Computer Vision 3D vidění

Zdenek Krnoul, Miroslav Jirik

Department of Cybernetics Faculty of Applied Sciences University of West Bohemia

ESF projekt Západočeské univerzity v Plzni reg. č. CZ.02.2.69/0.0/0.0/16_015/0002287







Obsah

Úvod do 3D vidění (Zdeněk Krňoul)

- Typy úloh
- Teorie 3D vidění
- Vstupní data a principy snímání
- Reprezentace 3D dat (Miroslav Jiřík)
 - Současné techniky





3D vidění - základní typy úloh:

► 3D Rekonstrukce

- Nalezení geometrických a fyzikálních parametrů objektů ve 3D scéně
- Možno rekonstruovat jejich 3D tvar (3D skenery, 3D tvář identifikace apod.)
- Rozpoznávání/odhad ve 3D
 - Nalezení a klasifikace objektů ve scéně (navigace auta, robota aj.)
 - Algoritmy slouží pro určení polohy, orientace, měřítka apod.

Pozn. V praktických úlohách se často snažíme vyhnout 3D vidění, převést úlohu na 2D. Důvodem jsou požadavky na robustnost, nízkou cenu a velkou efektivitu využití systémů počítačového vidění.







Teorie 3D vidění (Marrova teorie, r. 1982)

- 4 úrovně reprezentace 3D scény:
 - ▶ intenzitní obraz
 - prvotní náčrtek
 - 2 1/2 dimenzionální náčrtek
 - ▶ plná 3D reprezentace









Prvotní náčrtek:



- Obsahuje informace o velikostech a směrech významných jasových změn v obraze
- Jejich vzájemné geometrickém uspořádání
- Předpokládáme, že takto získané čáry a skvrny zachovávají informaci potřebnou pro pozdější 3D reprezentaci







2 1/2 rozměrný náčrtek:



Pro jeho získání se používá informace obsažená v prvotním náčrtku

Používají se různé techniky, souhrnně nazývané "tvar z X"

- Tvar ze stereovidění
- Tvar z pohybu (optický tok)
- Tvar z jasu
- Tvar z textury





Plná 3D reprezentace objektu:



- Opírá se o geometrické vlastnosti, které lze v obraze nalézt
- Geometrické vlastnosti jsou vyjádřené vzhledem k souřadnému systému vycházejícímu z tvaru objektu
- Základní geometrické vlastnosti:
 - střed (nejčastěji těžiště)
 - celková velikost
 - zobecněná osa symetrie (pokud je)







Vstupní data, možnosti snímání

- intenzitní obraz(y) (RGB) (často je potřeba více pohledů na scénu z různých úhlů)
- stereovidění právě dva intenzitní obrazy, kalibrované kamery
- hloubkové senzory jsou v dnešní době běžný postup







Hloubková mapa

- 1. Často používáme specializovaný senzor či kameru např. dobře známé MS Kinect, Intel RealSense aj.
- 2. Nebo technikou zaostřování objektivu
- Nebo výpočtem pouze z intenzitních obrazů, např. zmíněné stereovidění - tzv. husté stereo
- 4. A další techniky založené např. na odrazu světla, známé textuře aj.



Obrázek: Blickfeld, KITTI dataset







Hloubkové senzory:

- Strukturované osvětlení scéna osvětlena pomocný zdrojem světla
- Laserové podobný princip jako radary a sonary



Obrázek: Společný princip zařízeních při získávání hloubkové mapy.







