Matematika v računalniškem vidu

Marko Boben

Univerza na Primorskem – IAM in Univerza v Ljubljani – FRI

Izleti v matematično vesolje, 11 februar 2015

Matematika v računalniškem vidu

Marko Boben
Univerza na Primorskem – IAM in Univerza v Ljubljani – FRI

Iz časa dela v laboratoriju VICOS, FRI UL in sodelovanja s
S. Fidler in A. Leonardisom

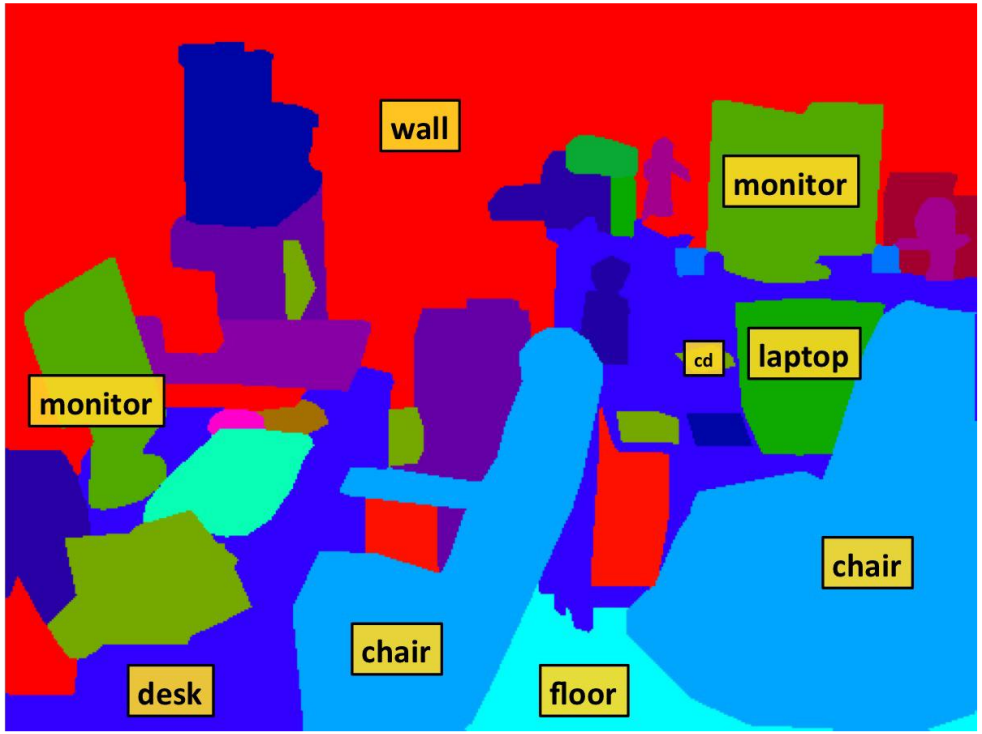
Kaj je računalniški vid?

- Področje računalništva, ki poskuša razviti algoritme, ki omogočijo, da računalnik “vidi”.

Del prosojnic je povzet po prosojnicah S. Fidler za predmet CSC 420: Introduction to Image Understanding na University of Toronto, Kanada

Kaj je računalniški vid?

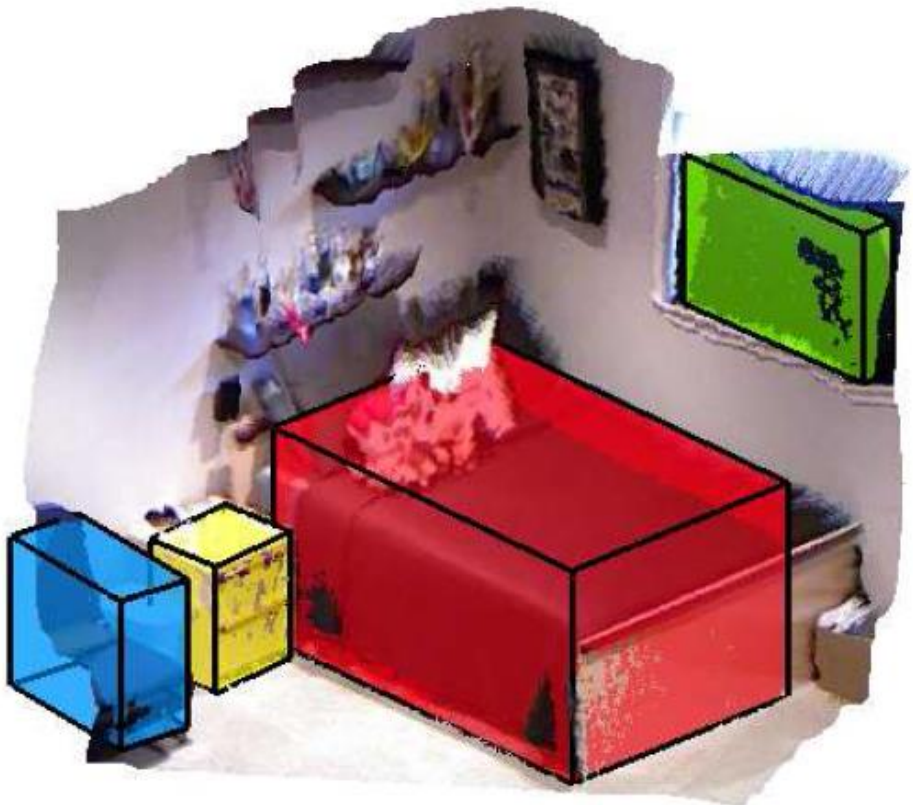
- Kaj pomeni “videti”?
  - Ugotoviti kateri predmeti so na sliki



Prosojnice: S. Fidler

Kaj je računalniški vid?

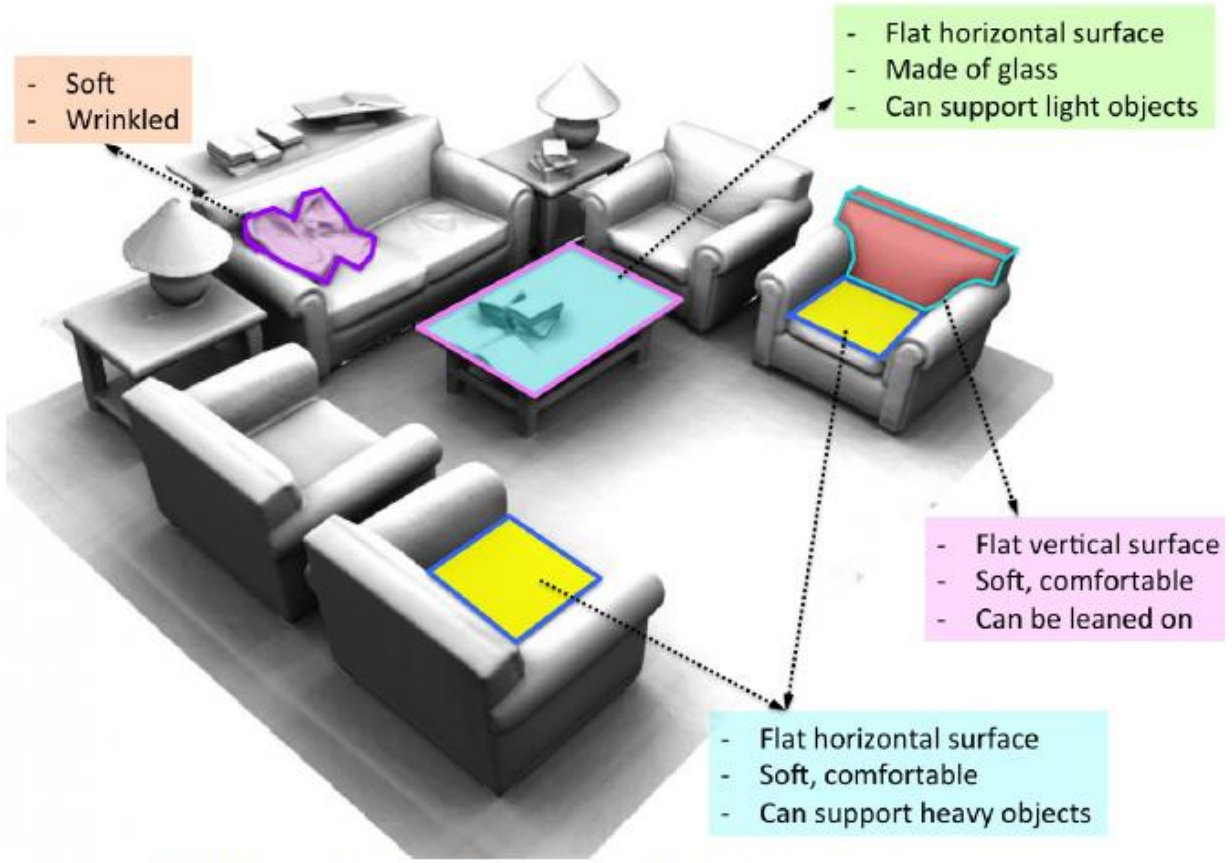
- Kaj pomeni “videti”?
  - Ugotoviti kateri predmeti so na sliki
  - Razumeti kje na sliki so predmeti



