

# Senzo-motorický prístup k rozpoznávaniu scény

**Andrej Lúčný**

**KAI FMFI UK Bratislava**

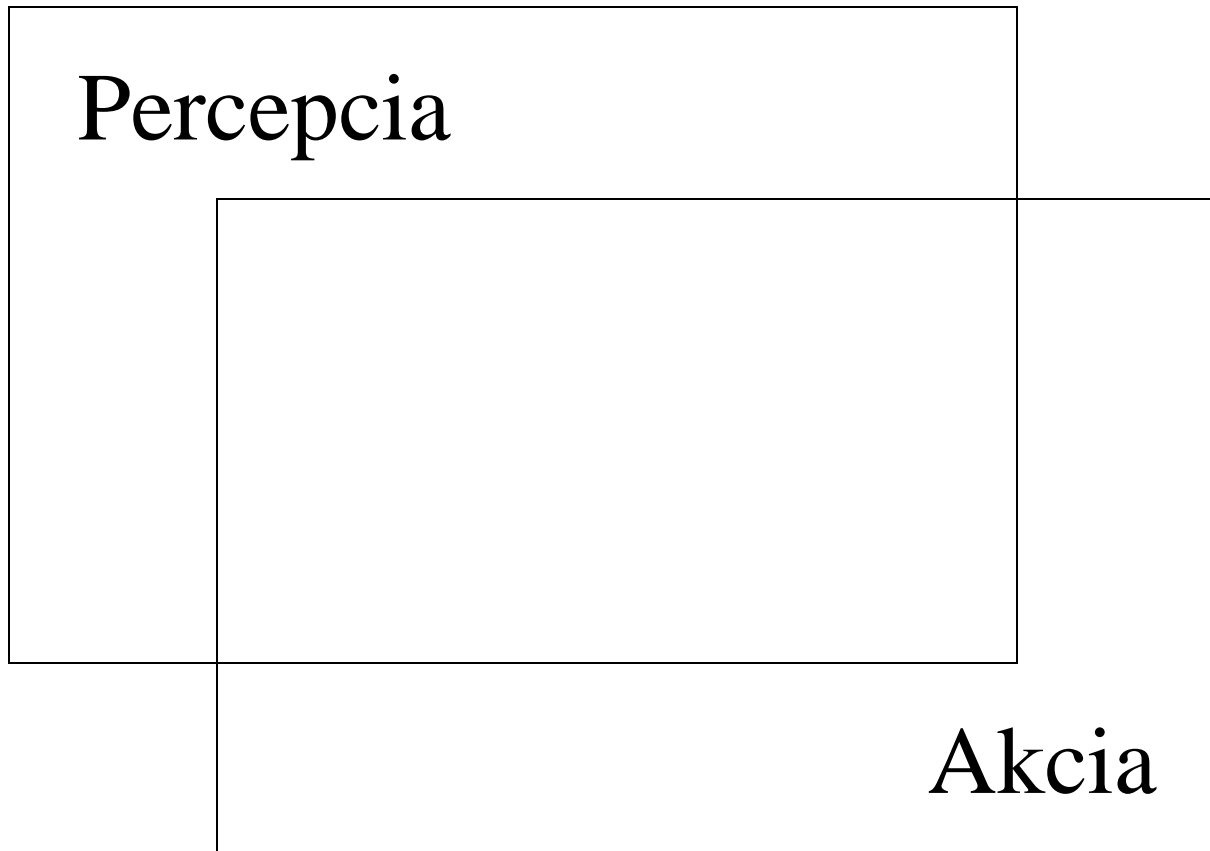
**andy@microstep-mis.com**

**<http://www.microstep-mis.com/~andy>**

# Tradičná predstava o kognícii

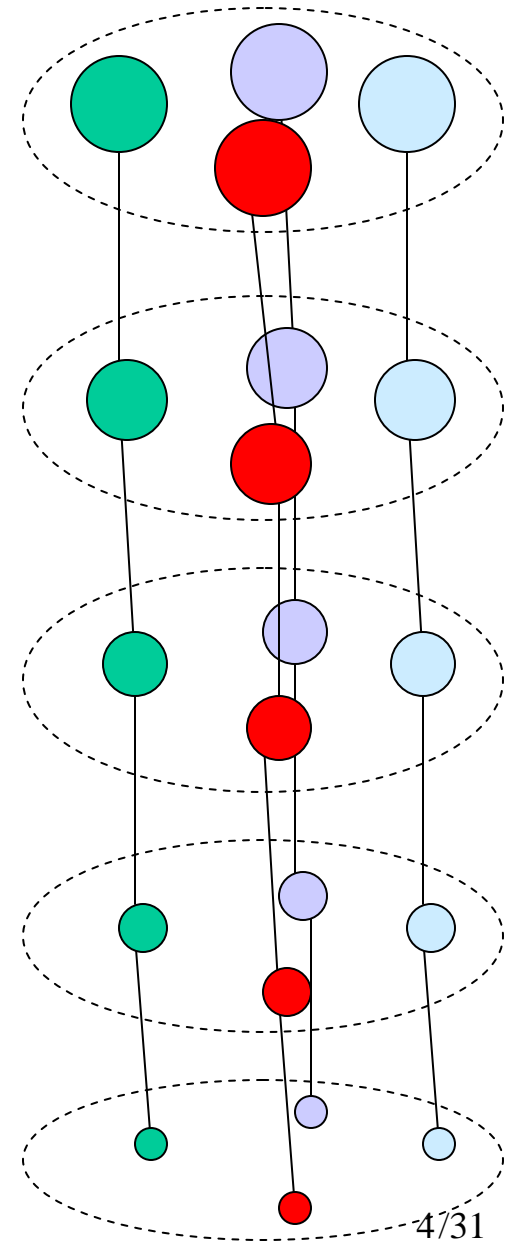
Percepcia	Kognícia	Akcia
-----------	----------	-------

# Moderná predstava o kognícii



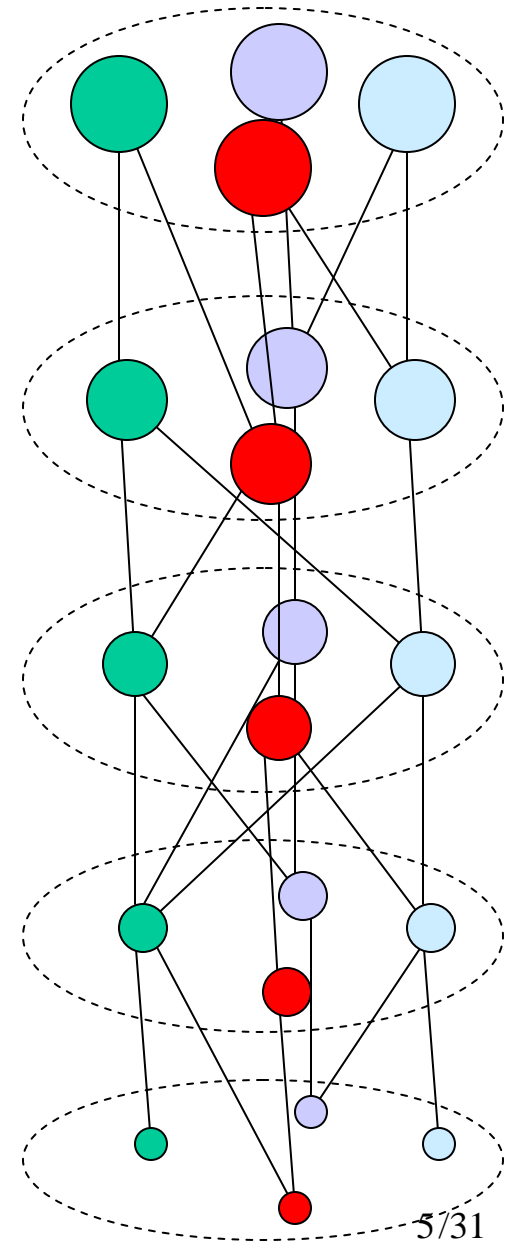
# Napodobňovanie evolúcie

- Živé systémy vyznačujúce sa kogníciou vznikli procesom evolúcie
- Všetky časti mysle sa pritom vyvinuli postupným a spoločným povstávaním z jednoduchších verzií



# Napodobňovanie evolúcie

- Pokrok jednej časti sa však pritom mohol oprieť o vtedajší stav všetkých ostatných častí
- K napodobneniu tohto procesu teda nemusí stačiť zaoberať sa jedinou časťou



