hybnosť, výmena energií, a najmä sú relatívne menej časovo náročné.

Systémy MC so sebou prinášajú množstvo výhod, ktoré našli uplatnenie napríklad vo filmovom odvetví. Od tradičného spôsobu animácie sa už upustilo, a nahradili ho tieto technológie hlavne kvôli tomu, že väčšina systémov pracuje v reálnom čase, takže možno

kontrolovať priebežne dosiahnuté výsledky a prípadná zmena v scenári znamená len natočenie novej akcie. Tak isto klesá časová náročnosť a dokonca jeden herec môže zahrať viacero rolí.

Nielen v počítačovej animácii našli MC systémy svoje uplatnenie. Aktívne sa využívajú pri analýze ľudského pohybu, v medicíne, pri kontrole výrobného procesu, alebo dokonca riadení robotov.

Existuje niekoľko rôznych princípov, na ktorých MC systémy pracujú. Najpoužívanejšie MC systémy sa delia na magnetické, optické a mechanické.

Magnetické MC systémy využívajú markery – prijímače priamo pripojené k počítaču, pohybujúce sa v nízkofrekvenčnom elektromagnetickom poli vytvoreným centrálnou umiestneným vysielačom. Dokážu pracovať v reálnom čase a nenastávajú problémy so stratou alebo zamenou markerov. Markery okrem polohy v priestore dokážu rozoznávať aj natočenie, tým pádom ich stačí menej pre výsledky zrovnateľné s optickým MC systémom. Na druhú stranu kvôli prípojným káblom tento systém obmedzuje pohyb a celý systém je náchylný k elektromagnetickému rušeniu spôsobenému metalickými objektmi, monitormi, svetlami alebo uložením káblov v miestnosti.

Mechanické MC systémy využívajú špeciálne obleky pripomínajúce kostru so zabudovanými mechanickými senzormi snímajúce uhly pri kĺboch herca. Ako sa herec pohybuje, senzory priamo snímajú zmeny uhlov a tým pádom ja hercov pohyb. Mechanické MC systémy sú pomerne lacné a tak isto ako magnetické pracujú v reálnom čase. Problematickejšie je snímanie interakcie dvoch alebo viacerých hercov.

Optické MC systémy sa stali v poslednej dobe veľmi populárnymi. Používajú reflexné alebo inak výrazné markery, ktoré sú sledované sústavou kalibrovaných kamier. Pozícia markeru v 3D priestore je vypočítaná na základe jeho pozícií v jednotlivých záberov z kamier. Natočenie markeru je určené až na základe polohy voči ostatným markerom. Napríklad natočenie v lakti sa určí pomocou markeru z pleca, lakt'a a zápästia. Systémy používajú väčšie množstvo kamier, aby pokryli snímanú miestnosť a aby markery prekrytím časťou tela nemizli.

Pasívne optické MC musia riešiť korešpondenciu markerov medzi zábermi z kamier. Tieto výpočty sú časovo náročné, tým pádom pasívne optické MC systémy nepracujú v reálnom čase, ale analyzujú nahrané videá.

Aktívne optické MC systémy používajú ako markery let diódy. Let diódy svetlo emitujú miesto toho aby ho odrážali, preto možno snímať hercov aj na väčšie vzdialenosti. Korešpondencia medzi zábermi z kamier je riešená tak, že v každom okamihu je rozsvietená práve jedna dióda. Na snímanie týchto diód je nutné použiť vysokorýchlostné kamery. Tieto systémy dokážu pracovať v reálnom čase.

MC systémy	Magnetické	Mechanické	Optické pasívne	Optické aktívne
Senzory:	elektromagnetické	mechanický skelet	reflexné markery	let diódy
Spracovanie:	v reálnom čase	v reálnom čase	spracovanie záznamu	v reálnom čase alebo spracovanie záznamu
Cena (USD):	5 000 – 15 000	25 000 – 75 000	50 000	400 000
Výhody:	- dáta o polohe a natočení markeru - stačí menší počet markerov - cena	- dáta o polohe a natočení markeru - herec nie je obmedzovaný káblami	- herec nie je obmedzovaný káblami - možnosť snímania interakcie viacerých hercov - presné polohy markerov	
Nevýhody:	- herec a veľkosť snímanej miestnosti sú obmedzené káblami - elektromagnetické rušenie	- problém snímania interakcie viacerých hercov	- v prípade pasívneho systému spracovanie nie je v reálnom čase - zložitá kalibrácia kamier - cena	

Tabuľka 1.: Rozdelenie MC systémov, ich výhody a nevýhody

Hlavným cieľom tejto práce je napodobnenie pasívneho optického Motion Capture systému pomocou relatívne dostupného vybavenia. Najdôležitejšie na dosiahnutie tohto cieľa bude rozdelenie celého procesu od získania zdrojových dát, až po rekonštrukciu pohybu na jednotlivé časti – zostavenie scény a získanie zdrojových dát, identifikácia pohybu markerov v obraze, nájdenie korešpondencie markerov medzi snímkami z kamier a v čase. Pri častiach zaoberajúcich sa korešpondenciami je najkomplikovanejší problém identity markerov. Keďže reprezentáciu pohybu tvoria práve markery ktoré majú jednotný vzhľad, bude potrebné riešiť problém hľadania obrazov markeru v snímkach.

V tejto práci je rozoberaná teória projekčnej geometrie, ktorou sa dá opísať proces premietania 3D scény na 2D rovinu, a tak isto je opísaný aj jednoduchý kamerový model a epipolárna geometria existujúca medzi dvomi pohľadmi. V ďalších kapitolách nasleduje presná špecifikácia vlastného MC systému a jednotlivých jeho súčastí.

1 Úvod do projektívnej geometrie a kamerových systémov

V nasledujúcich kapitolách sa zoznámime s teóriou potrebnou k pochopeniu procesu získavania dát počínajúc kalibráciu kamier, končiac 3D rekonštrukciou objektu a sledovaním markerov.

1.1 Projektívna geometria

Projektívna geometria je matematická sústava, ktorá je využívaná hlavne v oblasti počítačového videnia. Je dôležitá pri popise závislostí pri kalibrácii kamery, stereo rekonštrukcii, rozpoznávaní objektu, syntéze obrázku, alebo analýze tieňov.

Pozorovaný 3D priestor môžeme opísať geometriou. Euklidovská geometria opisuje uhly medzi pretínajúcimi sa priamkami, tvary objektov a priamky, ktoré sú rovnobežné sa nikdy nepretnú. Pri zobrazovacích procesoch s kamerou začína byť tento model geometrie nedostatočný. Uhly a tvary nemožno určiť, a rovnobežné priamky sa pretínajú. Pri Euklidovskej geometrii je možné riešiť tento problém definovaním, že dve rovnobežné priamky sa pretínajú v bode „v nekonečne“. Euklidovská geometria je podmnožina toho, čo sa nazýva *projektívna geometria*.

Projektívna geometria existuje v ľubovoľnom počte dimenzií. S použitím jednej získame projekčnú priamku ρ^1 , dvoch projekčnú plochu ρ^2 a troch projekčný priestor ρ^3 . Zobrazovací proces je projekcia z ρ^3 do ρ^2 , z trojrozmerného priestoru do dvojrozmerného.

1.1.1 Homogénne koordináty

Predstavme si bod (x, y) v Euklidovskej ploche. Aby sme reprezentovali ten istý bod v projektívnej geometrii, pridáme tretí koordinát „1“ na koniec: $(x, y, 1)$. Všeobecne je škálovanie nedôležité, takže $(x, y, 1)$ zodpovedá rovnakému bodu ako $(\alpha x, \alpha y, \alpha)$, kde $\alpha \neq 0$ a tieto koordináty sa nazývajú *homogénne*.

Ak by bolo $\alpha = 0$, x/α a y/α je nekonečno. Takto sú definované body v nekonečne. Sú to body reprezentované homogénnymi koordinátami, ktorých posledný koordinát je nulový.

Tým, že v Euklidovskom priestore zvolíme počiatok súradnicovej sústavy, sa môže zdať že zvolený bod nadobúda dôležitejšie postavenie ako ostatné body. Keďže priestor je homogénny, nezáleží na tom, ktorý bod bude zvolený ako počiatok súradnicovej sústavy. Zmenu počiatku súradnicovej sústavy môžeme jednoducho dosiahnuť rotáciou a posunom

súradnicových osí. Tieto operácie sú označované ako *Euklidovská transformácia*. Pri týchto transformáciách sú body v nekonečne zachované.

1.1.2 Projekčná plocha

Priestor ρ^2 je označovaný ako *projekčná plocha*. Body v tejto ploche sú definované vektormi:

$p = [\alpha x_1, \alpha x_2, \alpha]^T$ s Euklidovskou pozíciou $(x_1 / \alpha, x_2 / \alpha, 1)$.

Priamka je definovaná vektorom $l = [l_1, l_2, l_3]^T$ a rovnicou:

$$l_1 x_1 + l_2 x_2 + l_3 = 0 \quad (1.1)$$

Ak bod p leží na priamke l , platí:

$$p^T l = l^T p = 0 \quad (1.2)$$

Táto rovnica vyjadruje *princíp duality* – v projekčnej ploche ρ^2 nie je rozdiel medzi reprezentáciou bodu a priamky. Rovnica (1.2) sa nazýva *rovnica priamky*, čo znamená, že bod p je definovaný sériou priamok, ktoré cez bod prechádzajú, alebo sa rovnica nazýva *rovnica bodu*, čiže priamka je reprezentovaná sériou bodov.

1.1.3 Prienik dvoch priamok

Predpokladajme, že chceme nájsť prienik dvoch priamok $u = [u_1, u_2, u_3]^T$ a $v = [v_1, v_2, v_3]^T$. Priamky sa pretnú v bode $P = [u_2 v_3 - v_2 u_3, v_1 u_3 - u_1 v_3, u_1 v_2 - v_1 u_2]$. Tento vzorec sa dá zapísať ako vektorový súčin:

$$P = u \times v \quad (1.3)$$

Ak sú priamky rovnobežné, pretínajú sa v bode v nekonečne $P = [u_2 v_3 - v_2 u_3, v_1 u_3 - u_1 v_3, 0]$.