Prosojnice: S. Fidler

Kaj je računalniški vid?

- Kaj pomeni “videti”?
  - Ugotoviti, kateri predmeti so na sliki
  - Razumeti, kje na sliki so predmeti
  - Ugotoviti 3D lastnosti

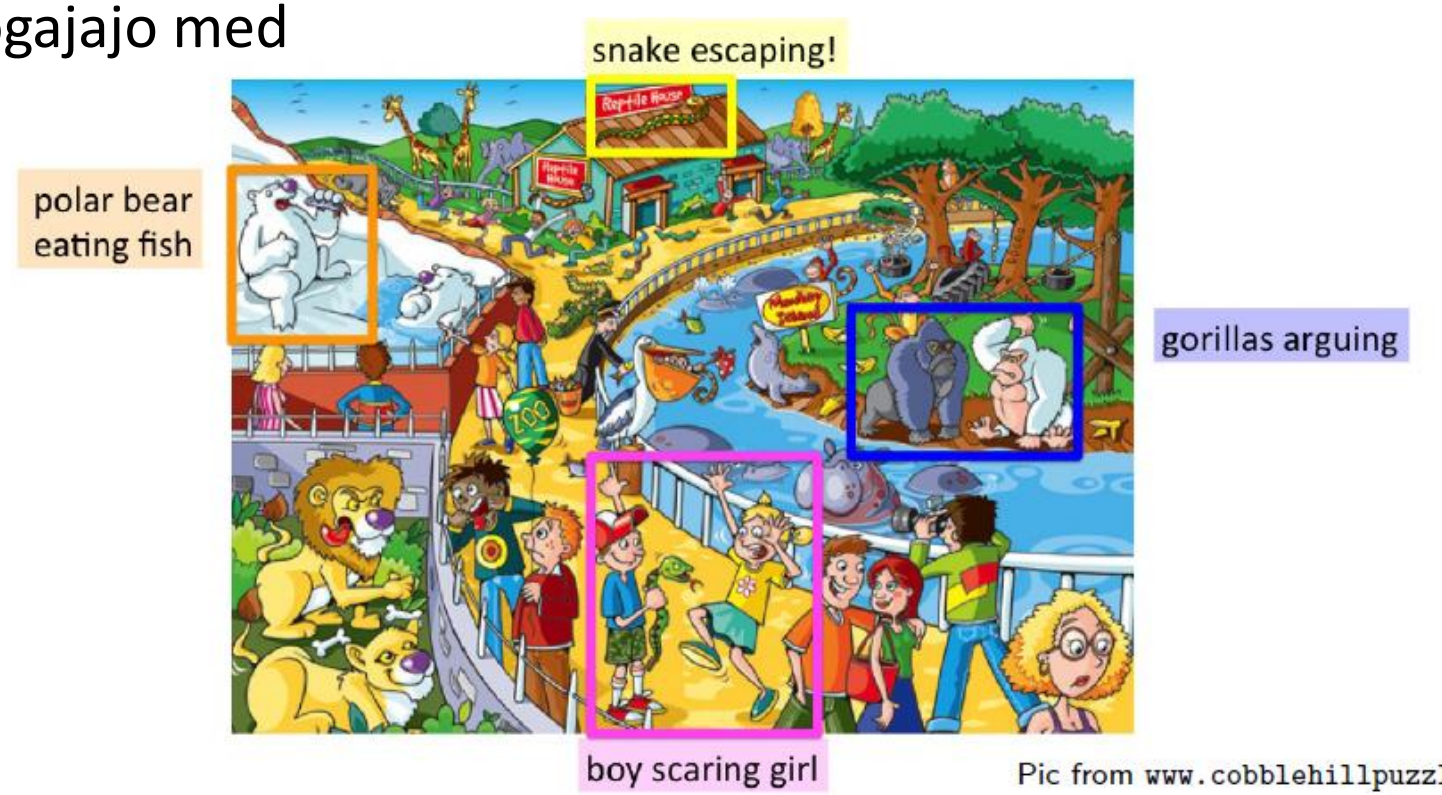


Depth pic from http://vladlen.info

Prosojnice: S. Fidler

Kaj je računalniški vid?

- Kaj pomeni “videti”?
  - Ugotoviti kateri predmeti so na sliki
  - Razumeti kje na sliki so predmeti
  - Ugotoviti 3d lastnosti
  - Ugotoviti akcije, ki se dogajajo med objekti na sliki



Zakaj računalniški vid?

- Da bodo v prihodnosti roboti zlagali perilo...

https://www.youtube.com/watch?v=gy5g33S0Gzo

https://www.youtube.com/watch?v=KKUaVzf3Oqw

Zakaj računalniški vid?

- ... nas vozili na delo

Amnon Shashua's Mobileye autonomous driving system
https://www.youtube.com/watch?v=4fxFDypHZLs

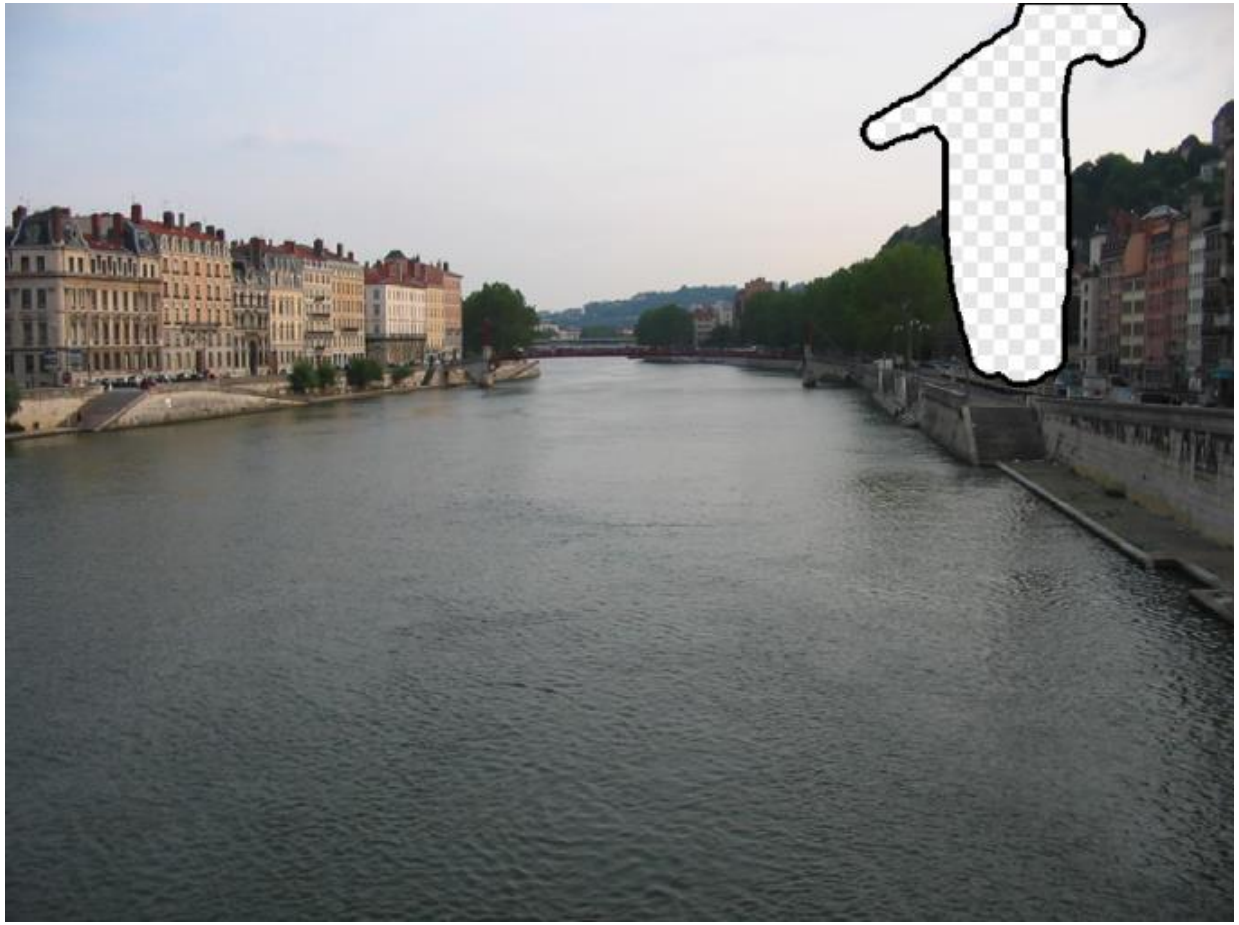
Zakaj računalniški vid?