# Evolúcia percepcie

- **A to platí aj percepciu !**
- **Východisko: senzomotorický prístup**

# Senzomotorický prístup

- Piaget: „Myslenie nevzniká len z vnímania, ale zo senzomotorickej činnosti“
- Špeciálne pre percepciu z toho vyplýva, že
  - sa opiera o akciu
  - neprebíha nezávisle na akciách
  - časť akcií konštruuje
  - nie je to pasívny ale aktívny proces

# Senzomotorický prístup

- Čo svedčí v prospech senzomotorického prístupu v iných vedných disciplínach ?
- Psychológia skúmajúca vývin kognície:
- Jean Piaget: senzomotorické obdobie vývinu dieťaťa
- Linda Smith: senzomotorické obdobie trvá svojím spôsobom po celý život



# Skúmanie vývinu kognitívnych schopností u detí

- Jean Piaget: „A not B error“  
Fundamentálnosť akcie



- Linda Smith: rozdiely zameriavania pozornosti na objekty v scéne u detí a dospelých



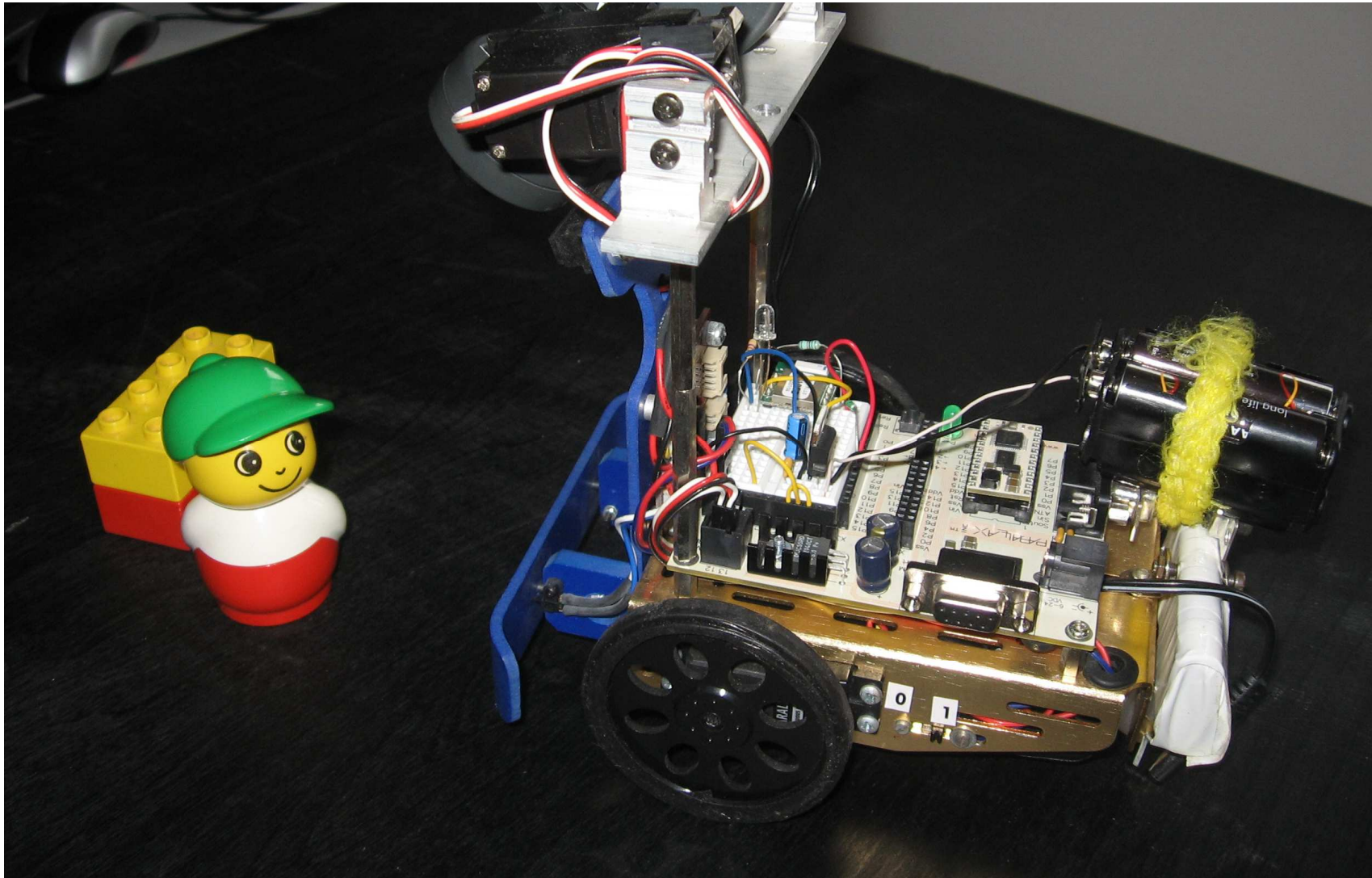
# Prístup k percepcii v počítačovom videní