1.1.4 Dvojpomer

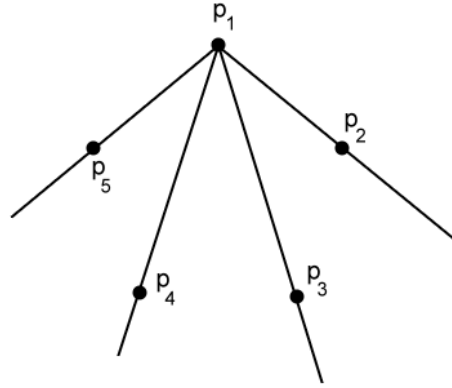
Projektívna geometria nezachováva vzdialenosti a ani pomery vzdialeností, ale *dvojpomer*, čo je pomer pomerov vzdialeností ostáva zachovaný. Vezmime si body p_1, p_2, p_3 a p_4 ležiacich na tej istej priamke a vzdialenosť medzi bodmi p_i a p_j označme $\lambda_{i,j}$. Definícia dvojpomeru je nasledovná:

$$Cr(p_1, p_2, p_3, p_4) = \frac{\lambda_{1,3} \lambda_{2,4}}{\lambda_{1,4} \lambda_{2,3}} \quad (1.4)$$

Euklidovská vzdialenosť medzi dvoma bodmi $p_1 = [x_1, y_1, \alpha]$ a $p_2 = [x_2, y_2, \alpha]$ je v 2D Euklidovskom priestore vypočítaná nasledovne:

$$\lambda_{p_1, p_2} = \sqrt{\left(\frac{x_1}{\alpha} - \frac{x_2}{\alpha}\right)^2 + \left(\frac{y_1}{\alpha} - \frac{y_2}{\alpha}\right)^2} \quad (1.5)$$

Existujú rôzne možnosti merania dvojpomeru, ktoré tak isto ostávajú nemenné po projektívnej transformácií. Princípom duality je možné počítať dvojpomer pre štyri priamky zbiehajúce sa v jednom bode, zámennou Euklidovskou vzdialenosťou uhlami, ktoré priamky zvierajú. Týmto spôsobom sa dá vypočítať dvojpomer aj pre päť bodov, ktoré neležia na jednej priamke, ako ukazuje obrázok (Obrázok 1.1: Dvojpomer pre päť nekolineárnych bodov).



Obrázok 1.1: Dvojpomer pre päť nekolineárnych bodov

1.1.5 Kužeľosečky

V Euklidovskej geometrii sú kužeľosečky – elipsy, hyperboly a paraboly ľahko definovateľné. Ale v dôsledku projektívnej geometrie kužeľosečky strácajú svoju charakteristickosť pretože sú projektívne ekvivalentné, čo znamená, že ľubovoľná forma môže byť projektovaná do inej.

Kužeľosečka v perspektívnej geometrii je definovaná ako súbor bodov, ktoré majú konštantný dvojpomer k štyrom určeným bodom, z ktorých žiadne tri neležia na priamke.

$$p^T C p = 0 \quad (1.6)$$

kde p je vektor veľkosti 3×1 a C je symetrická matica o rozmeroch 3×3 .

Tak isto aj pre kužeľosečky existuje princíp duality, duálna kužeľosečka C' , alebo jej obálka je definovaná ako súbor priamok vyhovujúcich rovnici:

$$l^T C' l = 0 \quad (1.7)$$

kde C' je symetrická matica o veľkosti 3×3 .

Závislosť medzi C a C' je nasledovná: ak sa mení bod p pozdĺž kužeľosečky a rovnica $p^T C p = 0$ je pravdivá, tak dotyčnica l kužeľosečky v bode p vyhovuje rovnici $l^T C^{-T} l = 0$. Porovnaním s rovnicou (1.7) zistíme, že dotyčnice ku kužeľosečke C patria duálnej kužeľosečke.

1.1.6 Projekčný priestor

Bod v ρ^3 priestore je určený vektorom o rozmeroch 4×1 $p = [x_1, x_2, x_3, x_4]$. Duálnou entitou k bodu v ρ^3 priestore je rovina Π , ktorá je taktiež reprezentovaná vektorom 4×1 $\Pi = [\Pi_1, \Pi_2, \Pi_3, \Pi_4]$ a rovnicou

$$x_1\Pi_1 + x_2\Pi_2 + x_3\Pi_3 + x_4\Pi_4 = 0 \quad (1.8)$$

a ak bod leží v rovine Π , musí byť pravdivá nasledujúca rovnica:

$$p^T \Pi = 0 \quad (1.9)$$

Všetky koncepcie doposiaľ diskutované v ρ^2 priestore majú svoju analógiu v ρ^3 priestore.

1.2 Jednoduchý kamerový model a epipolárna geometria

V tejto kapitole je opísaný perspektívny kamerový model, takzvaný „pinhole“ model, pri ktorom predpokladáme, že negatívne efekty na kvalitu obrazu ako je radiálna distordia šošoviek nie je relevantná a preto ju ignorujeme. Riešenie vplyvu radiálnej distorcie je rozoberané následne.

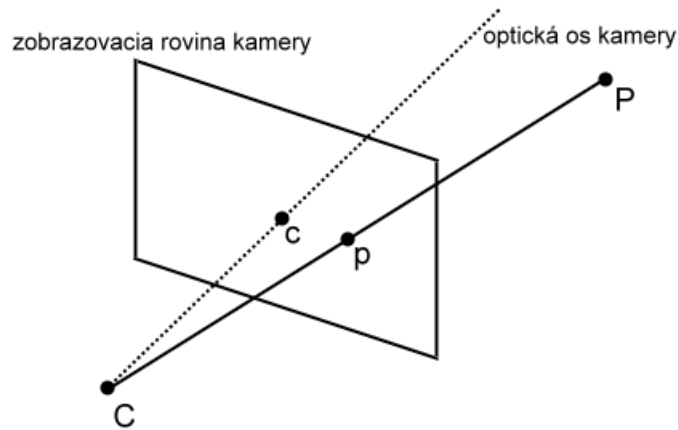
Ďalšou časťou kapitoly je epipolárna geometria, ktorá existuje medzi dvomi rôznymi pohľadmi. Zameriame sa na špeciálnu maticu, ktorá je algebrickou reprezentáciou epipolárnej geometrie a zohráva dôležitú úlohu pri hľadaní korešpondujúcich objektov medzi snímkami z dvoch pohľadov.

1.2.1 „Pinhole“ kamerový model

Existuje zobrazenie scény do zobrazovacej plochy kamery $\rho^3 \rightarrow \rho^2$. Obraz scény vytvorený projekciou 3D bodov do roviny zobrazenia použitím perspektívneho zobrazenia. Ak $P = [X, Y, Z]^T$ je pozícia bodu v svetovom súradnicovom systéme a $p = [u, v]^T$ je pozícia obrazu bodu P v súradnicovom systéme snímku platí:

$$s * p' = A * [R | t] * P \quad (1.10)$$

kde s je váhový koeficient, A sa nazýva kamerová kalibračná matica, alebo matica vnútorných parametrov kamery, matice R a t vyjadrujú vonkajšie parametre kamery, R je natočenie a t je translačný vektor. $P' = [X, Y, Z, 1]$ a $p' = [u, v, 1]$ sú homogénne koordináty vektorov P a p . Matica $P_p = A * [R | t]$ sa nazýva projekčnou maticou zobrazenia.



Obrázok 1.2: Perspektívna projekcia.

Optická os kamery prechádza centrom projekcie C a pretína zobrazovaciu rovinu kamery v bode c , ktorý sa nazýva hlavný bod.

Koordináty bodu p , ktorý je obraz bodu P , závisia aj od vzdialenosti centra projekcie C a zobrazovacej roviny kamery, ohniskovej vzdialenosti označovanej f . Ak umiestnime začiatok svetového súradnicového systému do centra projekcie a os z vyjadruje hĺbku, platí:

$$u = f \frac{x}{z}$$

$$v = f \frac{y}{z}$$

1.2.2 Vnútorne a vonkajšie parametre kamery

Každá kamera, ktorá je súčasťou kamerového systému má svoje vnútorné a vonkajšie parametre. Vnútorne parametre určujú ohniskovú vzdialenosť, hlavný bod (zvyčajne stred snímku), skosenie a šošovkovú distorciu. Vnútorne parametre kamery nie sú závislé na polohe kamery v priestore, preto sa dajú vždy znovu použiť, ak sa nezmenila ohnisková vzdialenosť. Všetky vnútorné parametre sú zhrnuté v kalibračnej kamerovej matici.

Vzorec (1.8) možno rozpísať:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} fx & 0 & cx \\ 0 & fy & cy \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_1 \\ r_{2,1} & r_{2,2} & r_{2,3} & t_2 \\ r_{3,1} & r_{3,2} & r_{3,3} & t_3 \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1.11)$$

Z kalibračnej matice kamery A parametre $[cx, cy]$ vyjadrujú hlavný bod, fx a fy sú ohniskové vzdialenosti delené veľkosťou pixlu, ak pixel nemá pomery šírky a výšky 1:1.

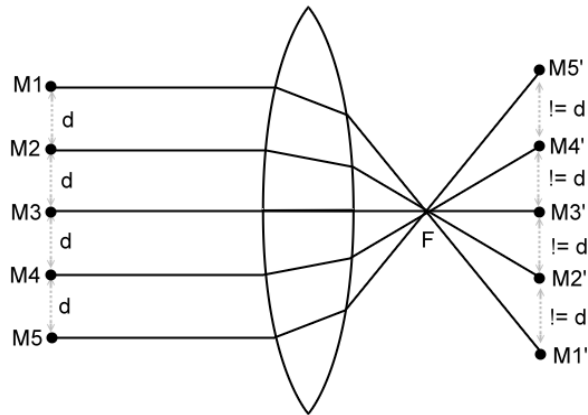
Matica, ktorá vznikne spojením matíc R a t sa nazýva matica vonkajších parametrov. Opisuje pohyb kamery voči statickej scéne, alebo naopak pohyb objektu pred statickou

kamerou. Matica $[R \mid t]$ transformuje súradnice bodu $[X, Y, Z]$ do súradnicového systému určeného vzhľadom na kameru:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R * \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + t \quad (1.12)$$

1.2.3 Radiálna distorcia

Pri zobrazovaní pomocou šošoviek dochádza k miernemu skresleniu obrazu, distorcii, ktorú treba zobrať v úvahu pri extrakcii polôh markerov zo záberov z kamier.



Obrázok 1.3: Radiálna distorcia

Ak premietame Bod M3, lúč dopadá na šošovku kolmo na jej povrch, prechádza skrz a nenastáva lom lúču, tým pádom sa zobrazí správne. Ale ak premietame bod M1, lúč dopadá pod istým uhlom, láme sa na šošovke, a keďže reálne šošovky nelámu lúč presne do ohniska, pri zobrazení bodu na premietaciu plochu vzniká vplyvom radiálnej distorcie isté posunutie.

Ak sú body $p_d = [x_d, y_d]$ ovplyvnené distorciou, môžeme získať „opravené“ body $p_u = [x_u, y_u]$ nasledovným predpisom:

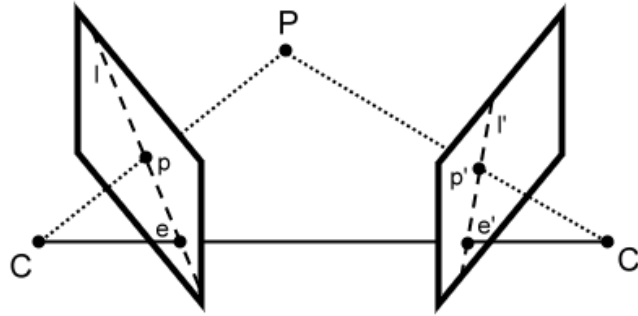
$$\begin{aligned} x_u &= c_x + (x_d - c_x)(1 + k_1 r_d^2 + k_2 r_d^4 + \dots) \\ y_u &= c_y + (y_d - c_y)(1 + k_1 r_d^2 + k_2 r_d^4 + \dots) \end{aligned} \quad (1.13)$$

kde (c_x, c_y) je centrum distorcie (hlavný bod), $r_d = \sqrt{x_d^2 + y_d^2}$ a k sú koeficienty radiálnej distorcie.