- Računalniški vid omogoča manipulacijo slik

Scene Completion using Millions of Photographs, Hays & Efros, SIGGRAPH 2007

Zakaj računalniški vid?

- Računalniški vid omogoča manipulacijo slik



Scene Completion using Millions of Photographs, Hays & Efros, SIGGRAPH 2007

Zakaj računalniški vid?

- Računalniški vid omogoča manipulacijo slik

Scene Completion using Millions of Photographs, Hays & Efros, SIGGRAPH 2007

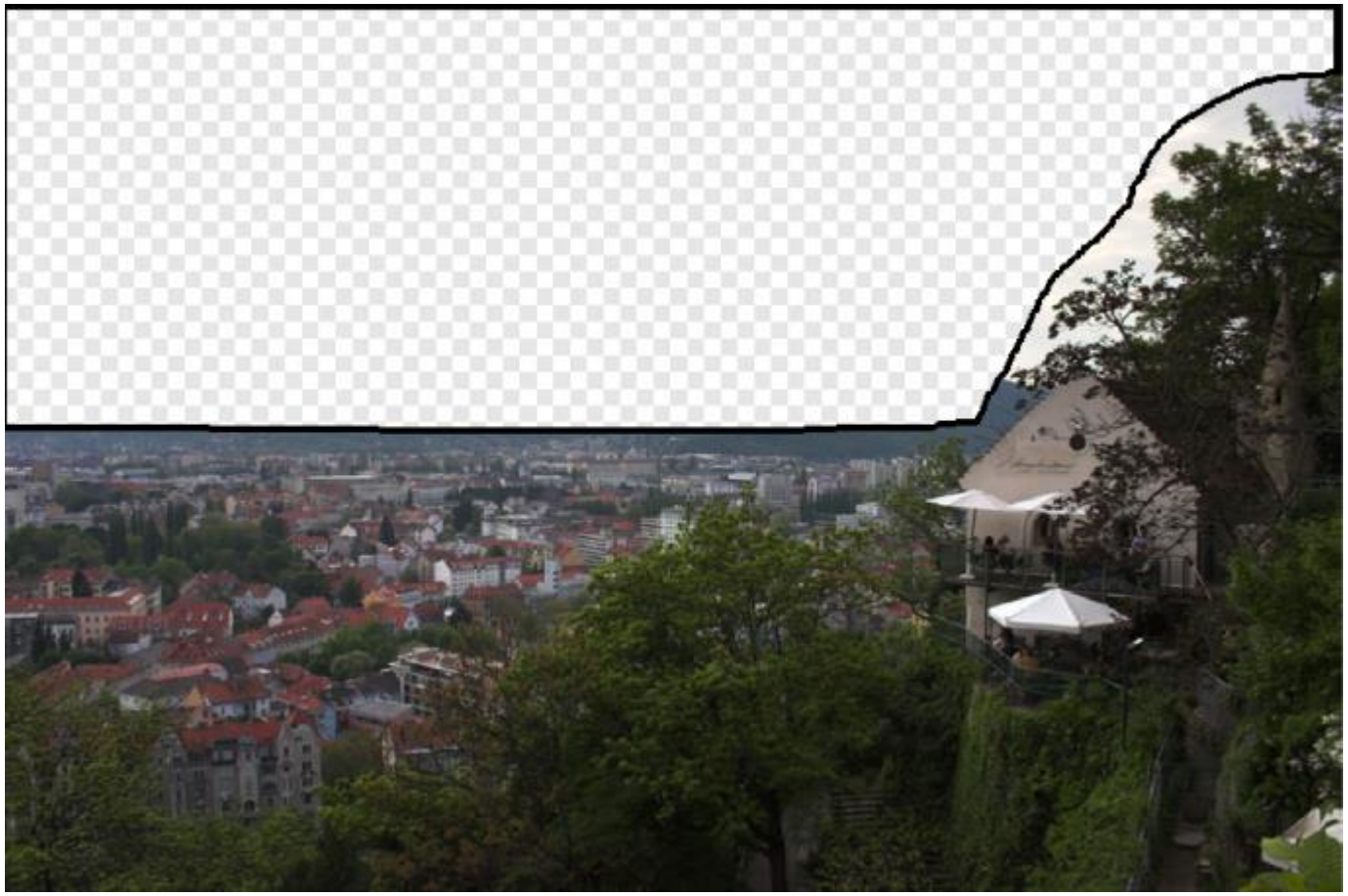
Zakaj računalniški vid?

- Računalniški vid omogoča manipulacijo slik

Scene Completion using Millions of Photographs, Hays & Efros, SIGGRAPH 2007

Zakaj računalniški vid?

- Računalniški vid omogoča manipulacijo slik



Scene Completion using Millions of Photographs, Hays & Efros, SIGGRAPH 2007

Prosojnice: S. Fidler

Zakaj računalniški vid?

- Računalniški vid omogoča manipulacijo slik

Scene Completion using Millions of Photographs, Hays & Efros, SIGGRAPH 2007

Zakaj računalniški vid?

- Računalniški vid omogoča manipulacijo slik

https://www.youtube.com/watch?v=ipTyCJi0t1Y

3D Object Manipulation in a Single Photograph using Stock 3D Models, Kholgade, Simon, Efros, Sheikh, SIGGRAPH 2014

Zakaj računalniški vid?

- Vizualizacije in analize v športu

Prosojnice: S. Fidler

Zakaj računalniški vid?