- Percepciu chápeme ako niečo uzavreté a oddelené od zvyšku mysle
- Tento prístup je typický pre počítačové videnie
- Tento prístup vyplýva z tradičného redukcionizmu

# Senzomotorický prístup k percepcii

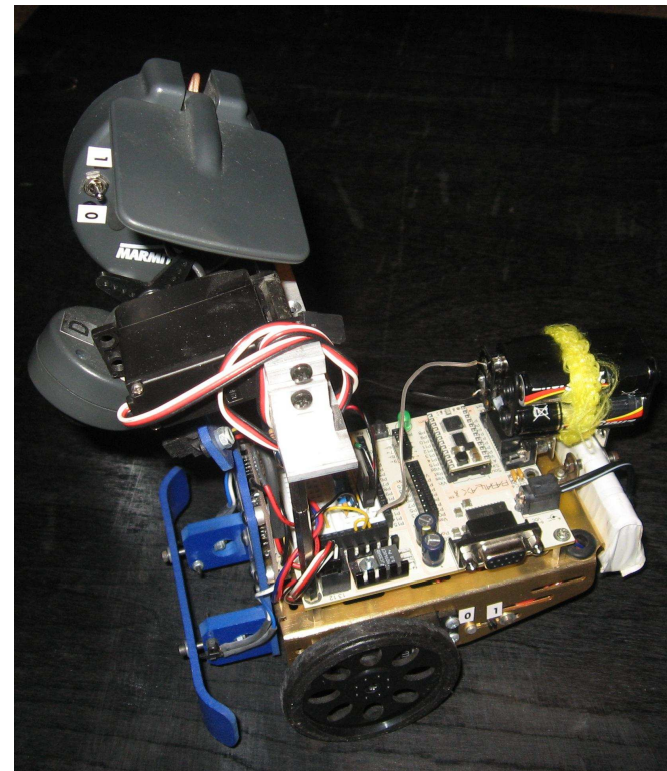
- Vedeli by sme senzomotorický prístup použiť ako nový impulz pre počítačové videnie?
- Áno, pokiaľ ukážeme, že si percepcia konštrukciou vhodných akcií môže významným spôsobom uľahčiť svoje postavenie

# Príklad: Rozpoznávanie scény



# Použité prostriedky

- Podvozok Boe-bot Javelin ovládaný cez bluetooth
- Bezdrôtová kamera XCam2
- Ovládanie z PC: JVM, architektúra Agent-Space ([www.agentspace.org](http://www.agentspace.org)), JMF, Java rxtx



# Ciel': zostrojit' kognitívny model videného



- Tvar
- Veľkosť
- Umiestnenie
- ...
- Vidíme jeden objekt alebo dva ?

# Tradičný prístup

- Tento problém sa dá riešiť vychádzajúc výlučne z daného obrázku.
- Môžeme použiť rôzne algoritmy vychádzajúce z Guzmanovej metódy.
- Vstupom je jediný obrázok a snažíme sa dômyselnými a výpočtovo náročnými algoritmami vyextrahovať z neho čo najviac informácie

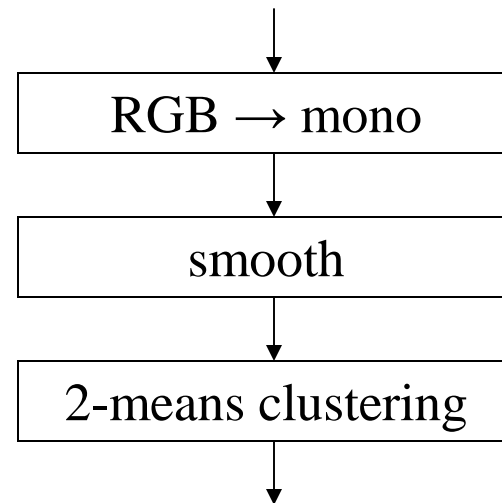
# Náš prístup

- Skonštruujeme vhodnú aktivitu robota
- Budeme pracovať s mnohými obrázkami z rôznych pohľadov
- Budeme používať len jednoduché algoritmy vykonateľné v reálnom čase.
- Budeme využívať znalosť akcie, ktorú konáme a pri spracovaní percepcie počítat' s dátami, ktoré pre potreby akcie sami generujeme