V praxi môže byť model radiálnej distorcie zapísaný pomocou nekonečného rozvoja, ale experimentálne bolo potvrdené, že viac ako 90% radiálnej distorcie sa dá odstrániť použitím prvých dvoch členov postupnosti.

1.2.4 Epipolárna geometria

Epipolárna geometria je projektívna geometria medzi dvomi pohľadmi. Nie je závislá na štruktúre scény, závisí na vnútorných a vonkajších parametroch kamier.



Obrázok 1.4: Epipolárna geometria.

Priamka prechádzajúca oboma centrami projekcie C a C' pretína projekčné plochy kamier v bodoch e a e' , ktoré sa nazývajú *epipóly*. Každá rovina, ktorá obsahuje túto priamku, sa nazýva *epipolárna rovina*. Bod P v priestore, sa zobrazí v zobrazovacích rovinách kamier do bodu p a p' . Epipolárna rovina obsahujúca bod P pretína zobrazovanie plochy kamier a vytvára priamky l a l' . Na týchto priamkach ležia obrazy bodu P .

Obraz p môže zodpovedať hocijakému bodu v priestore, ktorý leží na priamke $|CPl$. Keď premietneme priamku $|CPl$ do premietacej roviny druhej kamery, získame epipolárnu priamku, na ktorej sa budú nachádzať všetky obrazy všetkých bodov ležiacich na priamke $|CPl$. Tým pádom je obraz p' "nútený" ležať na epipolárnej priamke obrazu p , l' . Táto definícia je symetrická, takže aj pre obraz p platí, že bude ležať na epipolárnej priamke obrazu p' .

Vo väčšine prípadov epipolárne priamky sú rôznobežné. Iba ak by sa na snímání scény použili kamery ktorých projekčné plochy by ležali v jednej rovine. V tom prípade by priamka prechádzajúca C a C' by nepretínala ani jednu z rovín, epipóly by boli v nekonečne a epipolárne priamky by boli rovnobežné.

1.2.5 Fundamentálna matica

Fundamentálna matica je algebrickou reprezentáciou epipolárnej geometrie. Je to matica o rozmeroch 3×3 , a vyjadruje vzťah medzi zodpovedajúcimi si bodmi v snímkach z dvoch rôznych polôh.

Platí:

$$p^T F p' = 0 \quad (1.14)$$

$$l' = F p \quad (1.15)$$

kde p a p' sú zodpovedajúce si obrazy bodu P , F je fundamentálna matica a l' je epipolárna priamka vypočítaná pomocou bodu p .

Fundamentálnu maticu je možné zostaviť dvomi spôsobmi. Ak poznáme vonkajšie a vnútorné parametre kamery, matica sa vypočíta podľa nasledujúceho vzorca:

$$F = A_2^{-T} [t]_x R A_1^{-1} \quad (1.16)$$

kde A_1 a A_2 sú kalibračné matice kamier a R a t je transformácia (posunutie a rotácia), ktorá prenáša body zo súradnicovej sústavy jednej kamery do druhej.

Zápis $[t]_x$ je antisymetrická matica:

$$[t]_x = \begin{bmatrix} 0 & t_3 & -t_2 \\ -t_3 & 0 & t_1 \\ t_2 & -t_1 & 0 \end{bmatrix} \quad (1.17)$$

Druhý spôsob zostavenia fundamentálnej matice je možný aj keď nepoznáme vzájomnú polohu kamier, a ani vnútorné parametre kamier. Jediné, čo potrebujeme poznať, je minimálne osem párov zodpovedajúcich si bodov medzi snímkami. Je možné počítať fundamentálnu maticu nelineárne aj s poznaním len siedmich párov, no táto metóda je menej presná a výsledkom je niekoľko možných fundamentálnych matíc, z ktorých treba zvoliť najvhodnejšiu.

Ak poznáme bod $p_{1i} = [x_{1i}, y_{1i}, 1]^T$ ktorý zodpovedá bodu $p_{2i} = [x_{2i}, y_{2i}, 1]^T$ v druhom snímku, epipolárna rovnica bude vyzerat' nasledovne:

$$p_{2i}^T F p_{1i} = 0 \quad (1.18)$$

Tento vzorec môže byť prepísaný na:

$$u_i^T f = 0 \quad (1.19)$$

kde

$$u_i = [u_{1i} u_{2i}, v_{1i} u_{2i}, u_{2i}, u_{1i} v_{2i}, v_{1i} v_{2i}, v_{2i}, u_{1i}, v_{1i}, 1]^T \quad (1.20)$$

$$f = [F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33}]^T$$

Tento systém je následne riešený jedným z možných algoritmov, napríklad 8-point algoritmus, ktorý vyžaduje na správne riešenie práve osem alebo viac dvojíc korešpondujúcich si bodov.

1.3 Rekonštrukcia objektu do 3D priestoru

Pomocou obrazov bodov snímaného objektu, môžeme zrekonštruovať tento predmet do 3D priestoru. V obrázku (Obrázok 1.5: Princíp rekonštrukcie objektu) máme body P_1 a P_2 , ktoré majú obrazy p_1, p_2 a p_1', p_2' . Za predpokladu, že poznáme len správne dvojice zodpovedajúcich si obrazov bodu medzi snímkami z rôznych kalibrovaných kamier, môžeme nájsť bod P_1

vyslaním priamky z centra projekcie C , prechádzajúceho obrazom p_1 a priamky z centra projekcie C' , prechádzajúceho obrazom p_1' . Bod P_1 bude priesečník týchto priamok. Analogicky to platí aj pre bod P_2 .

V prípade nesprávneho určenia dvojice obrazov bodu, napríklad pomocou chybné dvojice p_1 a p_2' pre bod P_1 , nesprávne určíme pozíciu P_1 do bodu R_1 .

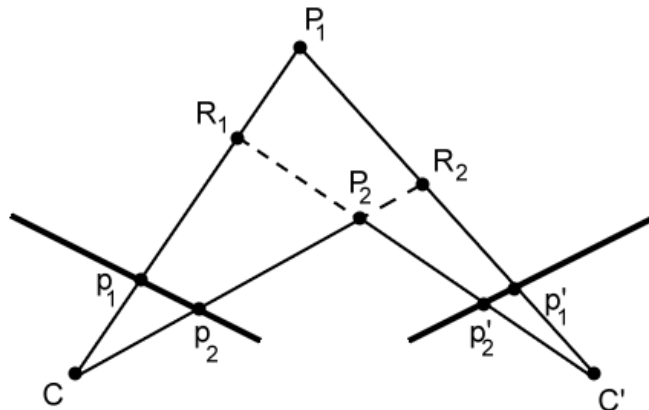
Na výpočte pozícií markerov v priestore sa priamo podieľajú vnútorné a vonkajšie parametre kamier. Pri vnútorných parametroch kamier poznáme pozíciu hlavného bodu $[cx, cy]$, a ohniskovú vzdialenosť f v pixloch. Ak uvažujeme súradnicovú sústavu kamery a pozícia obrazu p_1 je $[x_{p1}, y_{p1}]$, vektor $|p_1C|$ vyzerá nasledovne: $u = [x_{p1} - cx, y_{p1} - cy, f]$.

Kamery majú v priestore ešte vlastnú pozíciu a natočenie, určené vonkajšími parametrami kamier. Vektor u je potrebné ešte transformovať do svetovej súradnicovej sústavy:

$$u_w = Ru \quad (1.21)$$

kde R je rotačná matica kamery.

Pomocou takto získaného vektoru u_w a pozície centra bodu C získame rovnicu priamky, ktorá prechádza centrom projekcie C , obrazom p_1 a samozrejme hľadaným bodom P_1 .

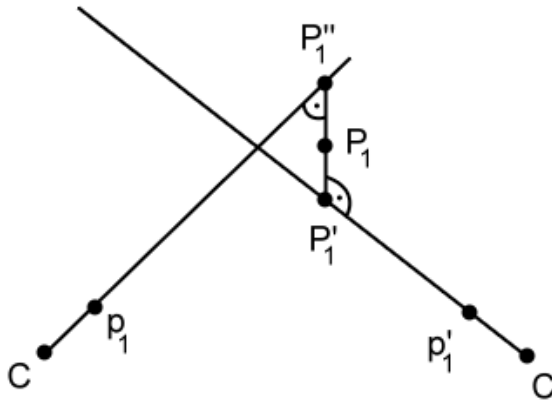


Obrázok 1.5: Princíp rekonštrukcie objektu.

V dôsledku šumu v snímkach, alebo nepresností v kalibračnej matici sa dve vyslané priamky vo všeobecnosti nepretnú v jedinečnom bode. Na trianguláciu je potrebné použiť metódu, ktorá relatívne presne odhadne najvhodnejší bod prieniku.

1.3.1 Mid-point triangulačný algoritmus

Bežne využívaná metóda na trianguláciu je nájdenie mid-pointu, bodu ktorý sa nachádza na najkratšej možnej úsečke spájajúcej lúče korešpondujúcich obrazov bodu. Najkratšia úsečka je kolmá na oba lúče. Táto metóda je nenáročná na výpočet.



Obrázok 1.6: Triangulácia metódou Mid-point.

Rovnice lúčov sú:

$$C + a(p_1 - C) = 0 \quad (1.22)$$

$$C' + b(p_1' - C') = 0$$

kde C a C' sú centrá projekcie a p_1 , a p_1' sú obrazy bodu P_1 .

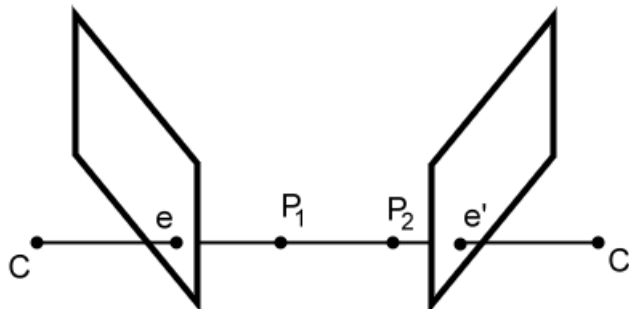
Riešením po nájdení a , b a c pomocou rovnice (tri koordináty, tri rovnice):

$$a(p_1 - C) - b(p_1' - C') = c((p_1 - C) \times (p_1' - C')) \quad (1.23)$$

môžeme vypočítať pozíciu bodu P_1 nasledovne:

$$P_1 = \frac{1}{2}((a(p_1 - C) + C) - (b(p_1' - C') + C'))$$

Môžu nastať prípady, keď pri určitej polohe kamier a pozícií obrazov sledovaného objektu nemožno použiť triangulačný algoritmus na získanie pozície objektu v 3D priestore. Tento prípad môže nastať, keď objekt leží na priamke spájajúcej centrá projekcie, a obrazy sa premietnu do epipólov. Túto situáciu znázorňuje obrázok (Obrázok 1.7).



Obrázok 1.7: Pozície bodov P_1 a P_2 nemôžu byť získané, keďže sa ich obrazy premietli do epipólov a vyslané lúče z centier projekcie C a C' by sa prekryvali po celej ich dĺžke.