- Posebni učinki v filmih

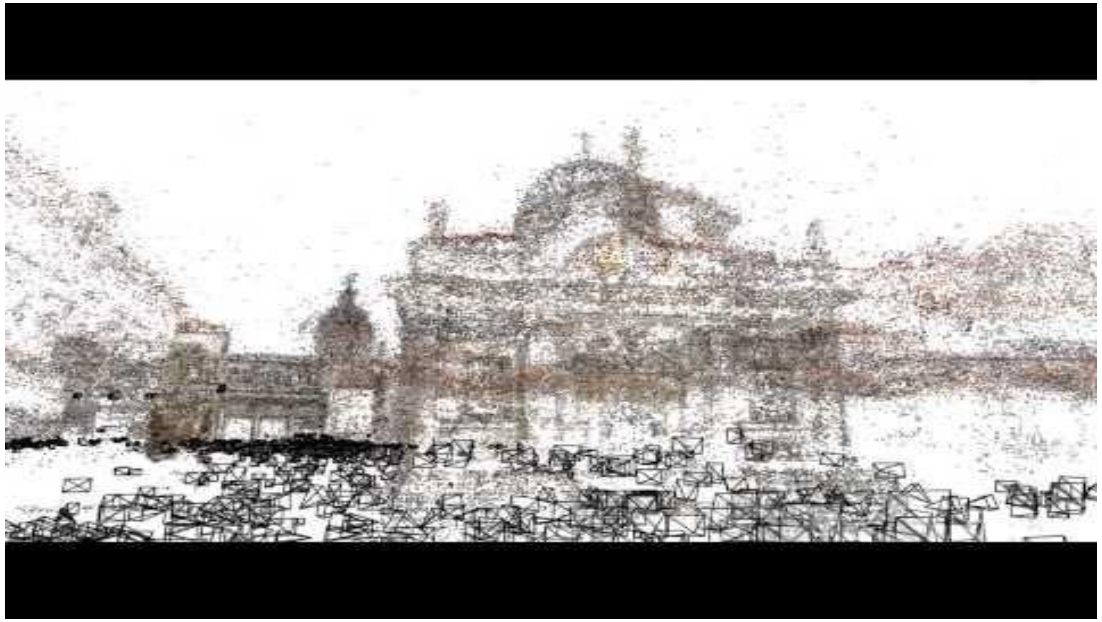
https://www.youtube.com/watch?v=WDwTQ57YyzI#t=21

http://vimeo.com/100095868

Zakaj računalniški vid?

- 3d rekonstrukcija objektov iz slik

https://www.youtube.com/watch?v=IgBQCoEfiMs

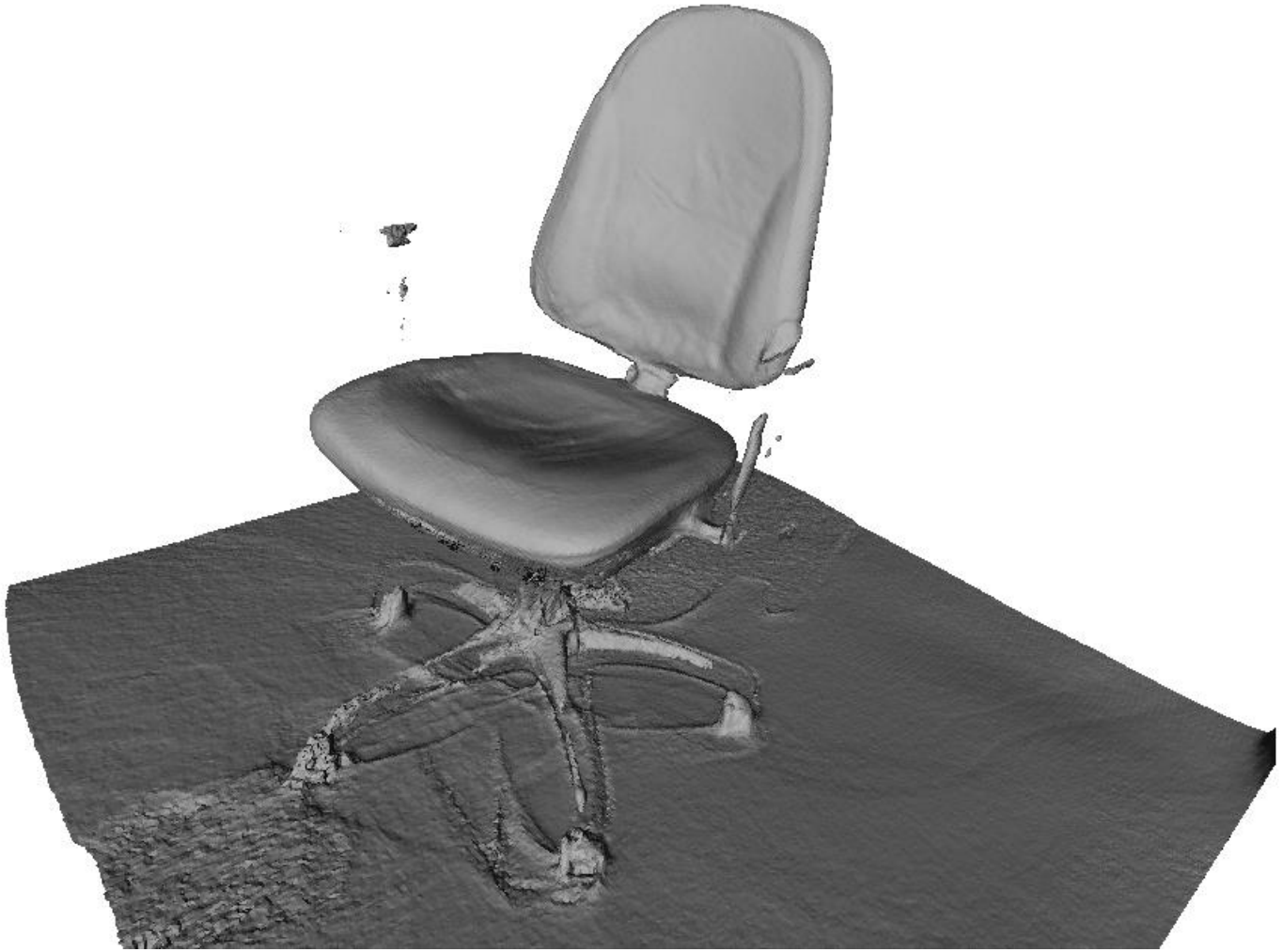


https://www.youtube.com/watch?v=sQegEro5Bfo

Photosynth, https://photosynth.net/

Zakaj računalniški vid?

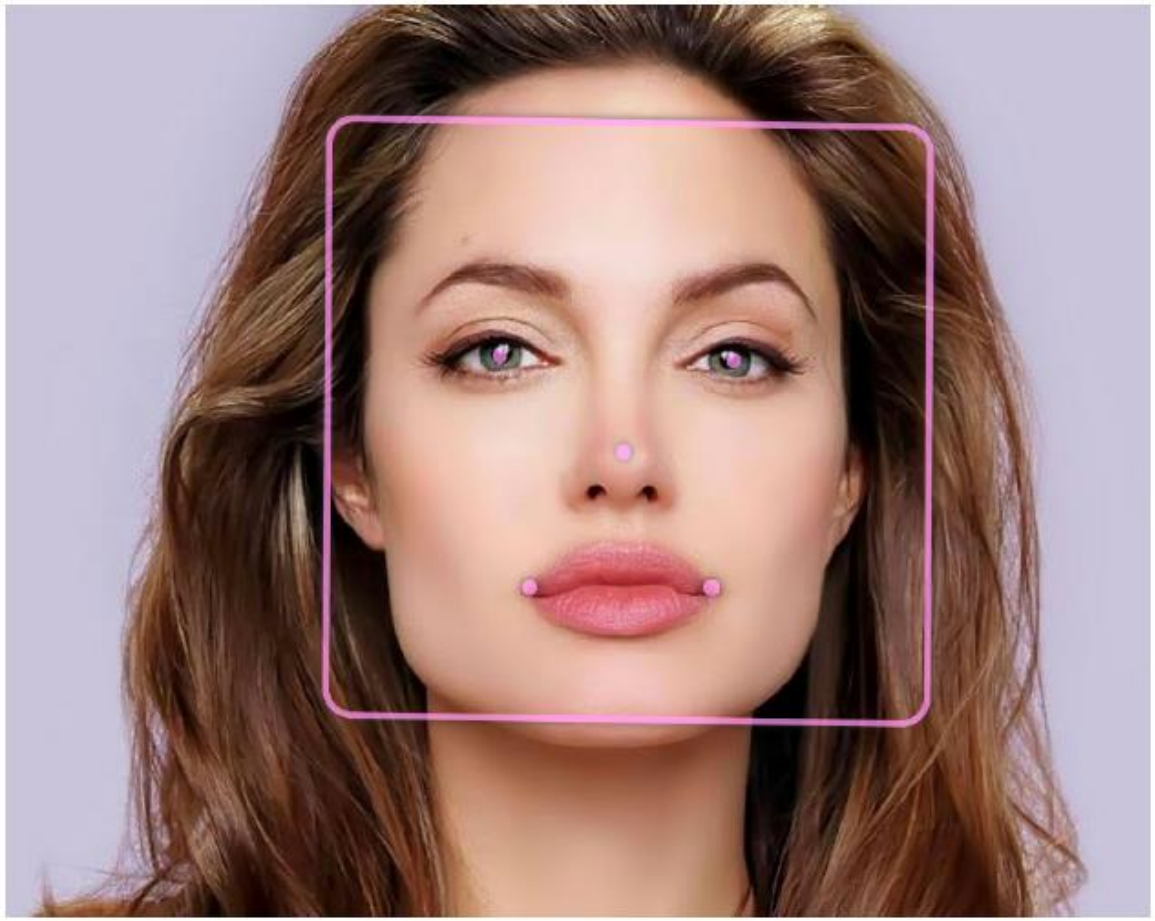
- 3d rekonstrukcija objektov iz globinskih slik