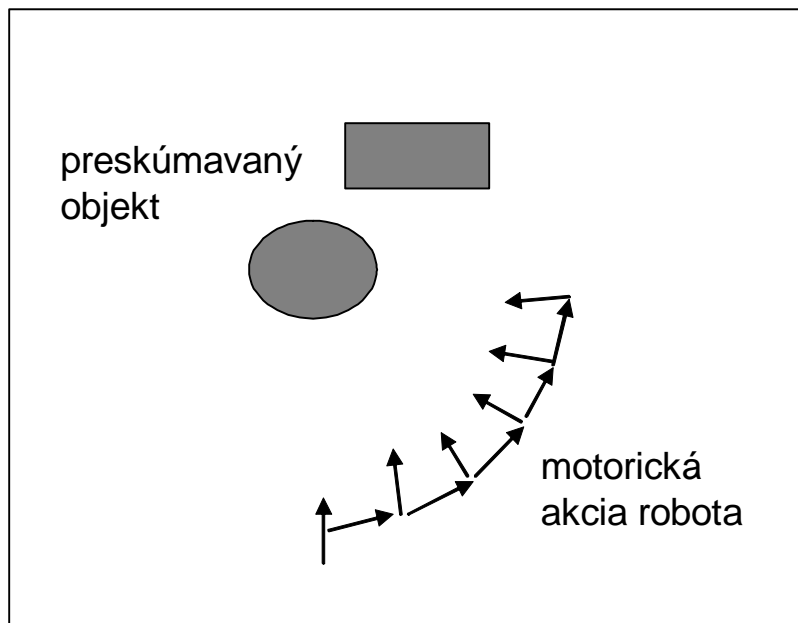
# Riešenie (1)

- Robot zaujme pozíciu kedy má objekt (zatiaľ vnímaný ako machuľa) v strede obrazu



# Riešenie (2)

- Spustí generovanie akcie pri ktorej objekt obchádza sprava, pričom sa pod chvíľou natáča k predmetu



## Riešenie (3)

- Keďže akciu generuje robot, môže ju generovať vždy rovnako
- Vďaka tomu sa na objekt pozerá vždy zhruba z rovnakých pozícií
- (aj keby to tak nebolo môže odometriou odhadnúť odkiaľ sa díval)
- (robot nevie hýbať kamerou, takže vždy je rovnaká i výška i uhol ktorým sa díva)

# Riešenie (4)

- Výsledkom akcie sú pohľady na robota z rôznych známych pozícií

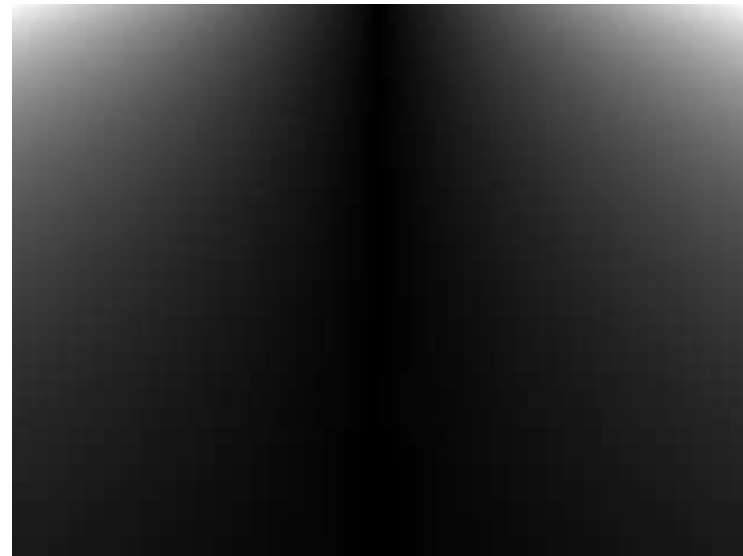


# Riešenie (5)

- Z toho môžeme pristúpiť k rekonštrukcii 3D objektov v scéne
- (Rekonštrukciu možno urobiť aj z pohľadov z neznámej polohy, je to však výpočtovo náročné. Naopak zo známej polohy sa dá ako-tak rekonštruovať v reálnom čase)