2 Špecifikácia vlastného MC systému

Rozhodli sme sa napodobniť jednoduchý *pasívny optický MC systém*, pozostávajúci z markerov, sústavy dvoch kamier a programu, ktorý dokáže spracovať dáta zo spomínaných kamier. Všetky MC systémy využívajú drahé laboratórne vybavenie ako vysokorýchlostné kamery alebo špeciálne obleky. Motion Capture systém sme sa snažili vytvoriť z relatívne dostupných a nie drahých prostriedkov. Systém by mal byť schopný analyzovať videá s jednoduchými pohybmi herca, vypočítať pozície markerov v 3D priestore vzhľadom na určenú svetovú súradnicovú sústavu, vysporiadať sa s prípadným zakrytím markerov a mal by im priradiť správnu korešpondenciu v čase pri ich znovuobjavení. Výstupom by mali byť trajektórie, ktoré markery opisovali.

Markery

V pasívnom optickom MC systéme má herec oblečený špeciálny oblek s markermi umiestnenými na dôležitých častiach tela, väčšinou na kĺboch. Tieto markery reprezentujú pohyb, takže detekcia pohybu v scéne sa vzťahuje na detekciu pozícií markerov. Ako markery sme sa rozhodli použiť polovice oranžových loptičiek na stolný tenis. Oranžová farba sa v obraze pomerne jednoducho deteguje (viď. 2.2 Extrakcia pozícií markerov z obrazu) a loptičky majú vyhovujúcu veľkosť. Podmienkou úspešného snímania pohybu je tiež aby sa v snímanej scéne nevyskytovali iné predmety farby podobnej ako markery. Keďže sú si všetky markery podobné, nastane problém korešpondencie markerov medzi zábermi z jednotlivých kamier a tak isto problém korešpondencie markerov medzi zábermi v čase.

Na snímanie pohybu sme použili desať markerov. Jeden umiestnený na čiapke na hlave, deväť na hornej časti tela – plecia, lakte, zápästia, stred hrudníka a boky.



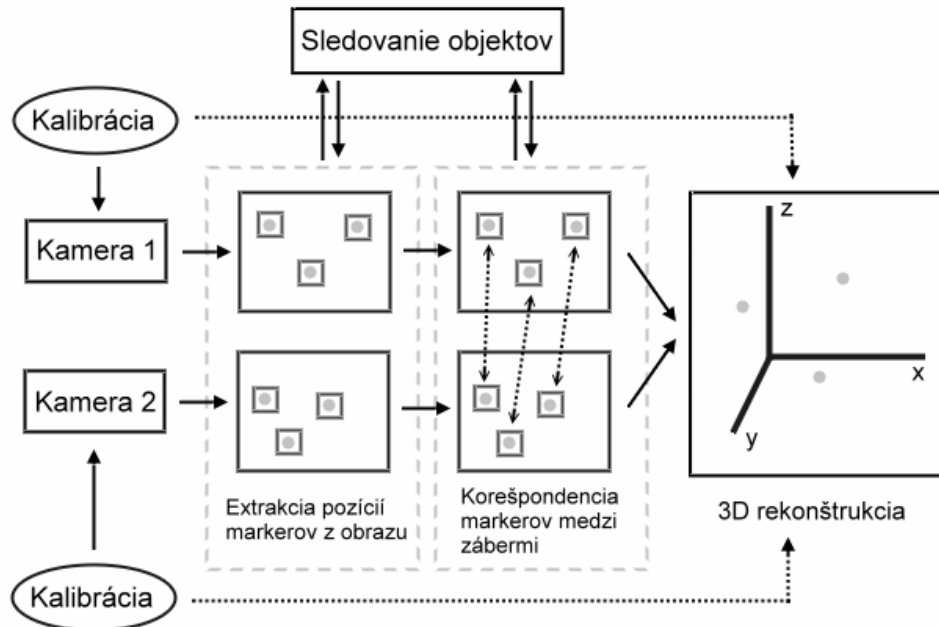
Obrázok 1.5: Rozmiestnenie markerov na hercovom tele.

Kamery

Na snímanie miestnosti sme použili dve kamery zapožičané Fakultou Informačných Technológií VUT v Brne.

System

Pohyb v počítačovom videní je tvorený sériou snímok, v ktorých sa sledovaný objekt nepohybuje. Ak chceme rekonštruovať pohyb, musíme zdrojové video rozdeliť na snímky a postupne spracovávať jednu za druhou. V nasledujúcich kapitolách opisujeme získavanie zdrojových videí a ich následné spracovanie.



Obrázok 2.1: Blokový diagram systému.

2.1 Kalibrácia kamier a určenie počiatku svetovej súradnicovej sústavy

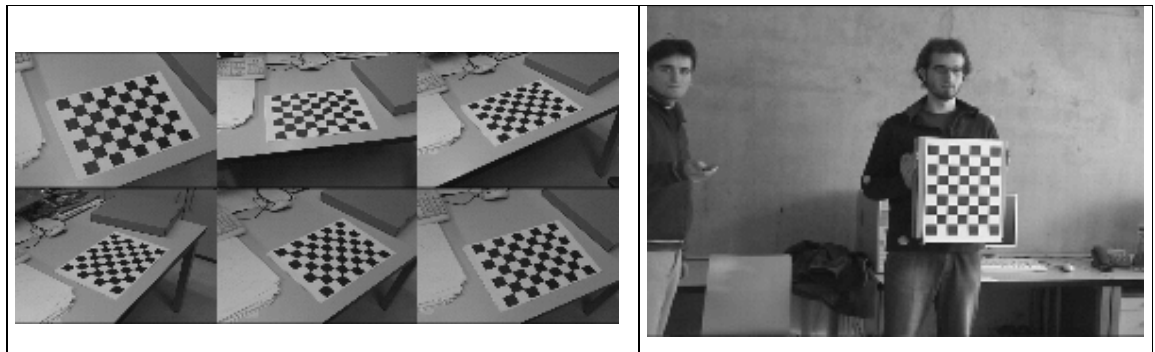
Celý proces snímania pohybu začína snímaním zdrojového videa, ktoré má byť spracované. Do miestnosti, v ktorej sa herec bude pohybovať je potrebné rozmiestniť kamery, ktorých parametre musíme poznať, aby sme boli schopní správne vypočítať pozície markerov. Kalibrovať kameru znamená nájsť jej vnútorné a vonkajšie parametre. Existuje viacero metód kalibrácie kamery.

Rozdeľujú sa na automatické a fotogrametrické. Automatické metódy nepotrebujú kalibračný predmet a odhadujú vnútorné parametre kamery pomocou nájdenia zodpovedajúcich si bodov v sledovanej scéne. Fotogrametrické metódy majú využívať predmety so známou geometriou, ktoré sú nasnímané z viacerých uhlov. Oproti automatickým kalibračným metódam dosahujú presnejšie výsledky, ale na druhú stranu sú náročnejšie na čas a laboratórne vybavenie.

Kalibrácia pomocou rovinného kalibračného predmetu

Na kalibráciu kamier používame fotogrametrickú metódu z knižnice OpenCV [2], ktorá využíva Zhangov kalibračný algoritmus. Pre dostatočnú presnosť je potrebných aspoň šesť snímok kalibračného predmetu (šachovnice) z rôznych uhlov (viď. Obrázok 2.2). Týmto spôsobom získame vnútorné parametre kamery a parametre distorcie šošoviek kamery.

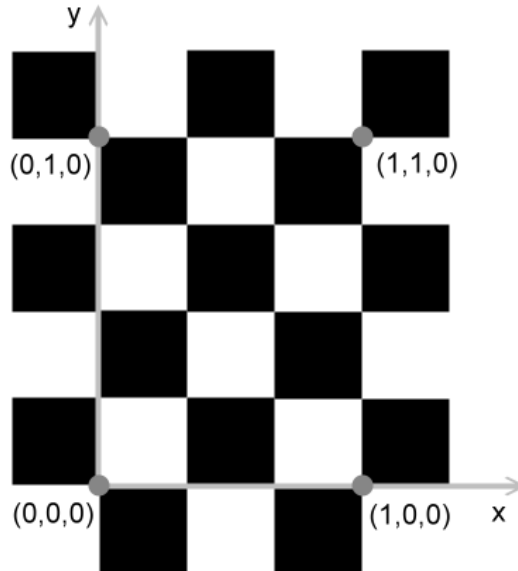
Vonkajšie parametre kamier môžu byť získané po rozmiestnení kamier do priestoru. Poloha a natočenie kamier by mala byť zvolená vhodne tak, aby kamery sledovali spoločnú oblasť, v ktorej sa herec bude pohybovať. Ak predpokladáme, že s kamery sú umiestnené fixne a už sa s nimi viac hýbať nebude, do miestnosti umiestnime kalibračný predmet. Kalibračné snímky zo všetkých kamier musia obsahovať celý kalibračný predmet, ktorý zároveň určuje súradnicový systém, vzhľadom na ktorý budú vypočítané pozície kamier a zároveň vzhľadom na ktorý budú vypočítané aj pozície markerov. Pomocou týchto snímok dostaneme ako výstup rotačnú maticu a translačný vektor. Kalibračný predmet môžeme zo snímanej oblasti odstrániť.



Obrázok 2.2: Vľavo séria obrázkov kalibračného predmetu na získanie vnútorných parametrov, vpravo kalibračná snímka na získanie aktuálnej polohy a natočenia kamery v priestore.

2.1.1 Určenie počiatku svetovej súradnicovej sústavy

Definovaním bodov *svetovej súradnicovej sústavy* do obrazov rohov kalibračného predmetu, ako zobrazuje obrázok (Obrázok 2.3: Na rovinnom telese určené body svetovej súradnicovej sústavy), určujeme počiatok svetovej súradnicovej sústavy a jej dimenzionálne osi. Súradnicové osi x a y môžeme určiť práve vektormi v jednotlivých obrazoch kalibračného predmetu a os z býva väčšinou definovaná smerom ku kamere.



Obrázok 2.3: Na rovinnom telese určené body svetovej súradnicovej sústavy.

Pri kalibrácii sme použili šachovnicu s dĺžkou strany štvorca 40 mm. Určením reálnych jednotiek umožnilo získať translačný vektor kamery v reálnych jednotkách. Toto sa dá využiť pri triangulácii bodov pomocou viacerých pohľadov, ak chceme poznať pozície bodov v priestore tak, aby sme mohli konfrontovať výsledky so skutočnosťou.

2.2 Extrakcia pozícií markerov z obrazu

Pri spracovaní videa je prvým krokom k úspešnej 3D triangulácii markerov do priestoru je čo najpresnejšia extrakcia ich 2D pozícií v snímkach z kamier. Rozlíšenie markerov od pozadia prebieha na základe rozpoznania farby. Pre ľahšiu detekciu oranžovej farby markeru sa snímka prevedie do farebného priestoru YUV .

