

Digitální zpracování obrazu × počítačové vidění

Zakotvení z pohledu digitální fotografie

Václav Hlaváč

České vysoké učení technické v Praze

Český institut informatiky, robotiky a kybernetiky

160 00 Praha 6, Jugoslávských partyzánů 1580/3

<http://people.ciirc.cvut.cz/hlavac>, vaclav.hlavac@cvut.cz

také z Centra strojového vnímání, <http://cmp.felk.cvut.cz>

Osnova přednášky:

- ◆ Digitální zpracování obrazů × analýza obrazů × počítačové vidění.
- ◆ Vidění × vnímání.
- ◆ Proč je vidění těžké?
- ◆ Interpretace, význam pro obrazy.

Co je počítačové vidění?

Počítačové vidění je jak vědou tak i technologií usilující o vytváření “strojů schopných vidět a vnímat”.

◆ **Vědecký obor:**

hledá teorie pro vytváření umělých systémů získávajících informace z obrazů.

◆ **Technický obor:**

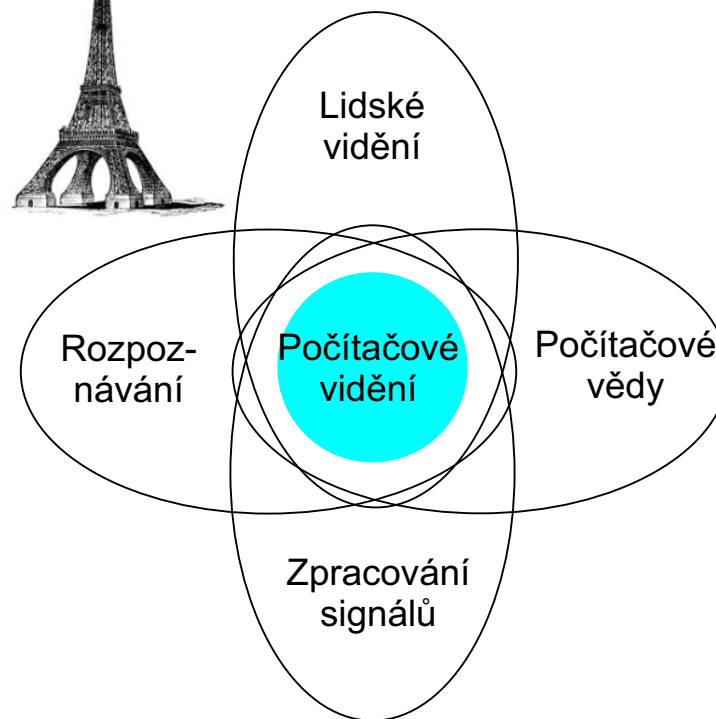
Počítačové vidění = kamera + počítač + ?

◆ **Obrazy (například):**

- pohledy z více kamer,
- video sekvence,
- vícerozměrná data z lékařského skeneru, např. tomografu.

Proč studujeme zpracování obrazu, analýzu obrazu a počítačové vidění?

- ◆ Počítačové vidění vyrostlo přinejmenším na čtyřech pilířích: (1) vědě o počítačích; (2) teorii signálů; (3) rozpoznávání; (4) porozumění lidskému vidění.
- ◆ Zkoumání započala v 60. letech 20. století.
- ◆ Má bohatou metodologii.
- ◆ Poskytuje a využívá zajímavé mezioborové vazby.
- ◆ Poskytuje vhled do lidského vidění.
- ◆ Je důležitým zdrojem informace v moderní informační době.



Na co se používá počítačové vidění?

- ◆ Pro detekci, segmentaci, popisování, rozpoznávání (vzletněji: porozumění, vnímání) objektů, které nás zajímají ve 2D či 3D obrazech nebo sekvencích.
- ◆ Pro detekci událostí (např. při sledování bezpečnostními kamerami, počítání lidí, při detekci startující balistické rakety ze senzorů na družici).
- ◆ Pro uspořádání informace na základě obrazů (např. pro indexování obrazových databází nebo video sekvencí).
- ◆ Jako součást řídicích systému (např. u průmyslových robotů nebo autonomně jedoucích aut).
- ◆ Pro modelování objektů nebo okolního světa (např. při obrazové kontrole kvality výrobku v průmyslu, při analýze lékařských obrazů, při získávání 3D modelu ze série 2D obrazů).
- ◆ Pro interakci mezi člověkem a strojem (např. ovládání počítačové hry pomocí hloubkoměru Kinect).
- ◆ atd.