# Riešenie (6)

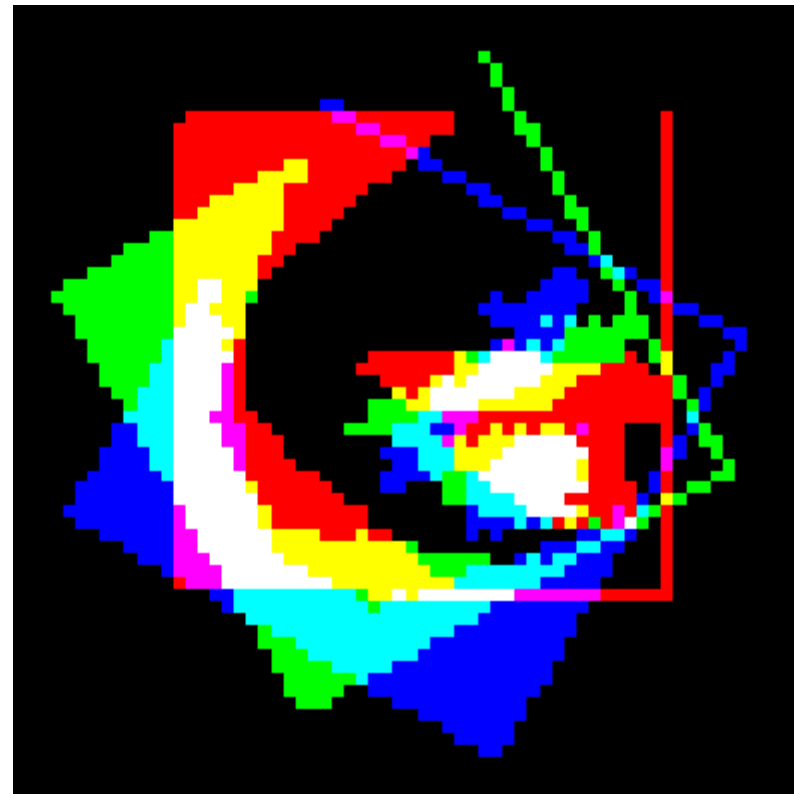
- Využívame pritom kalibráciu kamery urobenú off-line, ktorá dáva podlahové súradnice pre každý pixel na obraze



(použitá geometria je v článku)

# Riešenie (7)

- Pokiaľ robot vidí niekde podlahu, znamená to, že priamka ktorou sa díva neprechádza predmetom, môže ju teda využiť ako „rezací lúč“ v 3D priestore



(jeden z vodoravných rezov z 23/31  
troch pohľadov)

# Riešenie (8)

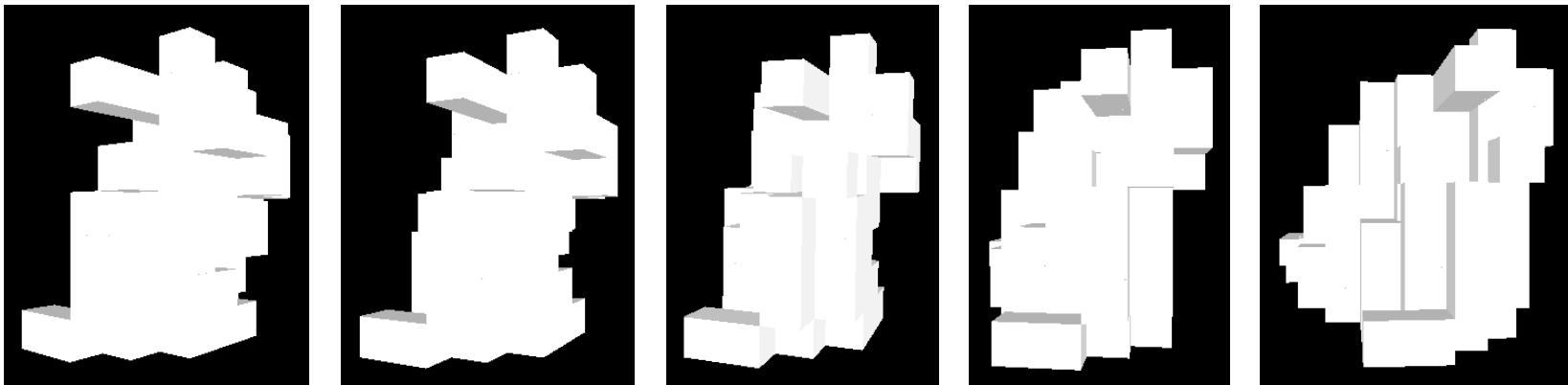
Výsledkom je určitá rekonštrukcia tvaru objektov