Princip strukturovaného osvětlení:

- Scéna/povrch osvětlen např. úzkým proužkem světla nebo jedinečným vzorem (MS Kinect I - infrared) a scéna snímána z jiného úhlu, (problém stínů)
- Nutná kalibrace projektoru a kamery (často zabudováno pevně v přístroji)
- Nutný algoritmus pro dekódování promítnutého vzoru
- Dále už jen princip stereovidění, viz níže
- Dobrá přesnost je získána pro jednobarevné nelesklé povrchy, u Kinectu ruší jiné IR záření (např. slunce)



Strukturovaného osvětlení: Moiré proužky:

- Obvykle promítnuty tři nebo čtyři fázově posunuté sinusové vzory na povrch
- Snímky deformovaných obrazců zachycené kamerou se používají k výpočtu fázové mapy, která obsahuje informace o výšce povrchu
- Vysoká přesnost je získána pro jednobarevné nelesklé povrchy



Pozn. Na podobném principu existuje mnoho dalších metod







Princip laserového měření:

- Měří fázový posun mezi vyslaným a přijatým signálem (např. bodově
 - lidar, nebo jako obraz ToF kamera MS Kinect II)



Obrázek: Převzato z [KOT+17]





Princip proměnného zaostření objektivu:



All-III-locus

Obrázek: Pro mobilní telefon, převzato z [SHS15]







Princip stereovidění:

- Využívá obecný model perspektivní kamery popis projekce 3D prostoru do 2D prostoru (obrazová rovina)
- Vždy se jedná o středovou projekci



Pozn. speciální případ ... střed projekce leží v nekonečnu \rightarrow afinnícnění tzv. paralelní projekce









- Perspektivni kamera má 11 stupňů volnosti: 1x ohnisková vzdálenost v pixelech + 1x poměr stran pixelu + 1x zkoseni os + 2x počátek obrázku + 3x posun + 3x rotace kamery = 11 DOFs (degree of freedom)
- Pro určení použijeme algoritmus kalibrace kamery (problém je lineární v homogenních souřadnicích)
- Teoreticky stačí 5,5 prostorových bodů







- Někdy je nutné provést korekci zkreslení souřadnic obrázku tzv. radialní zkreslení
- Efekt je způsoben nedokonalostí středového promítání, které je fyzicky provedeno přes čočku(y) objektivu





- U stereovidění máme dva pohledy na stejnou scénu z různých směrů
 - Pohledy mohou být získány souběžně v jeden okamžik (např. dvě kamery) (Stereovision, Binocular disparity)
 - Pohyb přístroje před objektem (např. pohyb robota) (Egomotion Estimation)
 - Pohyb objektu před přístrojem (např. otáčím jablkem před kamerou) (Structure from Motion - SFM)
- Tyto úlohy jsou duální a vedou na stejné řešení







Stereovidění - 3D rekonstrukce



- Pokud známe obě kamery zkalibrované
- A současně máme korespondující pár
- Pak můžeme určit 3D souřadnici neznámého bodu





3D tvar z textury:

- Předpoklad pravidelného vzoru
- ► 3D tvar z předpokládané projekce







3D tvar ze stínů (Shape from Shading - SFS):

- Předpoklad známé polohy zdroje světla
- Lambertovský povrch (odráží světlo do všech směrů stejně)



Obrázek: převzato z [PF05]





Lze rozdělit na:

- deskriptivní plně popisují tvar objektu, definována odrazivost a osvětlení, z takového modelu lze vytvořit syntetický intenzitní obraz i syntetickou hloubkovou mapu pro libovolné místo pozorování
- diskriminační slouží k odlišení objektů několika tříd

Používané typy reprezentace:

 drátový model - graf, jehož vrcholy odpovídají 3D bodům (často vrcholům objektu), hrany odpovídají hranicím (nespojitostem normál k povrchu). Nehodí se pro popis objektů s křivočarými povrchy







- CSG (constructive solid geometry) model používá jako základ množinu jednoduchých 3D objektů, jako hranoly, kužely, válce, koule, krychle, kvádry ap. a kombinuje je v určité pozici, zvětšení a orientaci pomocí jednoduchých množinových operací, jako průnik, sjednocení, rozdíl, ap.
- Model je reprezentován stromem; listy odpovídají jednotlivým elementárním tělesům, vyšší uzly představují množinové operace. Toto vyjádření je výpočetně velmi náročné.









