

Digitální zpracování obrazu × počítačové vidění zakotvení

Václav Hlaváč

České vysoké učení technické v Praze
Fakulta elektrotechnická, katedra kybernetiky
Centrum strojového vnímání

<http://cmp.felk.cvut.cz/~hlavac>, hlavac@fel.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:**

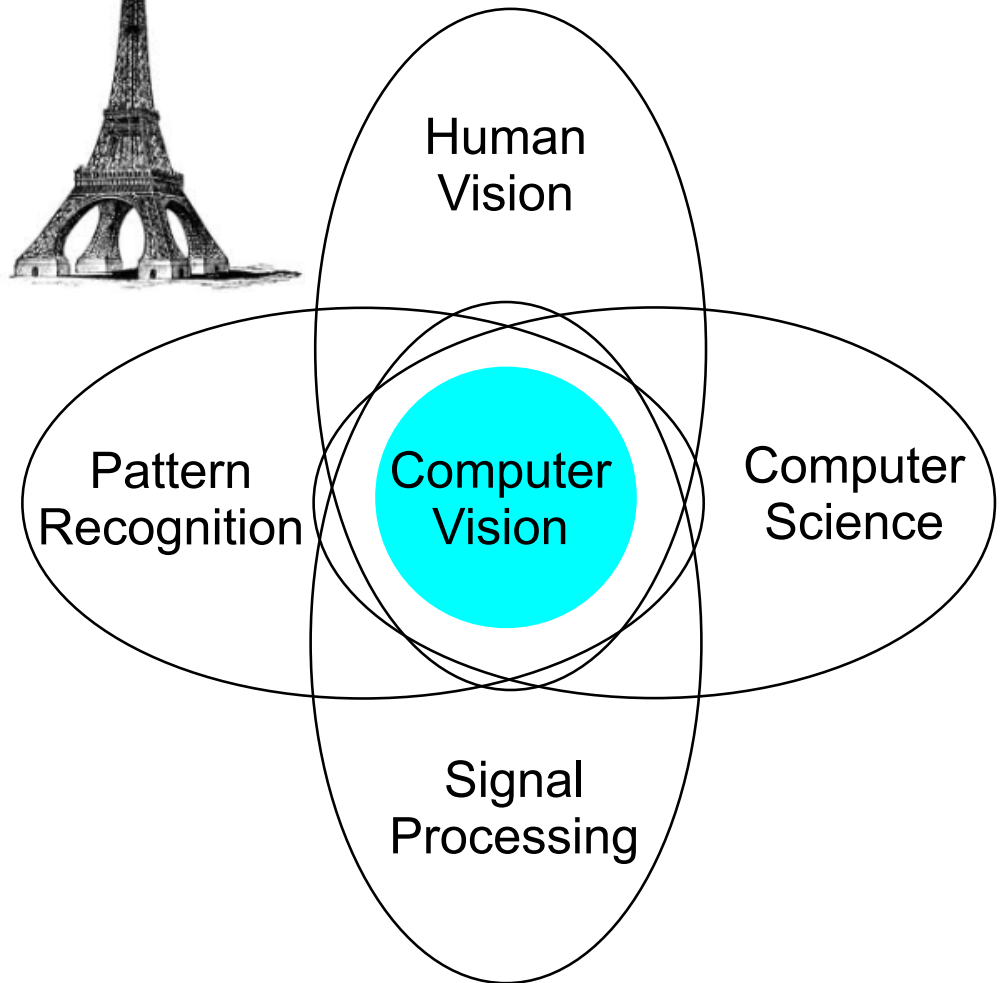
Počítačové vidění = kamera + počítač + ?

◆ **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?

- ◆ Jako součást řídicích systému (např. u průmyslových robotů nebo autonomně jedoucích aut).
- ◆ Pro detekci událostí (např. při sledování bezpečnostními kamerami, počítání lidí, při detekci startující balistické rakety ze senzorů na družici).
- ◆ Pro uspořádání informace (např. pro indexování obrazových databází nebo video sekvencí).
- ◆ 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.
- ◆ . . .