Farebný priestor YUV je odvodený od RGB modelu. R , G a B zložky sú pomerovo sčítané a vytvárajú zložku Y , ktorá vyjadruje celkovú jasnosť:

$$Y = 0.299 * R + 0.587 * G + 0.114 * B \quad (2.1)$$

U zložka je vypočítaná odčítaním Y od zložky B a zložka V odčítaním Y od R :

$$\begin{aligned} U &= 0.436 * (B - Y) / (1 - 0.114) \\ V &= 0.615 * (R - Y) / (1 - 0.299) \end{aligned} \quad (2.2)$$

V takto získanom farebnom priestore sa zodpovedajúca oranžová farba markeru nachádza v oblasti:

$$\begin{aligned} -0.436 < U < -0.150 \\ 0.150 < V < 0.436 \end{aligned}$$

a zložka Y závisí od osvetlenia aké pri snímaní videa použijeme.

Algoritmus na získanie platných pozícií markerov vyzerá nasledovne:

1. Prevod snímku z farebného priestoru RGB do priestoru YUV.
2. Ohodnotenie každého pixlu hodnotou '1', ak sa jeho farba nachádza vo vymedzenej oblasti pre oranžovú farbu, hodnotou '0' ak nie.
3. Extrakcia objektov zo získaného dvojúrovňového snímku pomocou semienkového algoritmu.
4. Eliminácia jednoznačne nepravdivých objektov (napríklad príliš veľký objekt - časť ľudskej kože, alebo objekty veľkosti jednotiek pixlov - šum)
5. Výpočet ťažiska pre každý nájdený objekt pomocou vzorca:

$$x_t = \frac{1}{n} \sum_{i=1}^n x_i \text{ a } y_t = \frac{1}{n} \sum_{i=1}^n y_i$$

kde x_t a y_t sú koordináty ťažiska objektu, x_i a y_i sú koordináty bodov objektu a n je ich počet.

2.2.1 Problémy pri detekcii objektov

Ak sa pri snímaní videa použije zdroj osvetlenia z jedného bodu, markery mávajú z rôznych uhlov rozličné odtiene oranžovej farby. Preto je lepšie posunúť hranice pre zložky Y , U a V ručne pre každú kameru, aby nenastávali nepravdivé detekcie napríklad ľudskej kože.

Ďalšie nejednoznačnosti vznikajú pri rýchlejších pohyboch herca. V dôsledku dlhšieho expozičného času a prekladaného riadkovania kamery je marker buď pretiahnutý, alebo vplyvom rýchleho posunu tvorí dva objekty blízko seba (vid'. Obrázok 2.4: Detekcia dvoch objektov miesto jedného vplyvom rozmazania). Ak je marker rozmazaný, vzniknú isté nepresnosti pri vypočítavaní ťažiska. Ohľad treba brať aj na to, že väčší objem objektu môže tiež znamenať že dva alebo viac markerov sú blízko pri sebe a „spájajú“ sa do jedného objektu. Zohľadnenie týchto skutočností je spomínané v nasledujúcej kapitole.



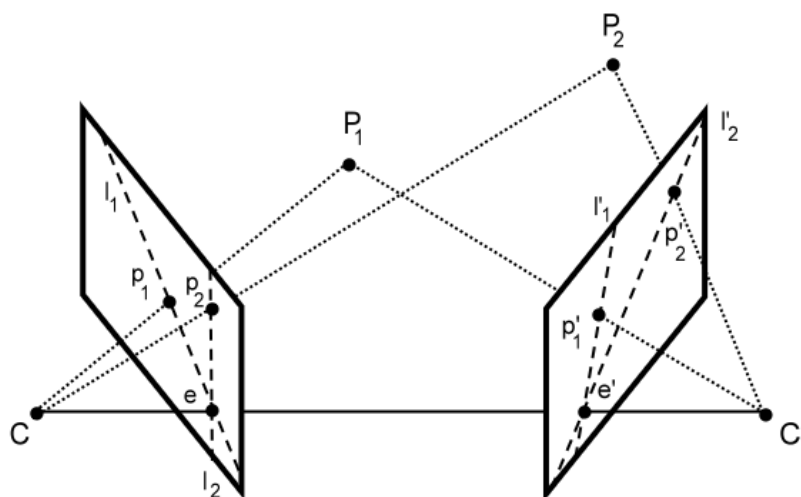
Obrázok 2.4: Detekcia dvoch objektov miesto jedného vplyvom rozmazania.

2.3 Korešpondencia objektov medzi snímkami z kamier

Informácia o korešpondencii objektov medzi snímkami je kľúčová k rekonštrukcii pozície markeru do 3D priestoru. Keďže všetky sledované markery majú rovnaký vzhlad, a nedá sa určiť na základe žiadnej vlastnosti, že sa jedná o ten istý marker, nastáva otázka: „Ktorý objekt v snímku z kamery1 zodpovedá ktorému objektu v snímku z kamery2?“ Ak v oboch snímkach existuje n obrazov, existuje $n!$ rôznych priradení obrazov z jedného snímku obrazom z druhého, ale len jedno je správne.

2.3.1 Určenie korešpondencie objektov medzi snímkami

Vychádzame z predpokladu, že na oboch snímkach z kamier sa nachádza rovnaký počet objektov. Priradenie korešpondencií objektov medzi týmito snímkami len na základe poznania rozmiestnenia markerov na hercovi by bolo veľmi nepresné a často by dochádzalo k zostaveniu nesprávnej dvojice korešpondujúcich bodov, čo by viedlo k nesprávnemu výpočtu pozície markeru v 3D priestore. Veľmi dobrý nástroj na jednoznačné určenie korešpondencie objektov medzi zábermi z rôznych kamier, sú poznatky z epipolárnej geometrie.



Obrázok 2.5: Určenie korešpondujúcich bodov.

Obrázok (Obrázok 2.5: Určenie korešpondujúcich bodov) ilustruje situáciu, kde máme v každom snímku obrazy dvoch bodov P_1 a P_2 . Za predpokladu, že scénu sme snímali kalibrovanými kamerami, môžeme pomocou vzorca (1.16) vypočítať fundamentálnu maticu F medzi týmito pohľadmi. Postup k nájdeniu korešpondujúcich objektov bude nasledovný:

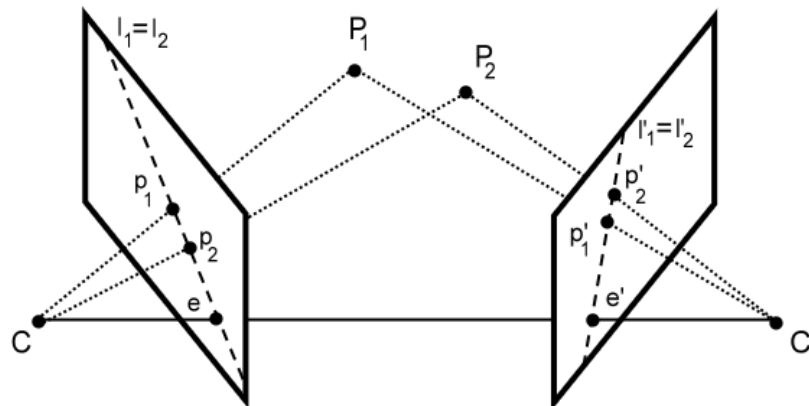
1. Pomocou vzorca (1.15) a pozície objektu p_i nájdeme epipolárnu priamku l'_i . Takto sa hľadanie zodpovedajúceho objektu redukuje na 1D priestor.

2. Nájďeme objekt v druhom snímku, ktorý sa nachádza najbliďšie k priamke l'_i . Označíme ho za priradený.
3. Postupujeme podobne pre ostatné objekty v prvom snímku.

Týmto spôsobom zostavíme dvojice korešpondujúcich objektov, ktoré využijeme pri triangulácii.

2.3.2 Problémy nejednoznačnosti priradenia objektu

Nejednoznačnosť priradenia objektov vzniká v prípade, keď pre zvolený objekt p_1 v snímku z prvej kamery, existuje viacero objektov ležiacich, v respektíve nachádzajúcich sa dostatočne blízko na to, aby boli považované za korešpondujúci objekt, na epipolárnej priamke l_1' v snímku z druhej kamery. Keďže leží viacero objektov na epipolárnej priamke l_1' vypočítanej pomocou pozície objektu p_1 zo snímku z prvej kamery, bude ležať rovnaký počet objektov (ak náhodou niektorý nie je prekrytý alebo mimo záberu) na epipolárnej priamke l_1 , vypočítanej pomocou pozície ľubovoľného bodu ležiaceho na priamke l_1' .



Obrázok 2.6: Obrazy bodov P_1 a P_2 ležia na tej istej epipolárnej priamke.

Správne priradenie objektu p_1 je možné dosiahnuť analýzou jeho polohy voči ostatným objektom na priamke l_1 . Zodpovedajúci objekt bude mať rovnakú polohu medzi objektmi ležiacimi na priamke l_1' . Na situácii v obrázku (Obrázok 2.6) sa od obrazu p_1 „vpravo“ nachádza jeden objekt, takže budeme hľadať objekt na druhom snímku ležiaci na epipolárnej priamke l'_1 , „vpravo“ od ktorého sa nachádza tiež jeden objekt ležiaci na tej istej priamke.

Tento spôsob priradenia korešpondencií objektov je možné najúčinnnejšie použiť na začiatku snímania, keď herec stojí v základnej polohe a približne všetky markery sú viditeľné a ešte nie je inicializovaný *Kalmanov filter* na sledovanie objektov.

2.3.3 Problémy pri rozdielnych počtoch objektov v snímkach

Pri snímaní videa môžu nastať nasledovné prípady prekrytia markeru, ktoré môžu spôsobiť rozdielne počty objektov v snímkach:

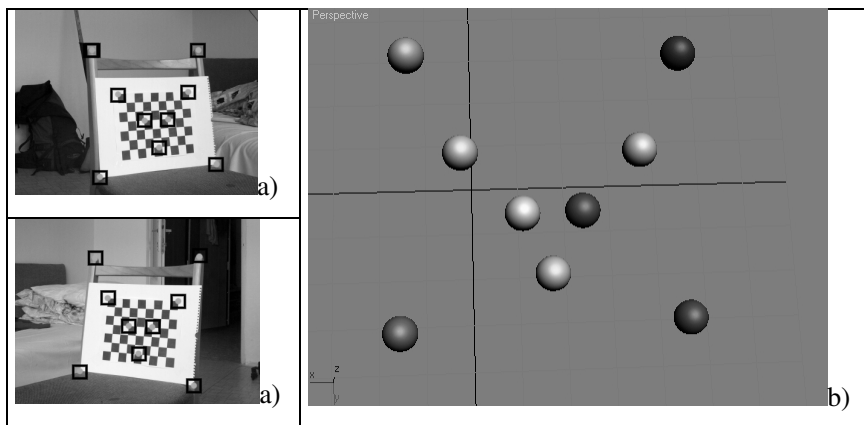
1. Dva, alebo viac obrazov markerov sa spoja do jedného objektu
2. Marker je čiastočne prekrytý, na jednom snímku má svoj obraz a na druhom nie.
3. Marker je prekrytý úplne, nemá obraz na žiadnom snímku

Situáciu, ktorá je spôsobená prvým prípadom, je možné riešiť najprv priradením korešpondencie objektom ktoré majú zodpovedajúce objekty v druhom snímku. Nasledovne skúsime pre objekty, ktoré nenašli korešpondenciu v prvom kroku, nájsť na epipolárnej priamke v druhom snímku objekty s väčším objemom ako je priemerný objem objektu. Tento objekt môže byť zodpovedajúcim objektom, pretože mohol vzniknúť spojením viacerých obrazov.