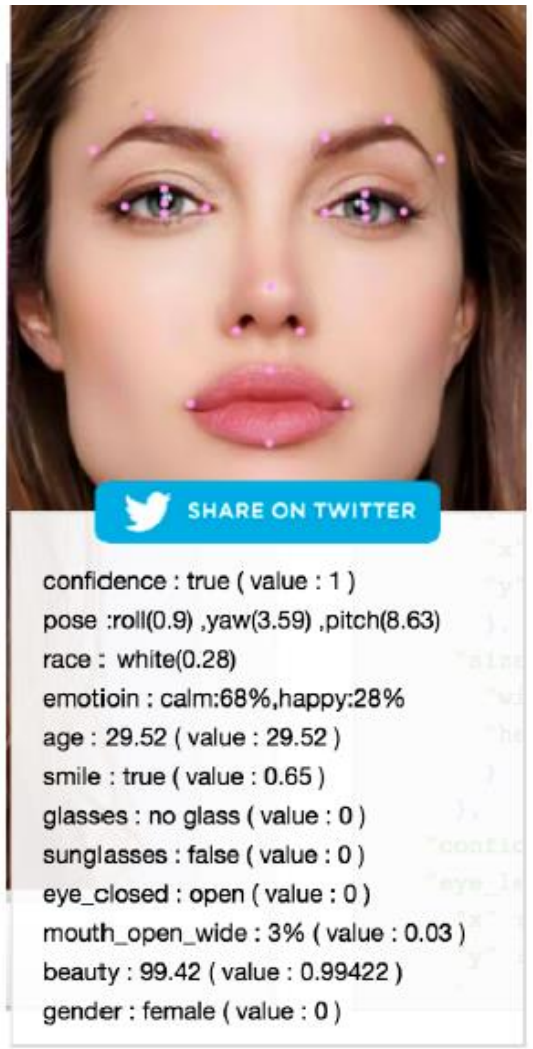
Prosojnice: S. Fidler

Zakaj računalniški vid?

- Detekcija in analiza obrazov



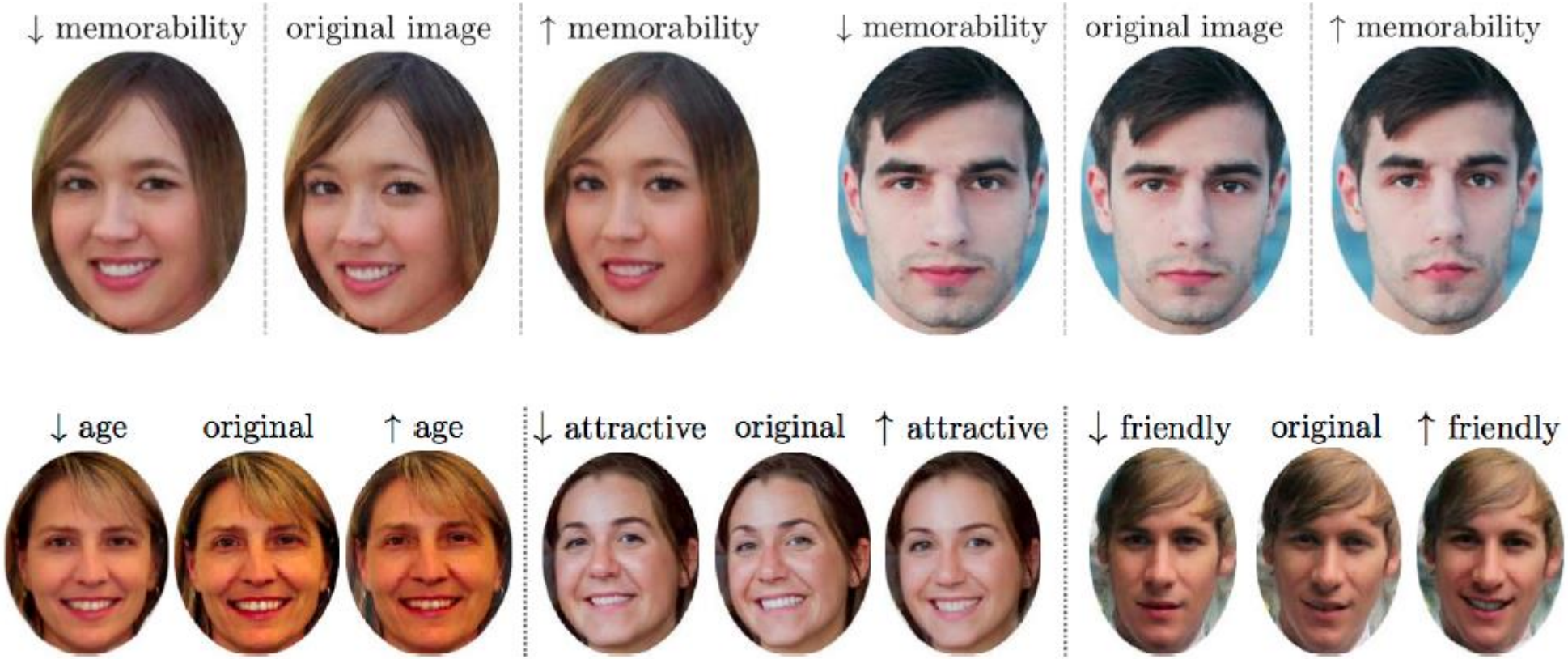
http://www.rekognition.com



Prosojnice: S. Fidler

Zakaj računalniški vid?