Vnímání

- ◆ Postupy k upoutání pozornosti a/nebo porozumění vjemům ze senzorů.
- ◆ Úloha je mnohem složitější, než si vědci byli schopni představit okolo let 1950 a 1960:
 - Tehdy: “Vytvoření strojů vnímat potvrvá zhruba jedno desetiletí.”
 - Přitom jsme od tohoto cíle stále velmi vzdáleni.
- ◆ Pět Aristotelových smyslů: zrak, slyšení, hmat, čich a chuť.
- ◆ Vnímání předpokládá dynamický vztah mezi: “reprezentaci světa v mozku” na základě ↔ smyslů, ↔ bezprostředního okolního světa, ↔ paměti.



Poděkování za obrázek:
<http://www.richardsonthebrain.com/>

Poznámky k lidskému (vizuálnímu) vnímání



6/12

Co vidíte na obrázku?



Poznámky k lidskému (vizuálnímu) vnímání

Co vidíte na obrázku?

Vidění a vnímání je pro člověka
a mnoho zvířat snadné a přirozené.



Poznámky k lidskému (vizuálnímu) vnímání

Co vidíte na obrázku?

Vidění a vnímání je pro člověka
a mnoho zvířat snadné a přirozené.

◆ Není to zadarmo:



Poznámky k lidskému (vizuálnímu) vnímání

Co vidíte na obrázku?

Vidění a vnímání je pro člověka a mnoho zvířat snadné a přirozené.

- ◆ Není to zadarmo:
 - Asi 50 % šedé mozkové kůry u primátů se věnuje zpracování vizuální informace (Felleman-van Essen 1991).



Poznámky k lidskému (vizuálnímu) vnímání

Co vidíte na obrázku?

Vidění a vnímání je pro člověka a mnoho zvířat snadné a přirozené.

◆ Není to zadarmo:

- Asi 50 % šedé mozkové kůry u primátů se věnuje zpracování vizuální informace (Felleman-van Essen 1991).
- Lidský mozek spotřebovává asi 20 % veškeré energie v lidském těle.



Poznámky k lidskému (vizuálnímu) vnímání

Co vidíte na obrázku?

Vidění a vnímání je pro člověka a mnoho zvířat snadné a přirozené.

◆ Není to zadarmo:

- Asi 50 % šedé mozkové kůry u primátů se věnuje zpracování vizuální informace (Felleman-van Essen 1991).

- Lidský mozek spotřebovává asi 20 % veškeré energie v lidském těle.

◆ Aby počítač viděl a vnímal jako člověk, museli bychom vyřešit většinu úkolů umělé inteligence (což je velmi těžké, blízké k nemožnému).



Poznámky k lidskému (vizuálnímu) vnímání

Co vidíte na obrázku?

Vidění a vnímání je pro člověka a mnoho zvířat snadné a přirozené.

◆ Není to zadarmo:

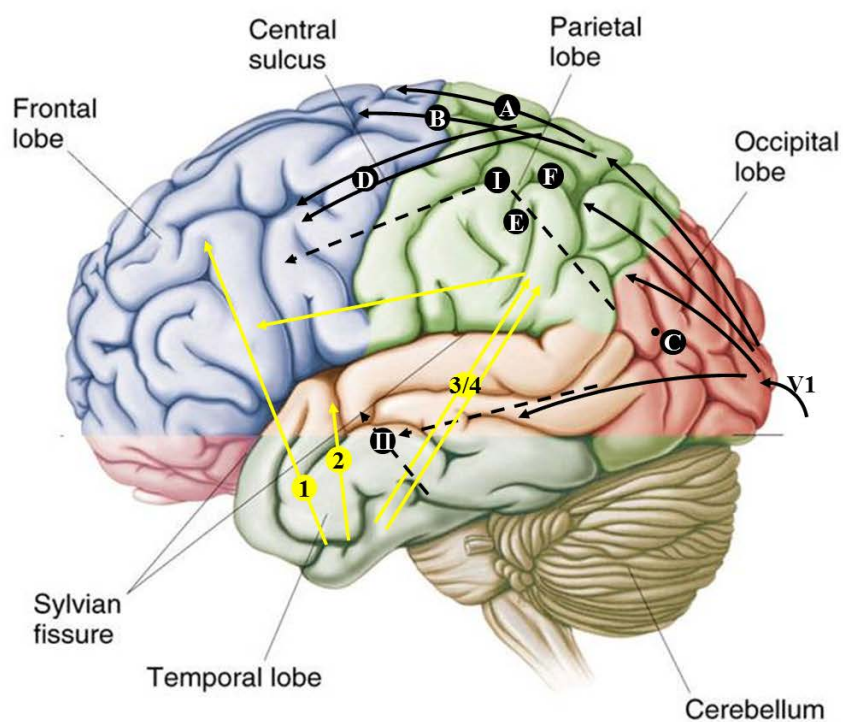
- Asi 50 % šedé mozkové kůry u primátů se věnuje zpracování vizuální informace (Felleman-van Essen 1991).
- Lidský mozek spotřebovává asi 20 % veškeré energie v lidském těle.

- ◆ Aby počítač viděl a vnímal jako člověk, museli bychom vyřešit většinu úkolů umělé inteligence (což je velmi těžké, blízké k nemožnému).
- ◆ Využívá se znalost vyšší úrovně, sémantická informace a kontext.



Lidské vidění

- ♦ Část šedé kůry mozkové věnující se vidění zaobírá 50 % mozku makaka.
- ♦ U člověka se větší část mozku věnuje vidění než jiným úkolům.



- A** Navigation, direction, obstacles
- B** Saccade control
- C** Pursuit control
- D** Understanding, focus, manipulation control
- E** Drawing, writing
- F** Calculation
- 1** Categorization
- 2** Recognition, objects in context
- 3** Calculation
- 4** Reading
- I** Recognition of actions
- II** Emotions

Lidské vidění na rozdíl od počítačového vidění

Vidění dovoluje člověku i zvířeti vnímat a porozumět světu, který je obklopuje.

Kognitivní vědy zkoumají také vidění v biologických systémech:

- ◆ Hledají empirické modely popisující biologické vidění.
- ◆ Někdy popisují vidění, jako by šlo o výpočetní systém.

Počítačové vidění usiluje o technická řešení, i když se někdy inspiruje v biologickém vidění:

- ◆ Biologické vidění zvládá úlohy, na něž je počítačové vidění stále krátké. Přesto poskytuje biologické vidění inspiraci i pro technická řešení.
- ◆ Technické požadavky na systémy počítačového vidění se často shodují s požadavky na biologické vidění.

Varování: Napodobování biologického vidění nemusí být nejlepším příkladem řešení technické úlohy.

Proč je počítačové vidění těžké ?
Najdeme alespoň 6 příčin.



Proč je počítačové vidění těžké ?

Najdeme alespoň 6 příčin.

3D → 2D přináší ztrátu informace díky vlastnostem perspektivní transformace (matematická abstrakce, dírková komora).

Měřený jas je dán složitým fyzikálním postupem vytváření obrazu. Zář (angl. radiance) (\approx jas) závisí na typu světelných zdrojů, jejich poloze, intenzitě, poloze pozorovatele, lokální geometrii povrchu a odrazivosti povrchu. Obrácená úloha je špatně podmíněna.

Nevyhnutelná přítomnost šumu v každém měření ve skutečném světě.