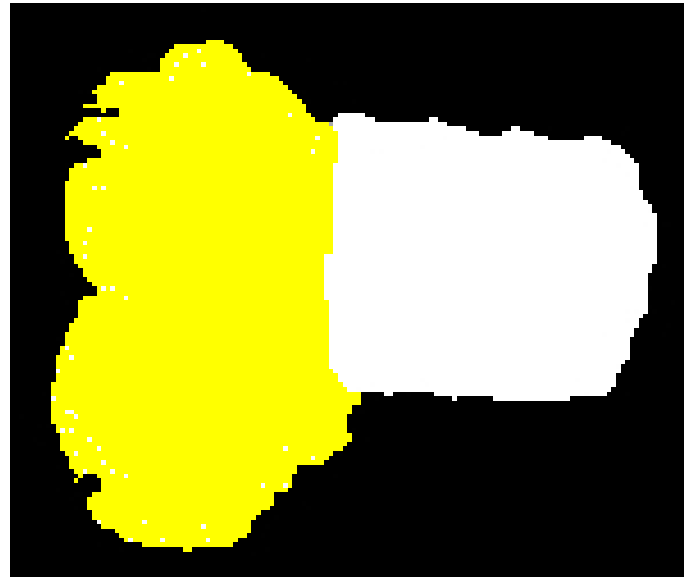
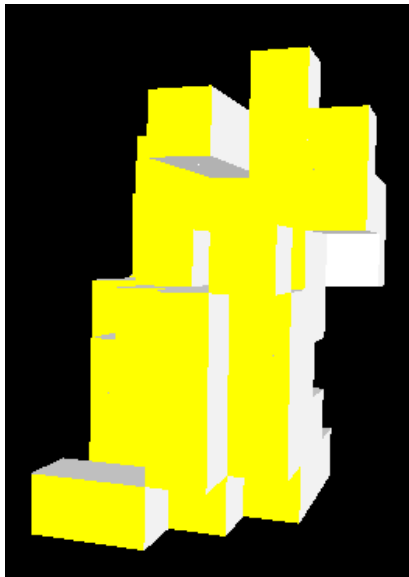
## Riešenie (9)

- Z rekonštrukcie vieme koľko objektov v scéne je (t.j. konkrétne, že sú dva)
- Teoreticky by sme teraz mohli renderovať všetky možné pohľady na jednotlivé objekty (keďže sa kamera nehýbe, nie je ich moc)



# Riešenie (10)

- Tieto pohľady by sme mohli hľadať na obraze a pokiaľ by došlo k určitej zhode, objekt by sme ďalej neskúmali



# Riešenie (11)

- Následkom toho môže samozrejme dôjsť i k seba klamu – ale rovnako sa mýlime i my.



# Negatíva

- Tento prístup je skôr ukážkou ako by sa na rozpoznávanie scény mohlo senzomotorickým prístupom ísť, než nejakým hotovým riešením, má nízku úspešnosť
- problémy: len jedna kamera, hardware úbohých parametrov, kontrast – čierna podlaha, tiene vyvolávajúce diery v objektoch, nepoužili sme neurčitosti
- V princípe ale myšlienku považujeme za správnu

# Pozitíva

- Práca v reálnom čase
- Jednoduché, výpočtovo nenáročné metódy
- Náznak toho, že senzo-motorický prístup naozaj zjednodušuje riešený problém

# Záver

- Budúcnosťou výskumu v počítačovom videní a ďalších oblastiach zameriavajúcich sa na percepciu je možno práve prekročenie hraníc daných úzkych domén smerom k motorickým činnostiam a komplexným systémom
- Motiváciou k takémuto posunu je kognitívna veda

**Ďakujem za pozornosť !**

Senzo-motorický prístup k  
rozpoznávaniu scény

Andrej Lúčný  
KAI FMFI UK, Bratislava  
andy@microstep-mis.com  
<http://www.microstep-mis.com/~andy>