GEON (z angl. geometrical ions)

- Zdůrazňuje kvalitativní charakter reprezentace objektů
- 3D objekty jsou složeny z několika sousedících GEONů
- Navrženo 36 základních GEONů
- Každý GEON je charakterizován čtveřicí kvalitativních vlastností:
 - hranice rovná / křivá
 - symetrie středová / osová / žádná
 - změna velikosti stálá / zvětšující se / zvětšující i zmenšující se
 - osa přímá / křivá









VOXEL (z angl. volumetrix element)



Obrázek: http://www.bilderzucht.de/blog/3d-pixel-voxel/







NURBS

Non-uniform rational B-spline [Pie91]







Geometry and Topology

Geometry

Geometry is a branch of mathematics concerned with questions of shape, size, relative position of figures, and the properties of space [DR15]. A mathematician who works in the field of geometry is called a geometer.

Topology

In mathematics, topology is concerned with the properties of a geometric object that are preserved under continuous deformations, such as stretching, twisting, crumpling and bending, but not tearing or gluing.







Obrázek: Seven Bridges of Königsberg [Eul53, Eul41, Giu20]



Content of example.obj can be displayed with ParaView application.

v	0.0	0.0	0.1
v	0.0	0.1	1.0
v	0.0	1.0	0.1
v	0.0	1.0	1.0
v	0.2	2.0	0.5

g cellobject f 1 2 4 3 f 3 4 5









Linear Algebraic Representation

Alberto Paoluzzi, Antonio DiCarlo [DPS14]







A complex C is a sequence
$$\cdots \longrightarrow C_{d+1} \xrightarrow{\partial_{d+1}} C_d \xrightarrow{\partial_d} C_{d-1} \longrightarrow \cdots$$

Chain and cochain complex

A chain complex C is a complex of chain spaces and boundary maps:



I-cell subsets), are the standard bases (Md rows) OFPORTMENTOF





✻

CYBERNETICS











Computer Vision

30/66

Point cloud



Triangulation (terrain map)









We first determine a triangulation of P: a planar subdivision whose bounded faces are triangles and whose vertices are the points of P. We then lift each sample point to its height, mapping every triangle in the triangulation to a triangle in 3-space. We get is a polyhedral terrain, the graph of a piecewise linear continuous function The polyhedral terrain as an approximation of the original terrain.







Let $P := \{p_1, p_2, ..., p_n\}$ be a set of points in the plane. To be able to formally define a triangulation of P, we first define a maximal planar subdivision as a subdivision S such that no edge connecting two vertices can be added to S without destroying its planarity. In other words, any edge that is not in S intersects one of the existing edges. A triangulation of P is now defined as a maximal planar subdivision whose vertex set is P.

Let P be a set of n points in the plane, not all collinear, and let k denote the number of points in P that lie on the boundary of the convex hull of P. Then any triangulation of P has 2n - 2 - k triangles and 3n - 3 - kedges.





pské strukturální a investiční fonc

Delaunay triangulation for a set P of discrete points in a plane is a triangulation DT(P) such that no point in P is inside the circumcircle of any triangle in DT(P). Delaunay triangulations maximize the minimum angle of all the angles of the triangles in the triangulation.









Delaunay Triangulation Properties

- The Delaunay triangulation is a triangulation which is equivalent to the nerve of the cells in a Voronoi diagram,
- it is the triangulation of the convex hull of the points in the diagram in which every circumcircle of a triangle is an empty circle an edge is illegal if we can locally increase the smallest angle by flipping that edge.
- A Delaunay triangulation is unique iff the circumcircle of every triangle contains exactly three points on its circumference: the vertices of the triangle.
- For instance, the Delaunay diagram of the four vertices of a square is a square, and can be converted into a triangulation in two different ways





Computing of The Delaunay Triangulation

It turns out that it is not necessary to compute the angles to check whether a given edge is legal. Instead, we can use the simple criterion stated in the next lemma. The correctness of this criterion follows from Thales's Theorem

Let edge $\overline{p_i p_j}$ be incident to triangles $p_i p_j p_k$ and $p_i p_j p_l$, and let *C* be the circle through p_i , p_j , and p_k . The edge $\overline{p_i p_j}$ is illegal if and only if the point p_l lies in the interior of *C*. Furthermore, if the points p_i , p_j , p_k , p_l form a convex quadrilateral and do not lie on a common circle, then exactly one of $\overline{p_i p_j}$ and $\overline{p_k p_l}$ is an illegal edge.