- Manipulacija obrazov

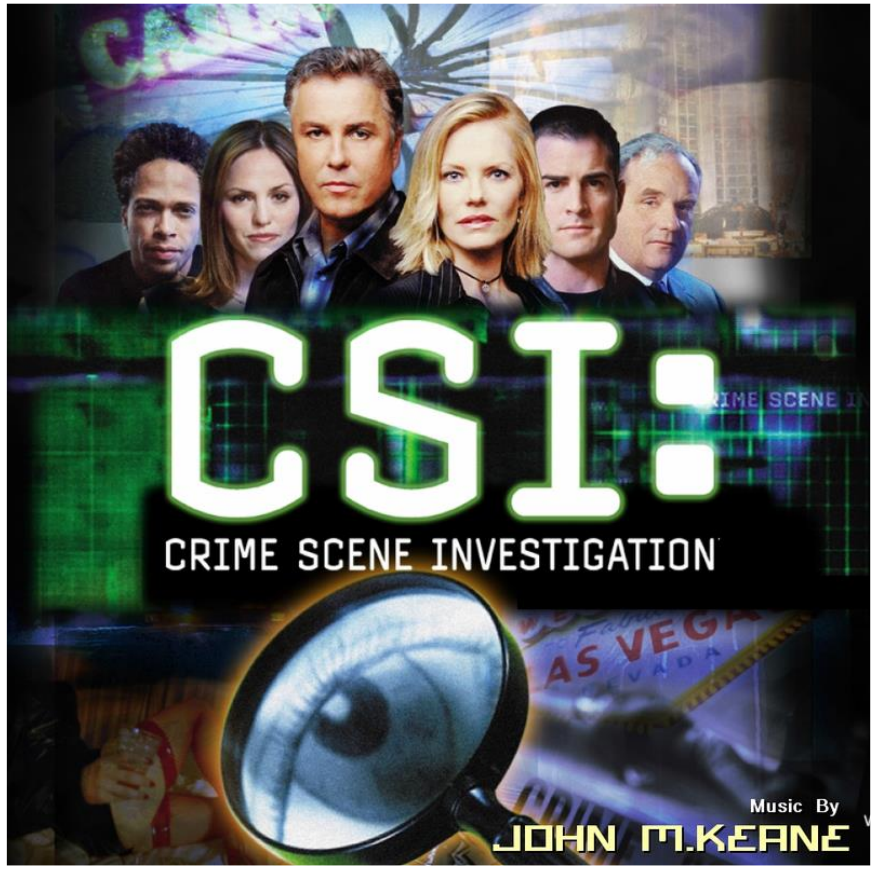


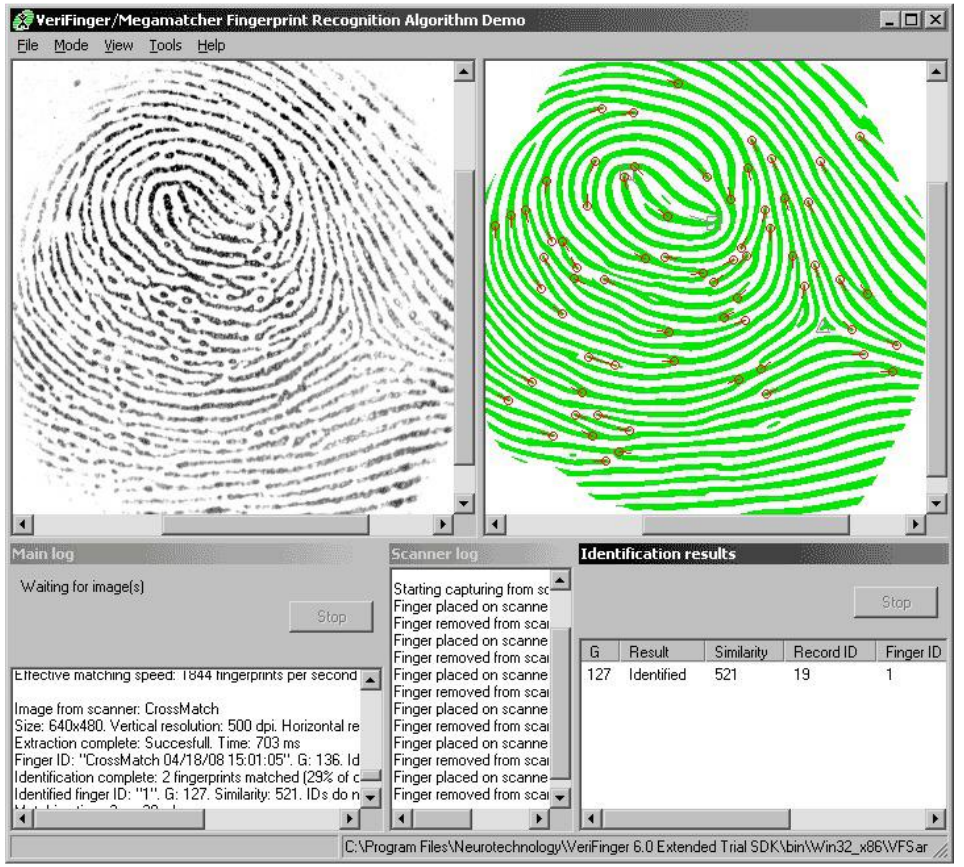
Khosla, Bainbridge, Oliva, Torralba, Modifying the Memorability of Face Photographs, ICCV 2013

Prosojnice: S. Fidler

Zakaj računalniški vid?

- Razpoznava prstnih odtisov





Vir: S. Lazebnik

Zakaj računalniški vid?

- Računalniške igre (detekcija položaja)

Vir: Microsoft Kinect

Računalniški vid

Kako se je začelo?

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group — July 7, 1966
Vision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

Vir: A. Torralba

Prosojnice: S. Fidler

Računalniški vid

Čez 50 let ...

Cyclist

Rank	Method	Setting	Code	Moderate	Easy	Hard	Runtime	Environment	Compare
1	pAUC			**38.03 %**	**51.62 %**	**33.38 %**	60 s	1 core @ 2.5 Ghz (Matlab + C/C++)	
S. Paisitkriangkrai, C. Shen and A. Hengel: Efficient pedestrian detection by directly optimizing the partial area under the ROC curve. ICCV 2013.									
2	DPM-C8B1			29.04 %	43.49 %	26.20 %	15 s	4 cores @ 2.5 Ghz (Matlab + C/C++)	
Anonymous submission									
3	LSVM-MDPM-us		code	29.88 %	38.84 %	27.31 %	10 s	4 cores @ 3.0 Ghz (C/C++)	
P. Felzenszwalb, R. Girshick, D. McAllester and D. Ramanan: Object Detection with Discriminatively Trained Part-Based Models. PAMI 2010.									

Car

Rank	Method	Setting	Code	Moderate	Easy	Hard	Runtime	Environment	Compare
1	SubCat			66.32 %	**81.94 %**	51.10 %	0.3 s	6 cores @ 2.5 Ghz (Matlab + C/C++)	
E. Ohn-Bar and M. Trivedi: Fast and Robust Object Detection Using Visual Subcategories. Computer Vision and Pattern Recognition Workshops Mobile Vision 2014.									
2	AOG		code	67.03 %	80.26 %	**55.60 %**	3 s	4 cores @ 2.5 Ghz (Matlab)	
B. Li, T. Wu and S. Zhu: Integrating Context and Occlusion for Car Detection by Hierarchical And-Or Model. ECCV 2014.									
3	SubCat-NoOcc			58.91 %	79.90 %	44.81 %	0.3 s	6 cores @ 2.5 Ghz (Matlab + C/C++)	
Anonymous submission									

Prosojnice: S. Fidler

Računalniški vid

Čez 50 let ...

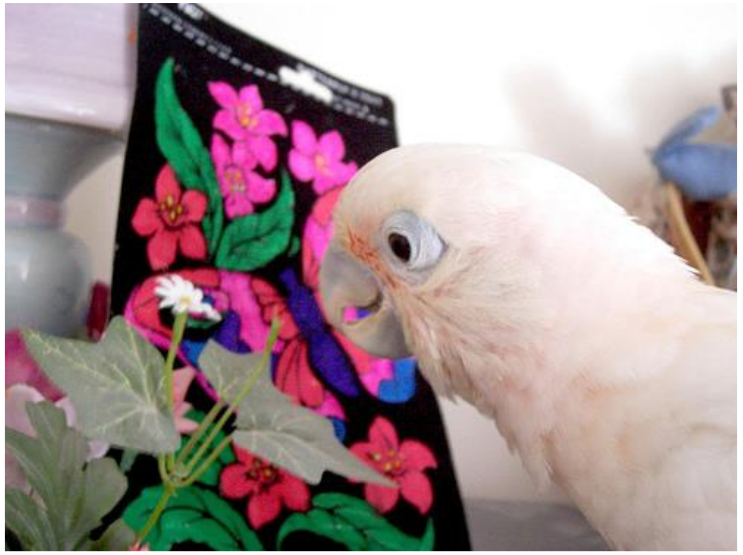
	mean	aero plane	bicycle	bird	boat	bottle	bus	car	cat	chair	cow	dining table	dog	horse	motor bike	person	potted plant	sheep	sofa	train	tv/ monitor	submission date
Feature Edit	56.4	74.8	69.2	55.7	41.9	36.1	64.7	62.3	69.5	31.3	53.3	43.7	69.9	64.0	71.8	60.5	32.7	63.0	44.1	63.6	56.6	2014-Sep-04
R-CNN (bbox reg)	53.7	71.8	65.8	53.0	36.8	35.9	59.7	60.0	69.9	27.9	50.6	41.4	70.0	62.0	69.0	58.1	29.5	59.4	39.3	61.2	52.4	2014-Mar-13
R-CNN	50.2	67.1	64.1	46.7	32.0	30.5	56.4	57.2	65.9	27.0	47.3	40.9	66.6	57.8	65.9	53.6	26.7	56.5	38.1	52.8	50.2	2014-Jan-30

Prosojnice: S. Fidler

Računalniški vid

Čez 50 let ...