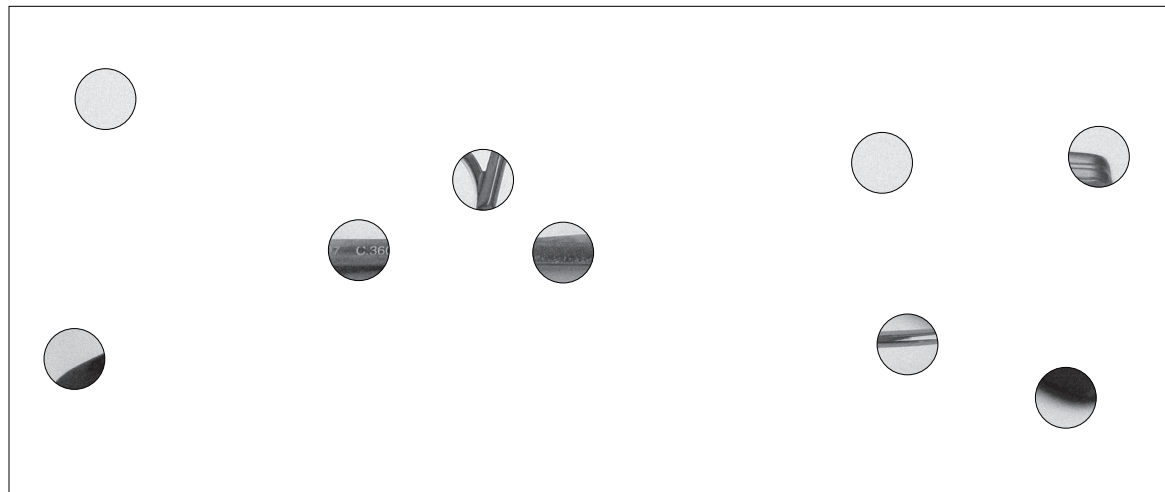
Příliš mnoho dat Stránka A4, 300 dpi, 8 bitů na pixel = 8,5 Mbajtů.

Neprokládané 4k video 3840×2160 , RGB (24 bitů) \approx 9.96 Gbitů/sekundu.

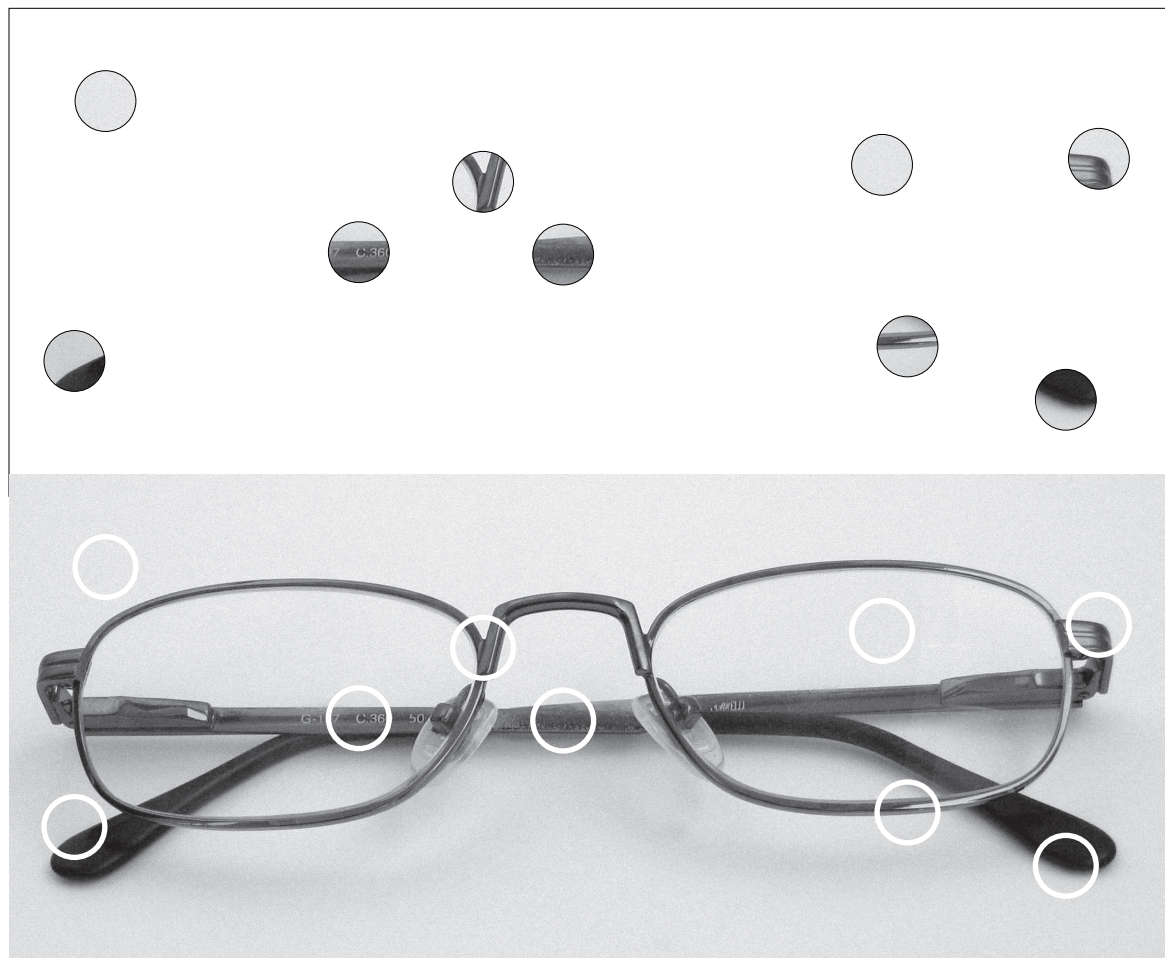
Nutnost zahrnout interpretaci (bude brzy diskutováno).