Vnímání

- ◆ Postupy k upoutání pozornosti a/nebo porozumění informacím ze senzorů.
- ◆ Úloha je mnohem složitější, než si vědci byli schopni představit okolo let 1950 a 1960:
 - “Vytvoření strojů vnímat potrvá 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.

Poznámky k lidskému (vizuálnímu) vnímání

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 velmi těžké, blízké k nemožnému).



Poznámky k lidskému (vizuálnímu) vnímání

Co vidíte na obrázku?

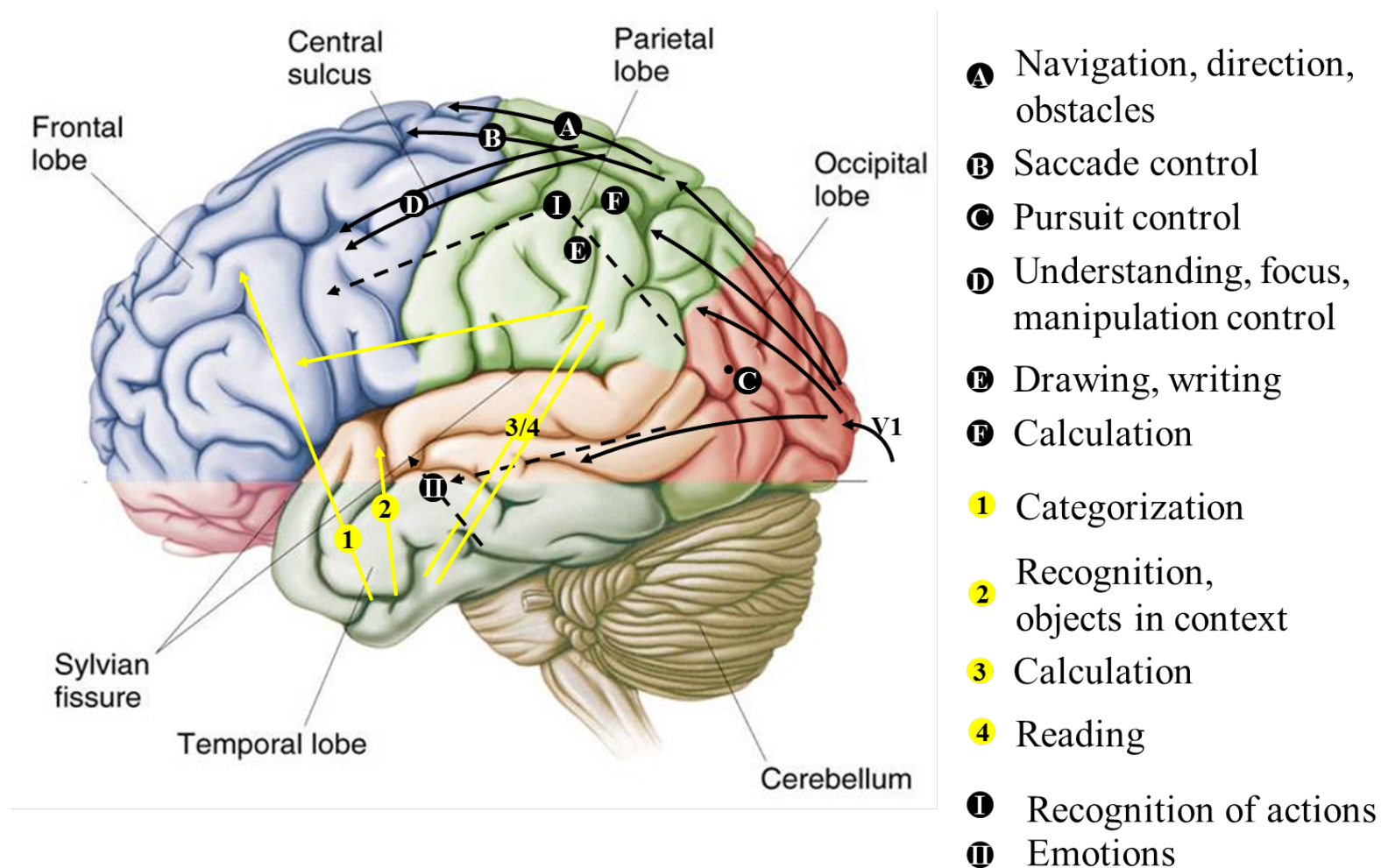
Vidění a vnímání je pro člověka a mnoho zvířat snadné a přirozené.