- Algoritmi delujejo že “precej” dobro
- Še vedno se dogajajo “nerodne” napake

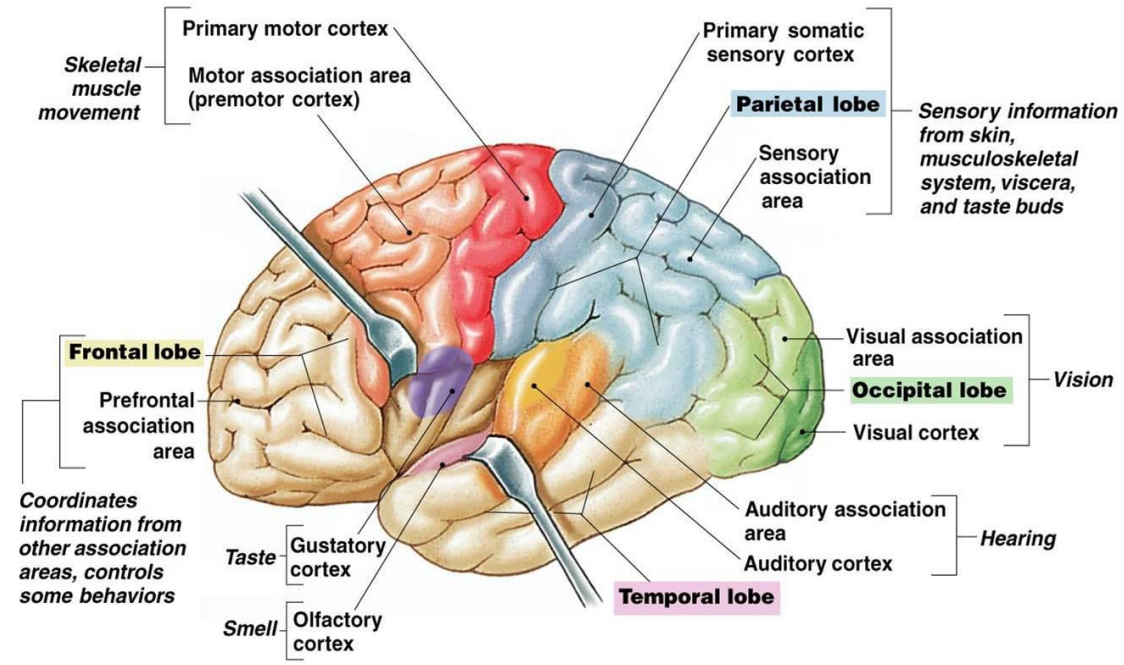


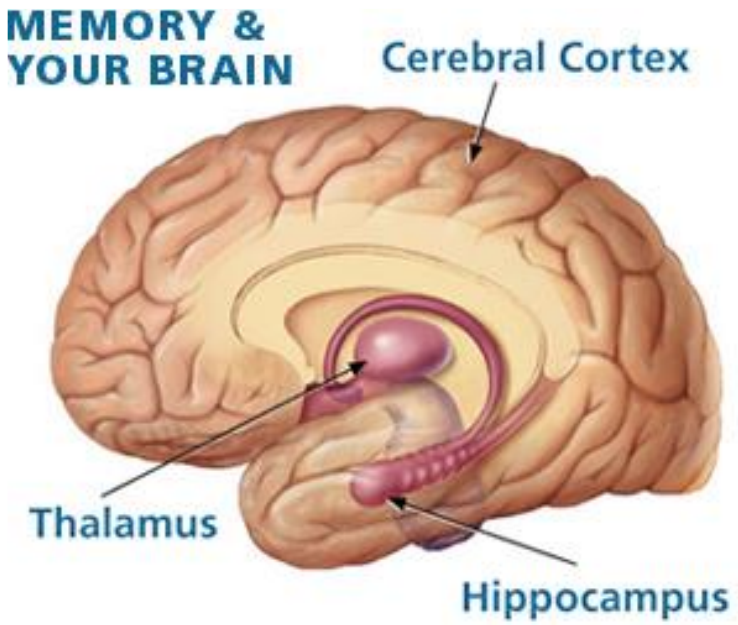
Rožnato območje na desni predstavlja detekcijo osebe na levi sliki

Prosojnice: S. Fidler

Zakaj je računalniški vid “težek”?

- Pol možganske skorje pri primatih je namenjene procesiranju vizualne informacije





Prosojnice: S. Fidler

Zakaj je računalniški vid “težek”?

- Velika količina podatkov
  - Slika vsebuje stotisoče (milijone) pikslov.
- Znotraj ene kategorije imamo lahko veliko variabilnost
  - Tole je pes:

Vir: R. Urtasun

Prosojnice: S. Fidler

Zakaj je računalniški vid “težek”?

- Veliko kategorij

Biederman, 1987

Vir: R. Urtasun

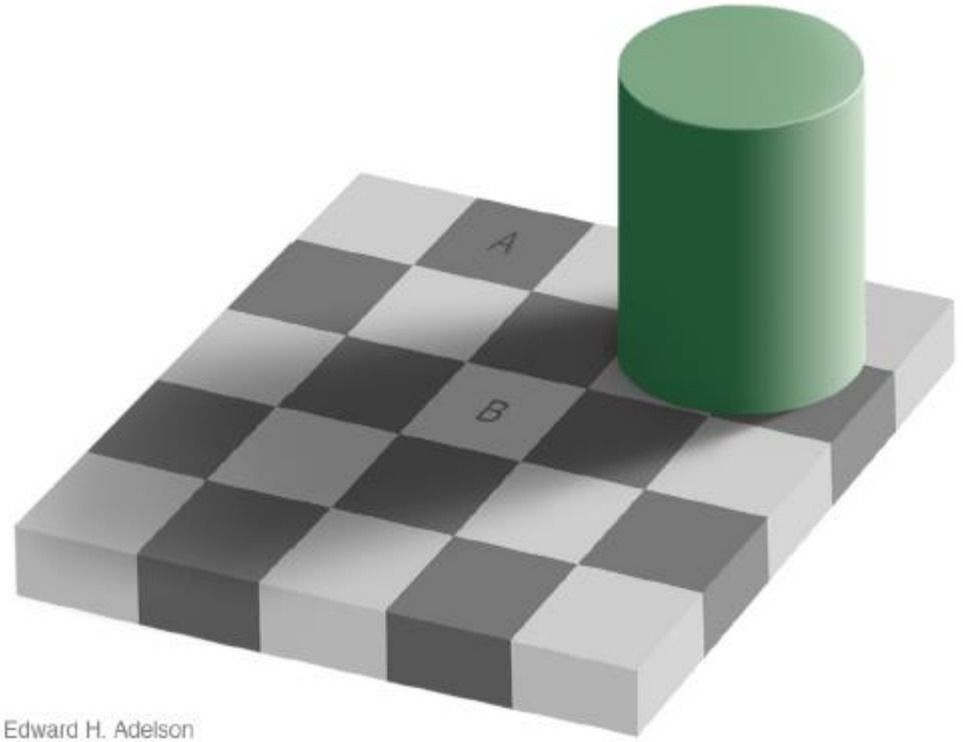
Kako dobri smo ljudje?

