

ONLINE 3D RECONSTRUCTION

Jiří Bastl

Master Degree Programme (2), FEEC BUT

E-mail: xbastl00@stud.feec.vutbr.cz

Supervised by: Petr Petyovský

E-mail: petyovsky@feec.vutbr.cz

Abstract: This article describes the implementation of 3D reconstruction, using passive triangulation. This means that the scene is scanned by two cameras and reconstruction is then calculated on the basis of the camera frame. The result is a 3D model of the captured scenes stored in the VRML format. The whole calculation model must be computed as fast as possible, because input is a live image stream from two cameras.

Keywords: 3D reconstruction, passive triangulation

1. ÚVOD

O problematice pasivní 3D rekonstrukce bylo napsáno již mnoho a srovnání některých algoritmů můžete nalézt například tady [1]. Většina zde uvedených algoritmů je ale výpočetně velmi náročná a pro zpracování živého videa by potřebovaly vysoký výpočetní výkon nebo výpočty paralelizovat. Tento článek se zaměřuje na omezení zpracovávaných dat a tím se pokouší zmenšit výpočetní náročnost.

2. ROZBOR

Velice zjednodušeně by se celý proces výpočtu dal shrnout do čtyř kroků:

1. Předzpracování snímků (filtrace, odstranění zkreslení, rektifikace)
2. Hledání významných bodů
3. Sesouhlasení bodů
4. Výpočet 3D

2.1. REKTIFIKACE SNÍMKŮ

Promítnutí bez zkreslení bodu z trojrozměrného prostoru do kamery popisuje rovnice (1):

$$s \cdot [u \quad v \quad 1]^T = \mathbf{W} \cdot [\mathbf{R} | \mathbf{t}] \cdot [X \quad Y \quad Z]^T, \quad (1)$$

kde \mathbf{W} je projekční matice a $[\mathbf{R} | \mathbf{t}]$ poloha kamery v prostoru (tj. transformace z obecných souřadnic do souřadnic kamery). Tato poloha je určena na základě snímku kalibračního objektu (šachovnice). Rektifikace se provádí za účelem zjednodušení problému hledání korespondencí v celém snímku na hledání v jednom řádku a srovnává snímky tak, že korespondující body mají stejnou souřadnici y . Transformaci si lze představit tak, že obrazové roviny obou snímků v prostoru vhodně natočíme a zpětně promítneme. Tento proces pro levý snímek popisuje následující rovnice (2):

$$s \cdot [u_{L,rect} \quad v_{L,rect} \quad 1]^T = \mathbf{W}_{new} \cdot \mathbf{R}_{L,rect} \cdot \mathbf{W}_L^{-1} \cdot [u_L \quad v_L \quad 1]^T, \quad (2)$$

kde \mathbf{W}_L je projekční matice, kterou se bod v prostoru promítl do snímku, $\mathbf{R}_{L,rect}$ je rotační matice a \mathbf{W}_{new} je projekční matice, kterou rotované body zpětně promítáme do roviny. Pro pravý snímek platí analogická rovnice.

2.2. HLEDÁNÍ VÝZNAMNÝCH BODŮ

Nalezením významných bodů a počítáním jen s jejich pozicemi se výrazně omezí množství informace, které je nutné zpracovat. Podmínkou pro kvalitní rekonstrukci je, aby v pravém i levém snímku bylo co možná nejvíce korespondujících bodů. To znamená, aby detektor významných bodů měl dobrou opakovatelnost. Samozřejmě je také nutné, aby celá detekce nebyla časově náročná. Byly testovány celkem tři detektory – Harrisův, Shi-Tomasi a FAST detektor. Jako nejlepší se ukázal být FAST detektor [2], který je byl desetkrát rychlejší než ostatní při srovnatelné, nebo lepší opakovatelnosti.

2.3. SESOUHLASENÍ BODŮ

Po nalezení významných bodů je nutné nalézt korespondence, které jsou určovány na základě normalizované korelace NCC mezi okolními jednotlivých významných bodů. Aby nebylo nutné porovnávat prvky každý s každým, jsou uvažovány následující omezující podmínky pro dvě paralelně umístěné kamery:

1. Rektifikace pozic významných bodů zaručí stejnou souřadnici y korespondujících bodů.
2. Korespondující bod v pravém obrázku bude ležet nalevo od jeho pozice v levém obrázku.
3. Pořadí korespondujících bodů v jednom řádku bude v levém i v pravém obrázku stejné.

První a druhá podmínka umožňuje korespondenci významného bodu v levém obrázku hledat pouze v úzkém pruhu v pravém obrázku. Třetí podmínka umožňuje odstranit špatně sesouhlasené body, protože ani vysoká korelace mezi okolními nezaručuje správnost korespondence. Princip odstranění chybně sesouhlasených bodů je ukázán na následujícím obrázku.