- ◆ 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 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.



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:

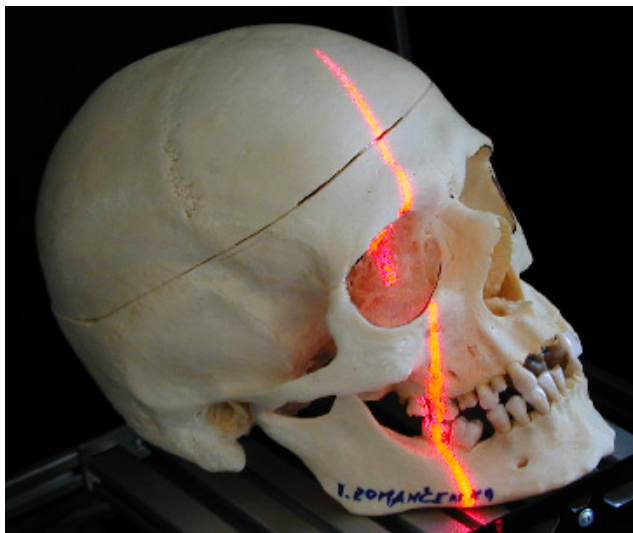
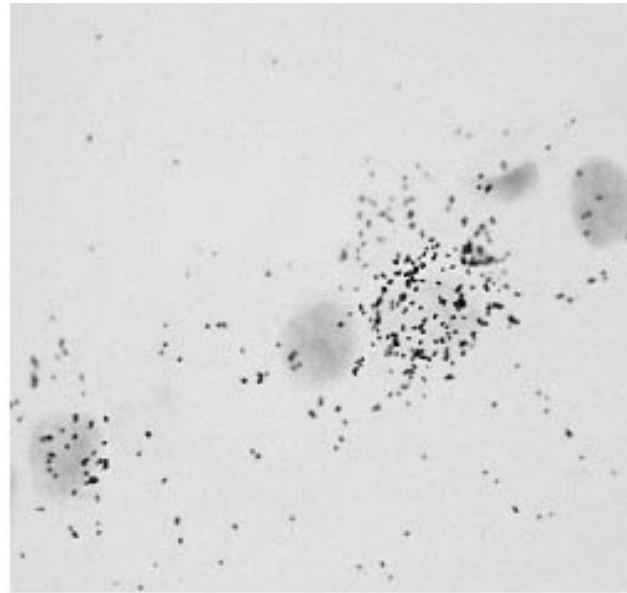
- ◆ Hledají empirické modely popisující biologické vidění.
- ◆ Někdy popisují vidění, jako by šlo o výpočetní systém.

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

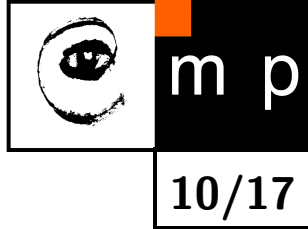
- ◆ Biologické vidění zvládá úlohy, na něž je počítačové vidění stále krátké. Přesto poskytuje biologické vidění inspiraci i pro technická řešení.
- ◆ Technické požadavky na systémy počítačového vidění se často shodují s požadavky na biologické vidění.