V druhom prípade nastáva situácia, pri ktorej ostane objekt s nepriradeným korešpondujúcim objektom (pretože nebol nájdený), tým pádom nedokážeme určiť 3D pozíciu markeru. V treťom prípade nám bude chýbať informácia o výskytu markeru úplne. Oba tieto problémy sa dajú riešiť pomocou sledovania objektov Kalmanovým filtrom (viď. 2.5.2 Využitie Kalmanovho filtru za zdokonalenie algoritmu priraďovania korešpondencií objektov medzi snímkami z rôznych kamier).

2.4 Výpočet 3D pozície markerov v priestore

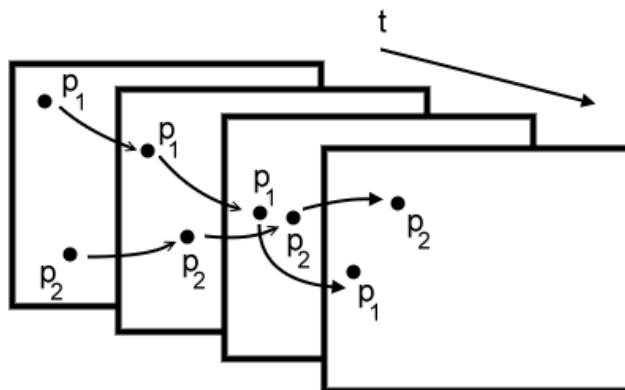
Zo známych dvojíc korešpondujúcich bodov medzi snímkami z kamier vypočítame pomocou triangulačného algoritmu (1.3.1 Mid-point triangulačný algoritmus) pozície markerov v 3D priestore, vzhľadom na počiatok svetovej súradnicovej sústavy (2.1.1 Určenie počiatku svetovej súradnicovej sústavy). Tieto markery rekonštruujú snímaný objekt v časovom okamihu, a je potrebné ich zaznamenávať, aby sme na konci spracovania videa mohli spojiť dáta so všetkých časových okamihov do ilúzie pohybu.



Obrázok 2.7: a) snímka z ľavej a pravej kamery, b) pozície markerov zobrazené v 3D priestore.

2.5 Korešpondencia objektov v čase

So zmenou času objekty menia svoju pozíciu. Postupnosť bodov, v ktorých sa objekt v istom čase nachádzal sa nazýva trajektória. Aby sme získali jednotlivé trajektórie objektov, musíme každému objektu nájsť správnu korešpondenciu s objektom v predchádzajúcom snímku. Nesmie nastať zámena medzi objektmi, inak by sa trajektórie „krížili“ a nezodpovedali by skutočnosti.



Obrázok 2.8: Korešpondencia objektov v čase.

Často používaný spôsob sledovania bodov je hľadanie k pozícii objektu v čase t najbližšiu pozíciu objektu v čase $t-1$. S touto jednoduchou metódou sme nedosiahli dostatočné experimentálne výsledky a často dochádzalo k zámene objektov. Preto je potrebné zvoliť metódu, ktorá berie v úvahu pozíciu a rýchlosť objektu a predpokladá jeho pozíciu v ďalšom časovom okamihu. Tieto vlastnosti spĺňa Kalmanov filter. V programe využívame implementáciu Kalmanovho filtru z knižnice Intel OpenCV [2].

2.5.1 Kalmanov filter

Kalmanov filter je efektívny rekurzívny filter, ktorý predpokladá stav dynamického systému na základe sérií nekompletných a zašumených meraní. Zostavuje na základe minulého a aktuálneho merania optimálny predpoklad budúceho stavu a zároveň počíta s neistotou merania. Kalmanov filter potrebuje počiatočný odhad, aby mohol správne pracovať.

Pre popis pohybu objektu vezmeme v úvahu jeho pozíciu, rýchlosť a zrýchlenie. Nasledujúci stav objektu popisuje rovnica:

$$\begin{bmatrix} x(t+1) \\ v(t+1) \\ a(t+1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} x(t) \\ v(t) \\ a(t) \end{bmatrix} \quad (2.3)$$

Kde x je pozícia objektu, v je jeho rýchlosť a a zrýchlenie. Tieto tri parametre určujú stavový vektor objektu.

Sledovanie korešpondencie objektov pomocou Kalmanovho filtru bude vyzerat' nasledovne:

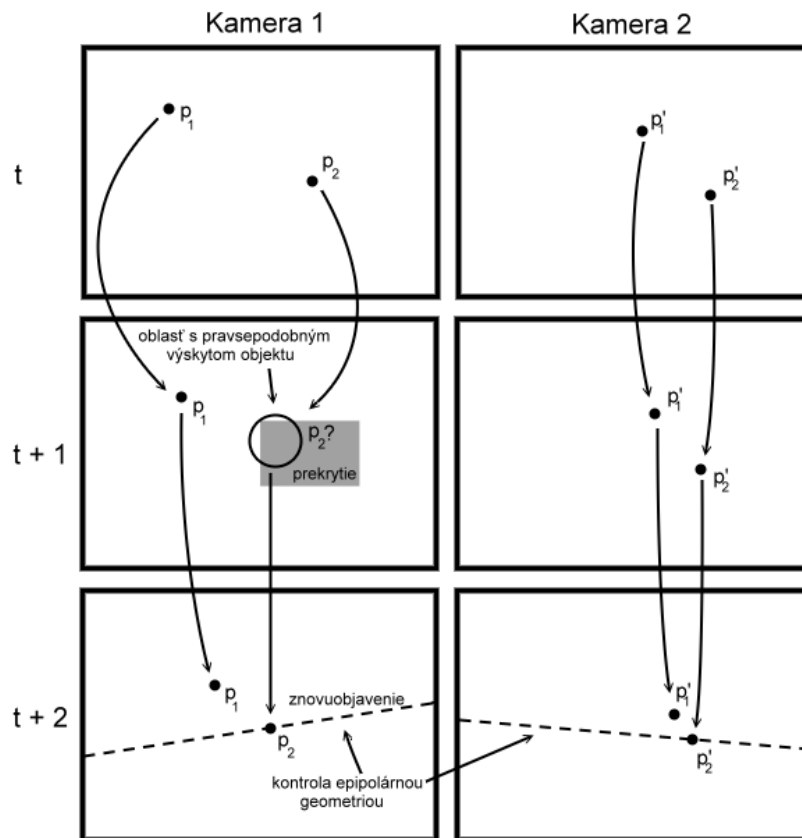
1. Inicializácia Kalmanovho filtra pre každý nájdený objekt použitím jeho polohy v čase t .
2. V čase $t+1$ Kalmanov filter vykoná predikciu polohy každého objektu. Zodpovedajúce objekty budeme hľadať v najbližšom okolí ich predpokladanej polohy a túto informáciu o korešpondencii v čase zaznamenávať.
3. Ak na predpokladanom mieste nenájdeme objekt, pravdepodobne bol marker v zábere prekrytý. Na tomto mieste vytvoríme objekt, počítame s ním ako keby bol naozaj detegovaný a budeme predpokladať, že v nasledujúcich snímkach sa objaví. Ak sa neobjaví po dlhšej dobe, vyhlásime objekt za „stratený“ a trajektóriu ukončíme.
4. Spravíme korekciu Kalmanovho filtru získanými pozíciami objektov.
5. Pokračujeme v bode 2.

Obrovskou výhodou sledovania objektov Kalmanovým filtrom je práve možnosť predpokladu pozície objektu ak bol marker prekrytý. Predstavme si, že herec vykoná pohyb rukou, pri ktorom si prekryje na niekoľko snímkov marker na tele. Bez informácií o obrazoch tohto markeru v snímkach by sme neboli schopní zrekonštruovať jeho pozíciu do 3D priestoru. Kalmanov filter ešte pred prekrytím mal informácie o polohe a rýchlosti markeru a po prekrytí na základe týchto informácií dokáže predpokladať jeho pozíciu. S využitím týchto informácií môžeme dočasne pracovať, kým sa marker znovu neobjaví.

2.5.2 Využitie Kalmanovho filtru za zdokonalenie algoritmu prirad'ovania korešpondencií objektov medzi snímkami z rôznych kamier

Predstavme si, že v snímkach z kamery1 a kamery2 sme našli v čase t objekty p_1, p_2 a p'_1, p'_2 (viď. Obrázok 2.9: Vyrovnanie sa s prekrytím objektu) zodpovedajúce obrazom markerov v scéne. Každému z nich je priradený Kalmanov filter. Korešpondenciu medzi objektmi medzi snímkami môžeme na začiatku snímania určiť pomocou epipolárnej geometrie, týmto získame dvojice objektov (p_1, p'_1) a (p_2, p'_2) . V nasledujúcom snímku, v čase $t+1$, stačí pomocou Kalmanovho filtra spraviť predikciu polôh dvojíc objektov, k týmto polohám nájsť objekty zo snímok a tým pádom získame korešpondujúce dvojice bodov medzi snímkami z kamier. Kontrolu správnosti dvojice je možné skontrolovať pomocou epipolárnej geometrie, kde predpokladáme, že prvý bod dvojice leží na epipolárnej priamke vypočítanej pomocou druhého bodu dvojice a naopak.

Tento spôsob hľadania korešpondencie sa pomocou predikcie Kalmanovho filtra dokáže vyrovnávať s krátkodobými prekrytiami markerov.



Obrázok 2.9: Vyrovnanie sa s prekrytím objektu.


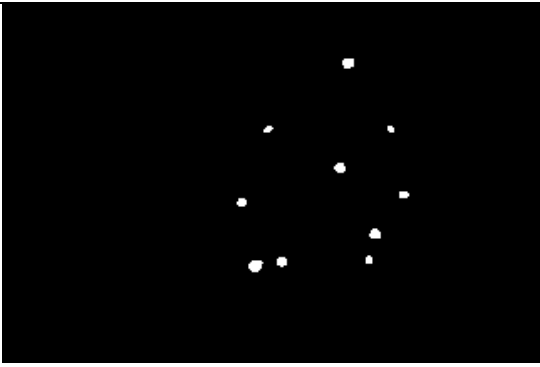
3 Dosiiahnuté výsledky

V nasledujúcej kapitole sú zhrnuté výsledky z jednotlivých častí programu, ktoré sme testovali na zdrojových videách. Zdrojové videá mali rozlíšenie 720 x 480 pixlov, a kamera mala snímáciu frekvenciu 25 snímok za sekundu. Na hercovi bolo umiestnených desať markerov. Počiatok svetovej súradnicovej sústavy bol určovaný pomocou kalibračnej šachovnice tesne pred hercom, ako zobrazuje obrázok (Obrázok 2.2 vpravo). Herec vykonával jednoduché, pomalé pohyby, aby nedochádzalo k častým zakrytiam markerov a ani ich rozmazaniu v snímkach. Zdrojové videá sú umiestnené na CD s programom v prílohe (Príloha 2). Na overenie správnosti jednotlivých častí spracovania dát boli použité výstupy programu.