Computing the Delaunay Triangulation

We define a legal triangulation to be a triangulation that does not contain any illegal edge. From the observation above it follows that any angle-optimal triangulation is legal. Computing a legal triangulation is quite simple, once we are given an initial triangulation. We simply flip illegal edges until all edges are legal.

```
Input: Some triangulation T of a point set P.

Output: A legal triangulation of P.

while T contains an illegal edge \overline{p_i p_j} do

begin Flip \overline{p_i p_j}

Let p_i p_j p_k and p_i p_j p_l be the two triangles adjacent to \overline{p_i p_j};

Remove \overline{p_i p_j} from T, and add \overline{p_k p_l} instead.

end

end

return T
```





Voronoi complexes







Computer Vision

39 / 66

Voronoi complexes

- Voronoi diagram is a partitioning of a plane into regions based on distance to points in a specific subset of the plane.
- That set of points (called seeds, sites, or generators) is specified beforehand, and for each seed there is a corresponding region consisting of all points closer to that seed than to any other. These regions are called Voronoi cells
- The Voronoi diagram of a set of points is dual to its Delaunay triangulation





strukturální a investiční fon



The post office problem



Obrázek: The post office problem [dBVKO⁺08]







Convex Hull







Given a discrete set S of points:

- intersection of all convex sets containing S;
- minimum convex set containing S
- set spanned by all convex combinations of S points









Jarvis algoritgm[Jar73]

Gift Wrap Algorithm (Jarvis March Algorithm) to find Convex Hull. O(nh) complexity, where n = #S, and h is the number of points on the convex hull. Possition is checked by cross product. Demo on Wikipedia



Chan's algorithm in the planar case: the algorithm combines an $O(n \cdot \log(n))$ algorithm (Graham scan-line, for example) with Jarvis march O(nh), in order to obtain an optimal $O(n \cdot \log(h))$ time. Demo on Wikipedia









Obrázek: Study the shape of a set of points[Ede14]





Alpha-complexes [Ede10, EKS83]









Voronoi decomposition

According to Delaunay triangulation, there is an edge between two points if their regions intersect in a common edge, and a triangle between three points if their regions intersect in a common point.





VROPSKÁ LINIE ské strukturální a investiční fond rační program Výzkum, vývoj a vzděláván

α -Complex



Obrázek: $\alpha\text{-Complexes}$ for varying α





α -shapes in 3D









α -shapes in 3D









Spatial Hierarchical Domain Trees

Foundations of Multidimensional and Metric Data Structures provides a thorough treatment of multidimensional point data, object and image-based representations, intervals and small rectangles, and high-dimensional datasets Hanan Samet is the inventor of such trees (quadtrees and octrees), which are the foundation of Geographical databases [Sam06]



EVROPSKÁ UNIE Evropské strukturální a investiční fondy Operační program Výzkum, vývoj a vzděláván 2^{n} -trees are ordered trees characterized by the property that each non-leaf node has exactly 2^{n} son nodes, respectively denoted as first, second, etc., and as 2^{n} -th son When n = 2 and n = 3 such trees are called quadtrees and octrees, and are used to represent hierarchical decompositions of the 2D or 3D space, respectively