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

Příklady vstupních obrazů

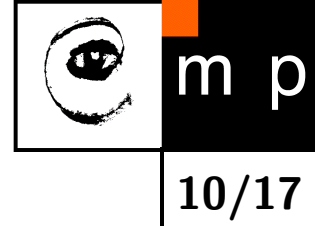


**Proč je počítačové vidění těžké ?
Najdeme alespoň 6 příčin.**



Proč je počítačové vidění těžké ?

Najděme 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.

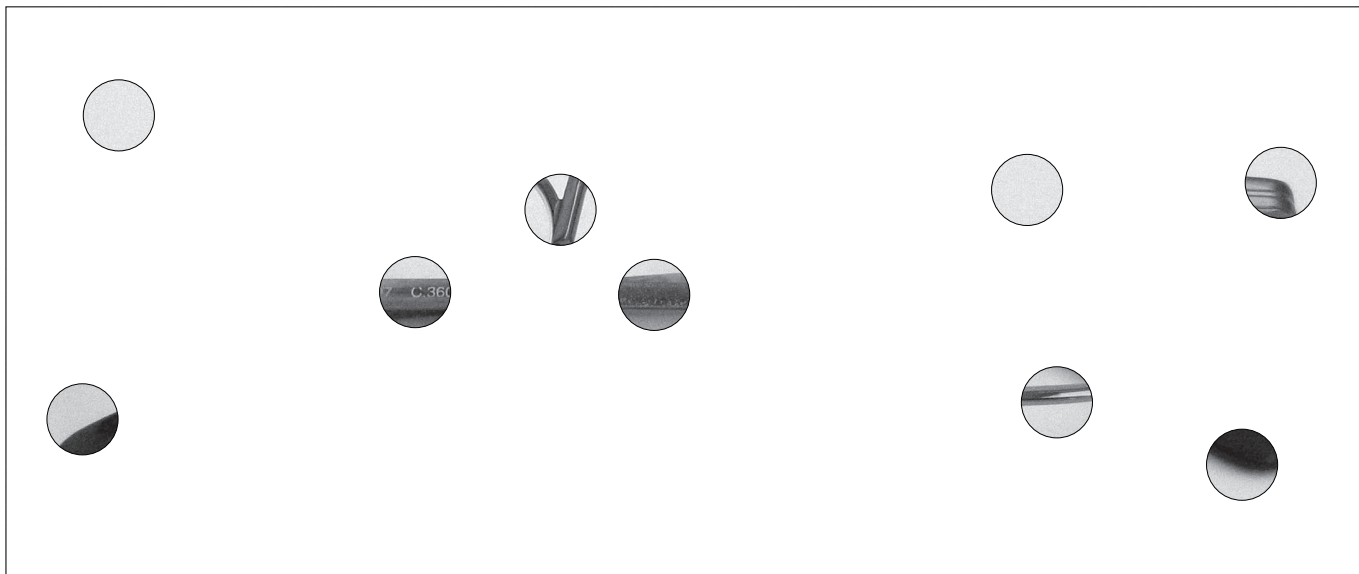
Příliš mnoho dat Stránka A4, 300 dpi, 8 bit per pixel = 8.5 Mbytes.

Neprokládané video 512 × 768, RGB (24 bit) = 225 Mbits/sekundu.