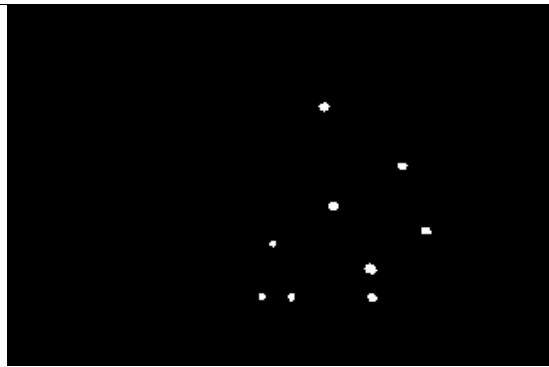
3.1 Test detekcie objektov v snímkach

V nasledujúcich troch testoch otestujeme úspešnosť detekcie rôzne osvetlených markerov oranžovej farby v snímkach z kamier. V prvých dvoch testoch sú parametre pre detekciu oranžovej farby nastavené optimálne, ako v kapitole 2.2 Extrakcia pozícií markerov z obrazu. V treťom prípade sú parametre upravené, aby bolo možné detegovať takmer zakrytý marker.


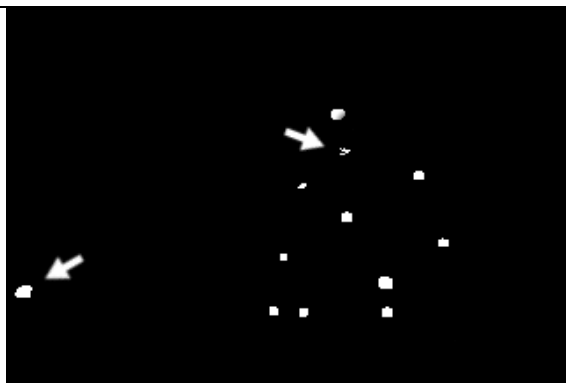
Test 1: Optimálne nastavené parametre pre oranžovú farbu – správna detekcia

	
Počet markerov v obraze	10
Počet detegovaných markerov	10
Úspešnosť	100%

Test 2: Optimálne nastavené parametre pre oranžovú farbu, jeden takmer zakrytý marker – jeden marker nedetegovaný.

	
Počet markerov v obraze	10
Počet detegovaných markerov	9
Úspešnosť	90%

Test 3: Upravený parameter V pre oranžovú farbu, jeden takmer zakrytý marker – vplyvom úprav nesprávna detekcia časti kože a časti vybavenia miestnosti

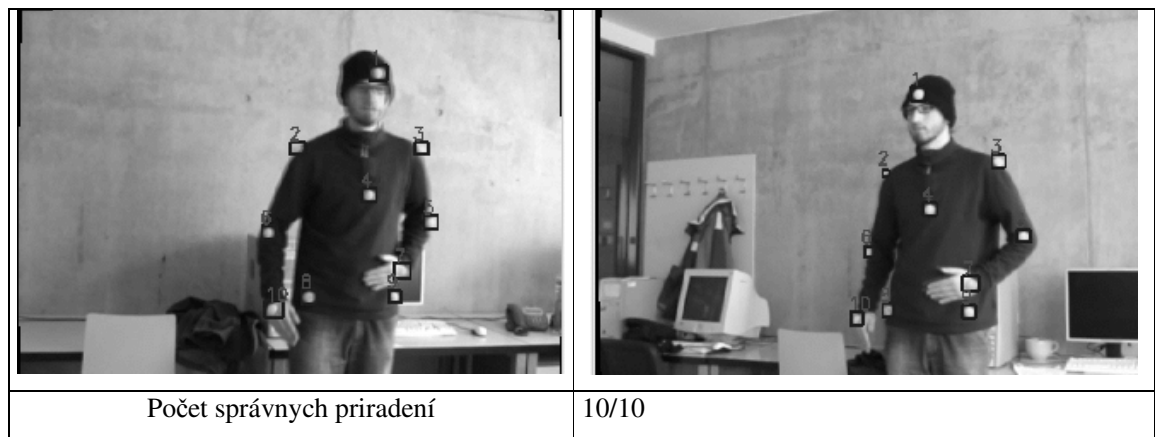
	
Počet markerov v obraze	10
Počet detegovaných markerov	12
Úspešnosť	-

V konečnom dôsledku je lepšie pre každú kameru na začiatku spracovania dát nastaviť ručne parametre pre oranžovú farbu, keďže z každého uhlu je scéna inak osvetlená. Falošné detekcie kože alebo častí scény sú väčšinou eliminované mechanizmom porovnávania veľkosti plôch objektov s priemernými veľkosťami. Ak niektoré falošné detekcie ostanú klasifikované ako platné objekty, pri ďalšom spracovaní nebudú mať korešpondujúci obraz v druhom snímku, a na záver spracovania môžu byť označené ako „neplatné“.

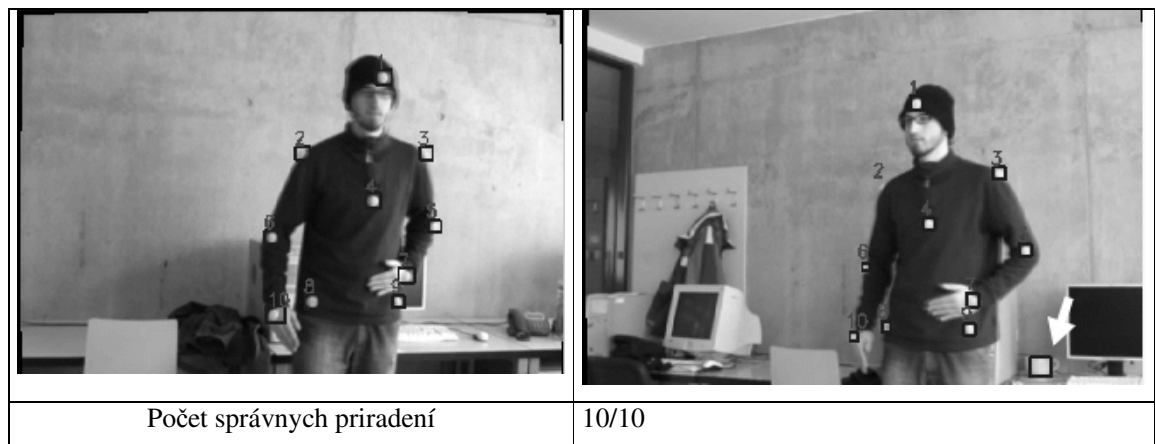
3.2 Test správneho priradenia korešpondencie medzi snímkami z kamier

Predpokladom k úspešnej 3D rekonštrukcii je správne priradenie korešpondencií medzi snímkami z kamier. Mechanizmus správneho priradenia tvorí epipolárna geometria a na jeho udržiavanie Kalmanov filter. V nasledujúcich testoch sme testovali rozličné prípady rozloženia obrazov a nejednoznačnosti priradenia korešpondencie.

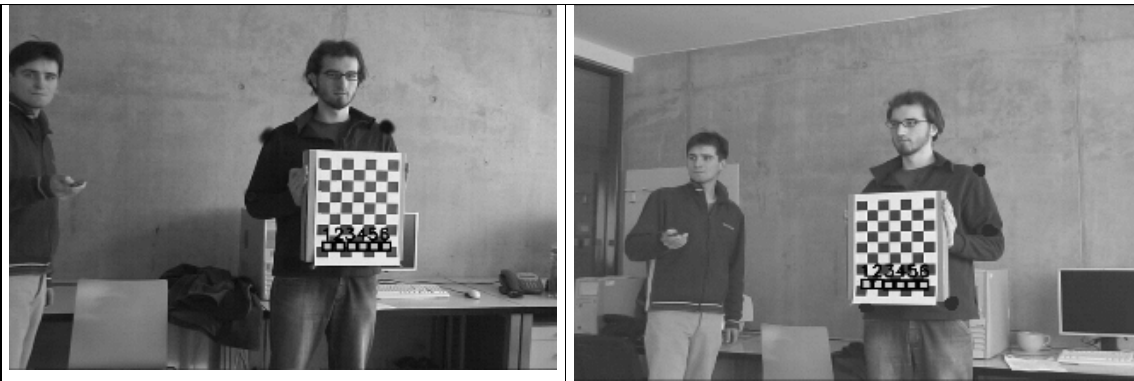
Test 1: Ideálne podmienky hľadania korešpondencie, keďže na oboch snímkach sú všetky obrázky markerov viditeľné.



Test 2: Aj keď bol detegovaný falošný objekt, nevplyva na priradenie ostatných objektov a korešpondenciu v druhom snímku nemá.



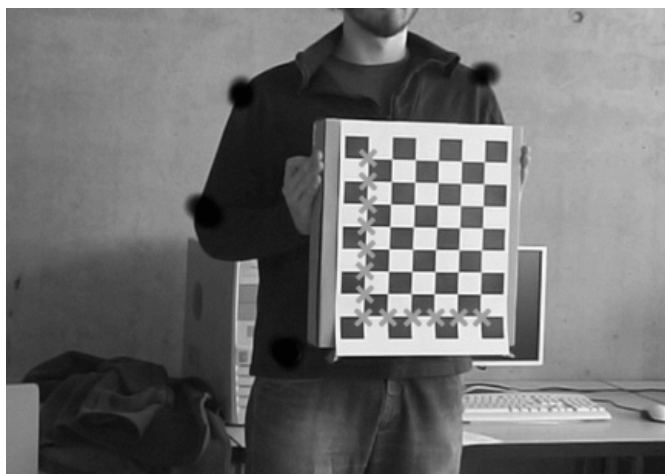
Test 3: Viac obrazov leží na jednej epipolárnej priamke – rozhodovanie podľa relatívnych pozícií.

	
Počet správnych priradení	6/6

Algoritmus pre priraďovanie korešpondencií pomocou epipolárnej geometrie pracuje spoľahlivo. Jediné prípady v ktorých dochádzalo k nesprávnym korešpondenciám boli ak bol korešpondujúci marker prekrytý, a na epipolárnej priamke sa nachádzal iný obraz markeru, alebo falošne detegovaný objekt. Po inicializácii a prvých pár krokoch Kalmanovho filtru už funguje aj kontrola pomocou predpokladaných pozícií objektov.

3.3 Test presnosti výpočtu triangulačného algoritmu

Pri tomto teste sme umiestnili markery na rohy kalibračnej šachovnice pozdĺž osí x a y , ako je zobrazené na obrázku (Obrázok 3.1: Rozmiestnenie markerov na kalibračnej šachovnici). Keďže poznáme skutočné vzdialenosti medzi markermi – 40 mm, môžeme výpočty triangulačného algoritmu konfrontovať so skutočnými hodnotami.



Obrázok 3.1: Rozmiestnenie markerov na kalibračnej šachovnici.