A	A	A	A	A	A	B	d	A	A	A	A	A	A	B	d
A	A	B	A	B	C	C	c	A	A	X	A	B	C	C	c
A	C	B	C	B	B	C	b	A	C	B	C	B	X	C	b
B	C	C	C	C	C	C	a	B	C	C	C	C	C	C	a
1	2	3	4	5	6	7		1	2	3	4	5	6	7	

Obrázek 1: Korelace bodů ležících v jednom řádku (vlevo).

Obrázek 1 vlevo znázorňuje tabulku korelací bodů v jednom řádku. Čísla řádků 1-7 odpovídají indexům významných bodů v levém snímku a písmena a-d indexům v pravém snímku. Značka A například u bodů [1, b] značí, že bod v pravém snímku ležel vpravo od bodu v levém snímku a korelace nebyla vůbec počítána. Pro dvojici bodů označenou písmenem C byla korelace spočítána, ale nepřekročila daný práh (okolí si nebyla podobná). Korelace mezi body označených písmenem B překročila práh, přičemž u bodů označených plnou značkou byla větší, než u bodů s prázdnou značkou. Kdyby byly za korespondence označeny body s největší korelací, nebyla by v tomto případě splněna podmínka stejného pořadí. Implementovaný algoritmus kontroluje body B s plnou značkou s body okolními, a pokud je pořadí špatné, hledá jiný přípustný bod (dvojice [3, c]). Když žádný přípustný není, korespondenci vyloučí (dvojice [6, b]).

2.4. VÝPOČET 3D

Z rovnic (1) a (2) lze odvodit, kam se promítne 3D bod $\mathbf{B}[X, Y, Z]$ do rektifikovaného snímku:

$$s \begin{bmatrix} x_{rect} \\ y_{rect} \\ 1 \end{bmatrix} = \mathbf{W}_{new} \cdot \mathbf{R}_{rect} \cdot [\mathbf{R} | \mathbf{t}] \cdot [\mathbf{B} | 1]^T = \mathbf{P}[\mathbf{B} | 1]^T = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \cdot [\mathbf{B} | 1]^T, \quad (3)$$

Pro pravý a levý snímek tak dostaneme dvě maticové rovnice s neznámými souřadnicemi X, Y, Z a neznámými měřítky s_L a s_R . Máme tedy pět neznámých a soustavu šesti rovnic (každý řádek představuje jednu rovnici o pěti neznámých), kterou lze řešit například SVD rozkladem.

3. DOSAŽENÉ VÝSLEDKY

Algoritmus byl testován na rovinné desce, na které byly vyznačené orientační body. Rovinná deska byla zvolena z důvodu snadné kontroly nalezených 3D pozic. V tabulce 1 jsou shrnuty informace o měření. Jako chybný bod byl vyhodnocen každý, který byl od roviny vzdálen více jak 20 mm. Statistické údaje jsou potom počítány z naměřených hodnot bez chybných bodů.

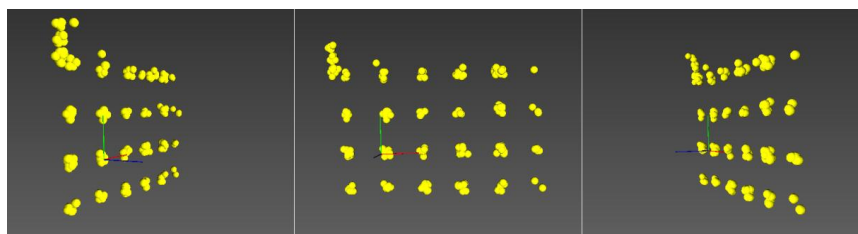


Obrázek 2: Rektifikované snímky s nalezenými korespondencemi

Počet korespondencí:	198	Průměrná chyba [mm]:	-3,7
Chybných bodů:	8	Rozptyl [mm ²]:	76,7

Tabulka 1: Výsledky měření

Průměrný výpočetní čas byl 64 ms na 2.0 GHz procesoru, při rozlišení videa 640x480 bodů a je potřeba zmínit, že výpočetní náročnost je závislá na počtu bodů. Na následujícím obrázku 3 je rozfázován pohled z několika úhlů na výsledný VRML model.



Obrázek 3: Model VRML

4. ZÁVĚR

Představený algoritmus dokáže být velmi rychlý a je dobře použitelný na scény s bohatou texturou. Současný stav neumožňuje pracovat se subpixelovou přesností, ale i přesto se dosažená přesnost blíží teoreticky předpokládané.

REFERENCE

- [1] SCHARSTEIN, D. - BLASIAK, A. The Middlebury Stereo Vision Page [online]. Verze 2. 2009 [cit. 2011-03-04]. Middlebury Stereo Evaluation. Dostupné z WWW: <<http://vision.middlebury.edu/stereo/eval/>>.
- [2] ROSTEN, E. - DRUMMOND, T. *Machine Learning for High-Speed Corner Detection*. Department of Engineering, Cambridge University, UK.