Quadtrees

- a quadtree is a quaternary tree (i.e. each non-leaf node has exactly 4 sons);
- ▶ the leafs are either white or black nodes (i.e. either empty or full);
- ▶ the non-leafs are gray nodes (i.e. neither empty nor full);
- ▶ the maximal depth of the quadtree is related to its resolution.
- The number of arcs on the path from the root to a node is called distance of the node from the root.
- ▶ Depth of a tree is the maximal distance of its nodes from the root.
- The resolution of the quadtree with squared bounding box of size L and depth m is clearly equal to

$$r = L/2^{m}$$





Quadtree encoding

hierarchical decompositive representation using a quadtree, and its actual encoding as a labeled tree, where black, white and gray nodes represent full and empty cells, and cells which are neither full nor empty. The sons of a gray node are clockwise ordered.



Obrázek: Quadtree encoding scheme: (a) 2D object (b) full cells (black), empty cells (white) and decomposed cells (gray)





Octrees

An octree is a tree data structure in which each internal node has exactly eight children. Octrees are most often used to partition a three-dimensional space by recursively subdividing it into eight octants.



Obrázek: Octree: partition of a 3D cell into 8 sub-cells generated by three orthogonal planes







Marching Cubes



Obrázek: Find the 0-surface (or any iso-surface) of a discrete 3D field [LC87]





Lar-surf



Obrázek: LAR surface extraction scheme





Surface extraction with LAR



Obrázek: Liver microstructure[PDFJ16]







Surface extraction with LAR









References I

- T. M. Chan.

Optimal output-sensitive convex hull algorithms in two and three dimensions.

Discrete and Computational Geometry, 1996.

- Mark de Berg, Marc Van Kreveld, Mark Overmars, O. Schwarzkopf, Otfried Cheong, Marc Van Kreveld, and Mark Overmars. *Computational geometry: Algorithms and applications*, volume 85. 2008.

Antonio DiCarlo, Alberto Paoluzzi, and Vadim Shapiro. Linear algebraic representation for topological structures. *Computer-Aided Design*, 46:269–274, 1 2014.







References II

Vincenzo De Risi. Mathematizing Space. The Objects of Geometry from Antiquity to the Early Modern Age. Birkhäuser, 2015.

Herbert Edelsbrunner. Alpha Shapes - a Survey. Tessellations in the Sciences, pages 1–25, 2010.

Herbert Edelsbrunner.

A short course in computational geometry and topology. Number Mathematical methods. Springer, 2014.

Herbert Edelsbrunner, David Kirkpatrick, and Raimund Seidel. On the shape of a set of points in the plane. *IEEE Transactions on information theory*, 29(4):551–559, 1983.







References III

Leonhard Euler. Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae*, pages 128–140, 1741.



Leonhard Euler.

Leonhard Euler and the Königsberg bridges. *Scientific American*, 189(1):66–72, 1953.



Bogdan Giușcă.

Map of Königsberg in Euler's time showing the actual layout of the seven bridges, highlighting the river Pregel and the bridges, 2020.







References IV

R. A. Jarvis.

On the identification of the convex hull of a finite set of points in the plane.

Information Processing Letters, 1973.



IPSJ Transactions on Computer Vision and Applications, 9, 12 2017.

William E Lorensen and Harvey E Cline. Marching Cubes: A High Resolution 3D Surface Construction Algorithm.

ACM siggraph computer graphics, 21(4):163–169, 1987.







References V

Alberto Paoluzzi, Antonio Dicarlo, Francesco Furiani, and Miroslav lirik

CAD models from medical images using LAR. Computer-Aided Design, 13(6), 2016.

Emmanuel Prados and Olivier Faugeras. Shape from shading. In in Mathematical Models in Computer Vision: The Handbook. Springer, 2005.



L Piegl.

On NURBS: a survey.

IEEE Computer Graphics and Applications, 11(1):55–71, 1991.







References VI

Hanan Samet.

Foundations of multidimensional and metric data structures. Morgan Kaufmann, 2006.

 Supasorn Suwajanakorn, Carlos Hernandez, and Steven M. Seitz. Depth from focus with your mobile phone.
 In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2015.