Lokální okno v kontrastu s potřebou globálního pohledu.

Ilustrace nedostatečnosti lokálního pohledu



Ilustrace nedostatečnosti lokálního pohledu



Interpretace a její role, sémantika

Jednou z lidských schopností je zasazovat vjemy do rámce (kontextu), který je konkrétnímu člověku známý. Tuto schopnost formalizujeme pojmem **interpretace**, který matematicky vyjádříme jako zobrazení.

Interpretace: pozorování \rightarrow logický model světa *nebo jinak* syntax \rightarrow sémantika

Příklady:

- ◆ Pohled z okna \rightarrow {prší, neprší}.
- ◆ Jablko na běžícím pásu \rightarrow {třída 1, třída 2, třída 3}.
- ◆ Dopravní scéna \rightarrow vyhledávání čísla auta.

Opora v teorii: matematická logika, teorie formálních jazyků.

Hluboká teoretická potíž: Gödelovy věty o neúplnosti (\approx 1930); Program Davida Hilberta usilující o nalezení úplné a konzistentní množiny axiomů pro celou matematiku je neuskutečnitelný. Přesněji řečeno, neexistuje soubor konzistentních axiomů, jehož tvrzení by šla sestavit algoritmem a uměla by dokázat všechna pravdivá tvrzení o aritmetice přirozených čísel (Peanova aritmetika).

Od nízké k vyšší úrovni zpracování z hlediska využívané apriorní znalosti

Nízká (až žádná) znalost \approx digitální zpracování obrazu

- ◆ Obrazy se neinterpretují, a proto jsou postupy nezávislé na konkrétní aplikační oblasti.
- ◆ Používají se metody zpracování signálů, např. 2D Fourierova transformace.

Střední znalost \approx analýza obrazu

- ◆ Často jen 2D obrazy, např. obrazy buněk v optickém mikroskopu.
- ◆ Interpretace přináší důležitou dodatečnou znalost umožňující řešit i úlohy, které by jinak řešit nešly.

Vyšší znalost \approx počítačové vidění, porozumění obsahu 3D scény z obrazů a videí

- ◆ Nejobecnější formulace úloh, 3D svět, měnící se scéna (videosekvence).
- ◆ Složité, využívá se interpretace, zpětné vazby a techniky umělé inteligence.
- ◆ Příliš ambiciózní cíle. Často špatně podmíněné a příliš těžké úlohy. Obvykle se musí radikálně zjednodušit.