	Skutočná pozícia markeru [x, y, z] [mm]	Vypočítaná pozícia markeru [x, y, z] [mm]	Odchýlka [mm]
1.	[0, 0, 0]	[-0.7, -0.2, 3.3]	3.37
2.	[40, 0, 0]	[39.7, 0.9, -1.6]	1.86
3.	[80, 0, 0]	[78.8, 1.1, -1.1]	1.56
4.	[120, 0, 0]	[120.6, -0.0, 0.8]	1.00
5.	[160, 0, 0]	[160.9, 0.2, 0.5]	1.04
6.	[200, 0, 0]	[199.7, 0.2, 3.2]	3.22
7.	[0, 40, 0]	[-1.4, 40.9, 1.7]	2.37
8.	[0, 80, 0]	[-2.0, 82.1, 0.1]	2.90
9.	[0, 120, 0]	[0.0, 120.6, -1.7]	1.80
10.	[0, 160, 0]	[-1.9, 162.0, 0.5]	2.80
11.	[0, 200, 0]	[-1.1, 200.6, 2.6]	2.88
12.	[0, 240, 0]	[0.2, 279.4, -1.0]	1.18
Priemerná odchýlka [mm]			2.165

Tabuľka 3.1: Porovnanie výsledkov so skutočnými hodnotami

Za týchto podmienok odchýlka od skutočných hodnôt nepresahuje štyri milimetre. Na jednej strane tieto výsledky potvrdzujú správnu kalibráciu kamier, v respektíve správne určenie vnútorných a vonkajších parametrov, keďže sa priamo podieľajú pri výpočte 3D pozície markeru. Na druhej strane ukazujú, že triangulačný algoritmus Mid-point je dostatočne presný.

Pri extrakcii pozícií markerov však vznikajú chyby spôsobené nejednoznačnou detekciou ťažiska markeru napríklad v prípade, keď je viditeľná len časť markeru, alebo má vplyvom rozmazania väčší objem. V týchto prípadoch je vypočítaná pozícia posunutá voči reálnej aj o niekoľko jednotiek centimetrov.

3.4 Test korešpondencie markerov v čase

Korešpondencia markerov v čase a vyrovnávanie sa s miznutím a znovuobjavením markerov je dôležitým predpokladom k úspešnému zostaveniu trajektórií. V nasledujúcich testoch budeme overovať správnosť priradenia identity markeru po prekrytí. V tabuľkách sú uvedené hodnoty pozície zmiznutia objektu, predpokladanej pozície objektu v snímke tesne pred objavením a skutočnej pozície po objavení. Znak 'x' v obrázkoch zobrazuje predpokladanú pozíciu markeru.

Test 1: Zakrytie markeru na dobu – 9 snímok. Úspešné znovu priradenie.



Pozícia zmiznutia [x, y]	Predpokladaná pozícia objavenia [x, y]	Skutočná pozícia objavenia [x, y]
[360, 278]	[329, 269]	[351, 279]

Test 2: Zakrytie markeru telom na krátku dobu – 21 snímok. Úspešné znovu priradenie.



Pozícia zmiznutia [x, y]	Predpokladaná pozícia objavenia [x, y]	Skutočná pozícia objavenia [x, y]
[602, 176]	[508, 158]	[527, 162]

Test 3: Zakrytie markeru telom na dobu – 32 snímok. Neúspešné priradenie.







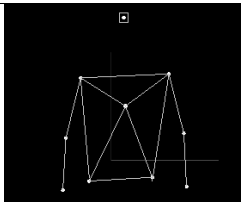
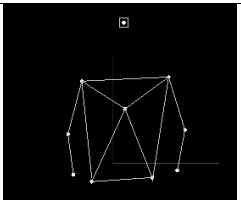
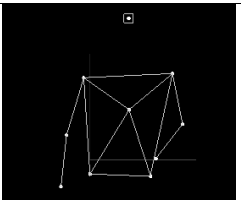
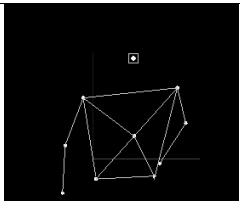
Pozícia zmiznutia [x, y]	Predpokladaná pozícia objavenia [x, y]	Skutočná pozícia objavenia [x, y]
[616, 163]	[468, 142]	[569, 152]

Algoritmus sa dokáže vyrovnávať s krátkodobými prekrytiami markeru. Nejednoznačnosť nastáva v momentoch ako v teste 3, keď prekrytý marker opisoval inú trajektóriu, akú predpokladal Kalmanov filter.

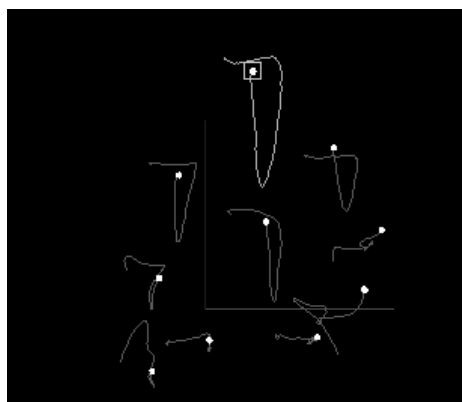
3.5 Test rekonštrukcie pohybu

Cieľom tohto testu bolo z krátkych zdrojových videí rekonštruovať pohyb, ktorý vykonával herec. Dôležitým faktorom bolo tiež, aby nedochádzalo k zámene identít markerov a rekonštrukciám nesprávnych trajektórií.

Test 1: Spracovanie videa s pohybom „poklona“.

Snímka 1	Snímka 6	Snímka 23	Snímka 48
			
			

V predchádzajúcej tabuľke sú zobrazené v hornom rade niektoré snímky z videa s pohybom „poklona“ a v dolnom rade zrekonštruované situácie. Na snímku 6 je možné vidieť zakrytie markeru na pravom boku herca rukou. Pre výpočet 3D pozície zakrytého markeru je použitý predpoklad Kalmanovho filtru. V ďalšom snímku je už marker znovu viditeľný a správne priradený.



Obrázok 3.2: Trajektórie markerov pohybu „poklona“.

Záver

Hlavný cieľom tejto práce bolo preštudovať dostupnú literatúru zaoberajúcou sa procesmi identifikácie pohybu a jeho následnou rekonštrukciou. Na základe získaných poznatkov zvolil vhodný prístup k problematike, a zostaviť systém snímajúci pohyb, ktorý by využíval dostupné laboratórne vybavenie.

K riešeniu detekcie pohybu v obraze sme zvolili spôsob optického snímania markerov, ktoré reprezentujú pohyb, pomocou dvoch kamier. Optický systém je výhodný hlavne v tom, že je nenáročný na laboratórne vybavenie a metódy výpočtu 3D pozície markeru využívajúce dva pohľady sú veľmi presné. Na druhú stranu softvér potrebný k spracovaniu dát je implementačne náročný a musí sa dokázať vyrovnáť s prekrytiami markerov a musí zachovávať identitu markerov medzi snímkami v čase.

Navrhnutý systém opísaný v druhej kapitole dokáže na základe kalibračných snímok samostatne zistiť vnútorné parametre kamier a následne vypočítať aj ich pozíciu a natočenie v priestore. V zosynchronizovaných zdrojových videách sú markery detegované na základe ich farby. Tento spôsob sa ukázal byť veľmi účinný, hoci dochádza k nejasnostiam a skresleným dátam vplyvom rozmazania rýchlym pohybom.

Hľadanie korešpondencií objektov medzi snímkami z dvoch kamier pomocou epipolárnej geometrie je najúčinnnejšie, keď sú viditeľné všetky obrazy markerov v oboch snímkach. Preto je vhodné, keď je na začiatku snímania pohybu herec v postoji, v ktorom sú viditeľné všetky markery v snímkach z oboch kamier. Po prvých pár snímkach je táto korešpondencia udržiavaná hlavne pomocou Kalmanovho filtru, ktorý už je inicializovaný.

V konečnom dôsledku implementovaný program dokáže zrekonštruovať jednoduché pohyby herca. Problémy so zachovaním trajektórií nastávajú pri zakrytí alebo zmiznutí viacerých markerov a ich neskoršom objavení na vzdialenejšom mieste. K ďalšiemu vývoju programu navrhujeme zaoberať sa nasledujúcimi bodmi:

- Modifikácia programu, aby dokázal spracovávať synchronizované videá nielen z dvoch, ale aj z viacerých kamier. Tým pádom by dochádzalo k menej často k prekrytiam a miznutiam markerov, keďže viacero kamier by dokázalo pokryť miestnosť z každého uhla.
- Definovať „kostru“ človeka – závislosti polôh a vzdialenosti markerov. Pracovala by ak kontrola identity markerov
- Podpora použitia špeciálne označených markerov, pomocou ktorých by sa dalo zistiť okrem ich pozície aj ich natočenie.
- Optimalizácia algoritmov, aby mohol program pre menší počet markerov spracovávať dáta v reálnom čase.

Po zdokonalení a ďalšom vývoji by program mohol nájsť uplatnenie nielen v mnohých oblastiach počítačovej grafiky a multimédií ako napríklad pri počítačovej animácii, vytváraní simulátoru ktorý vyžaduje realistické pohyby postavy, ale aj pri biomechanickom výskume, kontrole výrobného procesu, detekcií pohybových väd, alebo dokonca pri ovládaní robota.

Zaujímavé využitie by mohlo byť aj pri rozoznávaní druhu pohybu herca podľa jeho trajektórií.

Literatúra

- [1] Hartley, R.~I., Zisserman, A. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000, ISBN: 0521623049
- [2] Intel OpenCV documentation, dokument je dostupný na URL <http://opencvlibrary.sourceforge.net>
- [3] Zhang Z. A Flexible New Technique for Camera Calibration. Technical Report MSRTR-98-71, Microsoft Research, December 1998.
- [4] Frew, E., Bister, A., LeMaster, E. Image Thresholding for Object Detection. 1997, dokument je dostupný na URL <http://sun-valley.stanford.edu/projects/helicopters/final.html>
- [5] Birchfield, S., An Introduction to Projective Geometry, 1998, dokument je dostupný na URL <http://robotics.stanford.edu/~birch/projective/projective.pdf>
- [6] Henrichsen, A., 3D Reconstruction and Camera Calibration from 2D Images, University of Capetown, 2000, dokument je dostupný na URL <http://www.dip.ee.uct.ac.za/Publications/Theses/MScArne.pdf>

Zoznam príloh

Príloha 1. Slovník pojmov

Príloha 2. CD, obsahujúce program Mocap, zdrojové kódy, dokumentáciu k programu, výsledky a zdrojové dáta.

Príloha 1: Slovník pojmov

Motion Capture – zaznamenávanie pohybu

3D – trojrozmerný priestor v akom sa pohybuje človek

2D – dvojrozmerný priestor, rovina

marker – značka umiestnená na povrchu ľudského tela

objekt – dôležitá oblasť v snímku, v našom prípade obraz markeru

svetový súradnicový systém – súradnicový systém reálneho sveta, priestoru v ktorom sa herec pohybuje. Jeho počiatok je určený pri kalibrácii kamier.

korešpondencia objektov medzi snímkami – v priestore pozorujeme niekoľko markerov a každý z nich má svoj obraz (objekt) v snímkach zo všetkých kamier. Korešpondujúce objekty medzi snímkami sú obrazy toho istého markeru.

korešpondencia objektov v čase – korešpondujúce objekty v čase sú obrazy toho istého markeru v snímkach z tej istej kamery