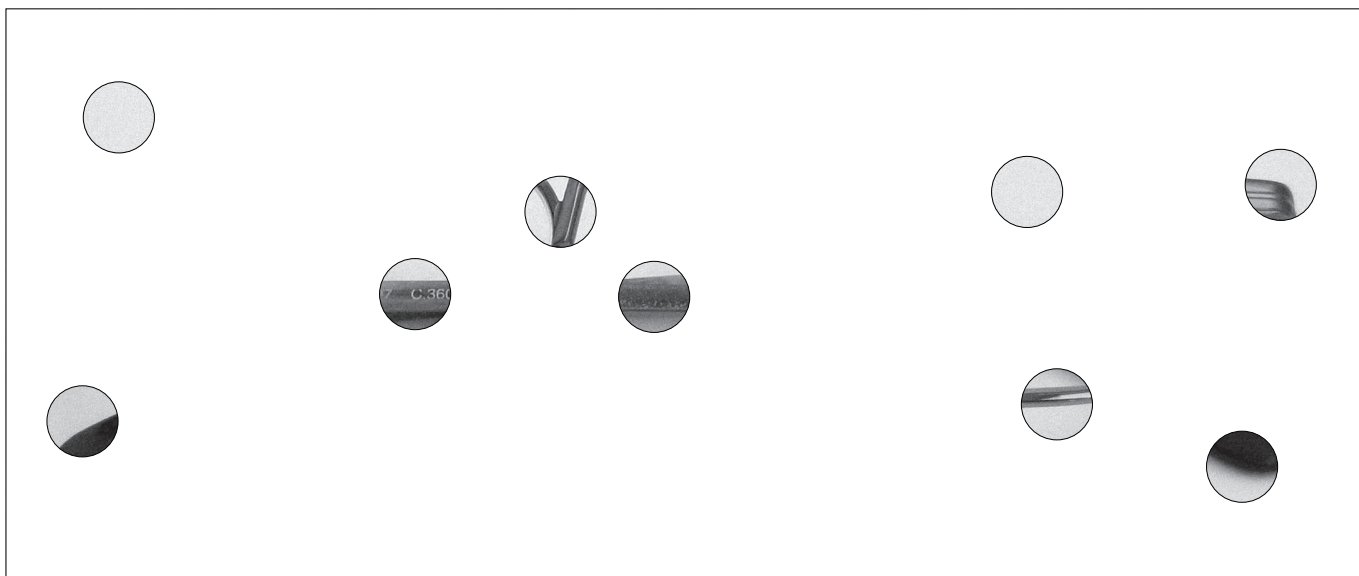
Nutnost zahrnout interpretaci (bude brzy diskutováno).

Lokální okno v kontrastu s potřebou globálního pohledu

Ilustrace nedostatečnosti lokálního pohledu



Ilustrace nedostatečnosti lokálního pohledu



Interpretace a její role, sémantika

Interpretace: pozorování \rightarrow model světa
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, neformálně: logický systém s kvantifikátory \forall , \exists nemůže být dokázán ani vyvrácen.

Od nízké k vyšší úrovni zpracování z hlediska využívané apriorní znalosti

Nízká (až žádná) znalost = 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 = 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 = 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.

Role apriorní znalosti, protipříklad

- ◆ Apriorní znalost “našeho světa” umožňuje člověku porozumět i mnohoznačným obrázkům.
- ◆ Ovšem, apriorní očekávání mohou také přivést k chybné interpretaci ...
- ◆ Protipříklad: Amesova židle.



Vidíme židle.

Role apriorní znalosti, protipříklad

- ◆ Apriorní znalost “našeho světa” umožňuje člověku porozumět i mnohoznačným obrázkům.
- ◆ Ovšem, apriorní očekávání mohou také přivést k chybné interpretaci ...
- ◆ Protipříklad: Amesova židle.



Vidíme židle.



Ve scéně židle nejsou.

1966 M. Minsky zadává úlohu počítačového vidění bakalářskému studentovi jako prázdninový projekt.

- ~**1960** Interpretace v omezeném umělém světě, např. svět kostek robotu.
- ~**1970** Jistý pokrok v interpretaci obrazů v omezeném světě.
- ~**1980** Umělé neuronové sítě přišly a odešly; posun zájmu ke geometrii a rigoróznějšímu použití matematiky; inspirace biologickým viděním (D. Marr a spolupracovníci).
- ~**1990** Detekce a rozpoznávání lidských obličejů; růst popularity statistické analýzy; zájem o geometrické úlohy vidění.
- ~**2000** Rozpoznávání ve větším; k dispozici začínají být rozsáhlé anotované databáze; počátek prakticky použitelných metod analýzy videa.

Rozpoznávání na základě obrazů hierarchie reprezentací