- Deluje dokaj dobro...
- ...lahko pa ga prelisičimo

Prosojnice: S. Fidler

Kako dobri smo ljudje?

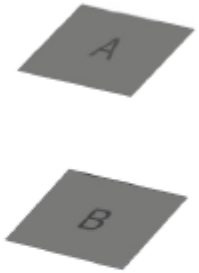
- Kateri kvadrat je svetlejši, A ali B?



Prosojnice: S. Fidler

Kako dobri smo ljudje?

- Kateri kvadrat je svetlejši, A ali B?

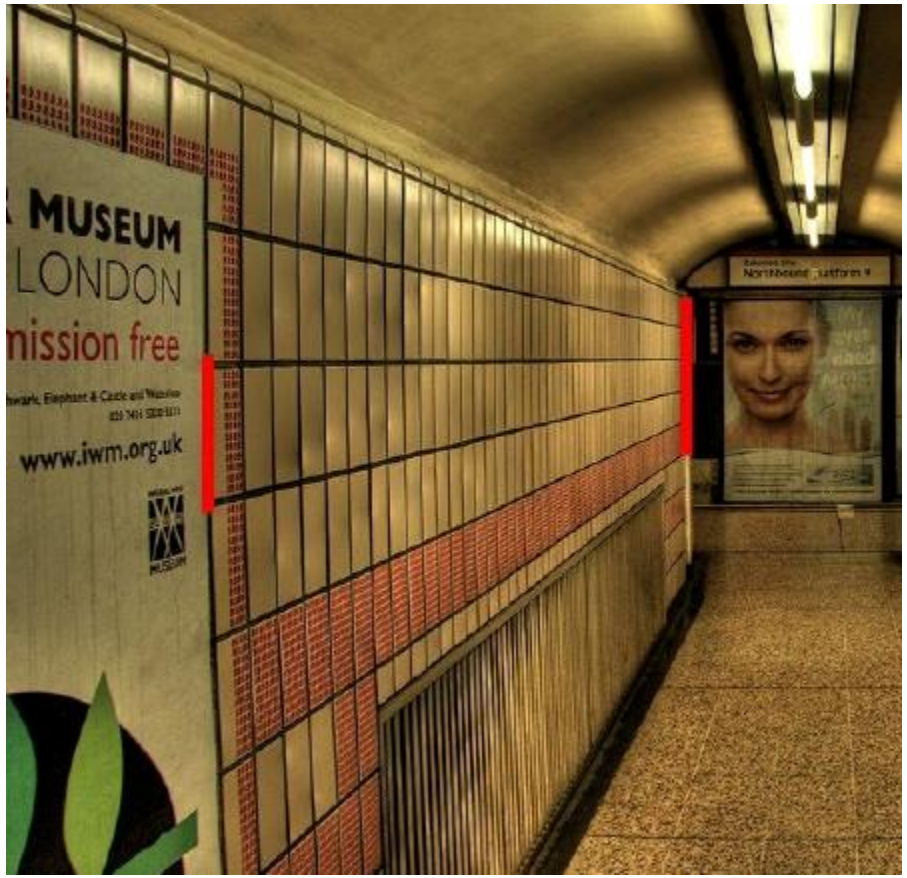


Edward H. Adelson

Prosojnice: S. Fidler

Kako dobri smo ljudje?

- Katera rdeča črta je daljša?



2006 Walt Anthony

Kako dobri smo ljudje?

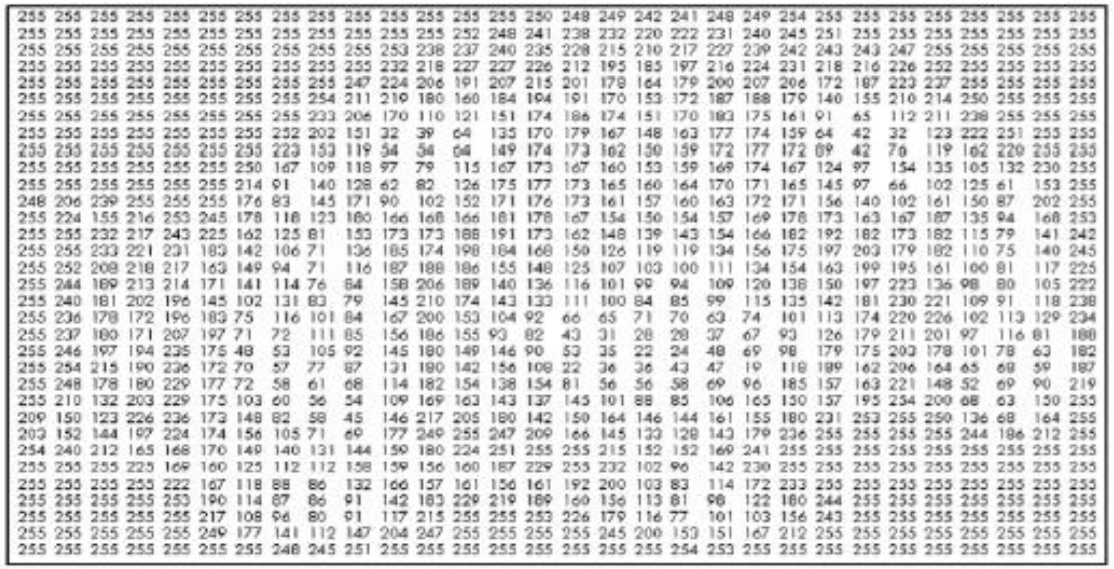
- Katera rdeča črta je daljša?

2006 Walt Anthony

Prosojnice: S. Fidler

Slike in filtri

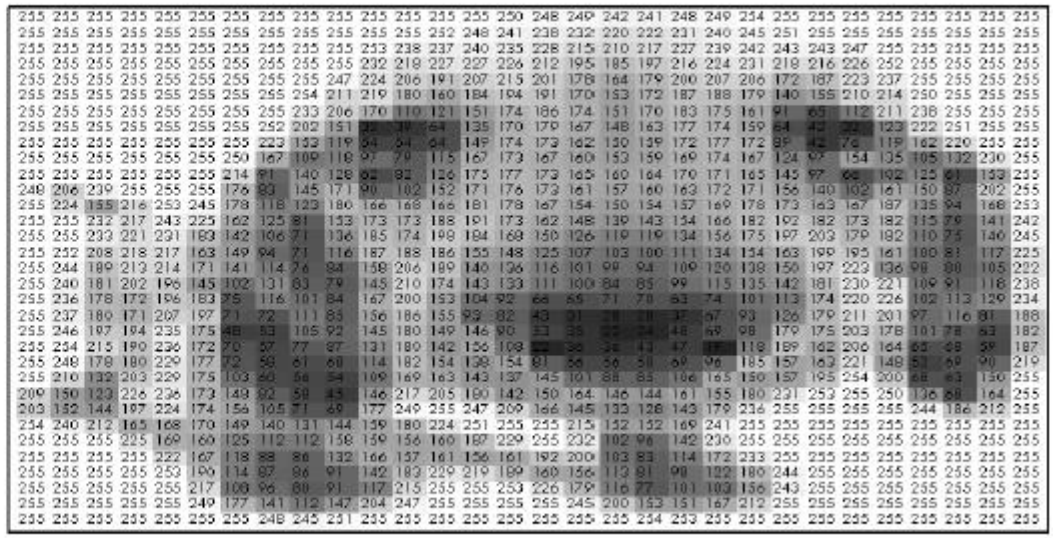
- Kako predstavimo sliko?
  - Slika je matrika I z elementi, ki so cela števila (običajno iz intervala [0, 255])
  - Elementu $I(i, j)$, $i = 1, \ldots, m$, $j = 1, \ldots, n$ rečemo *intenziteta*



Prosojnice: S. Fidler

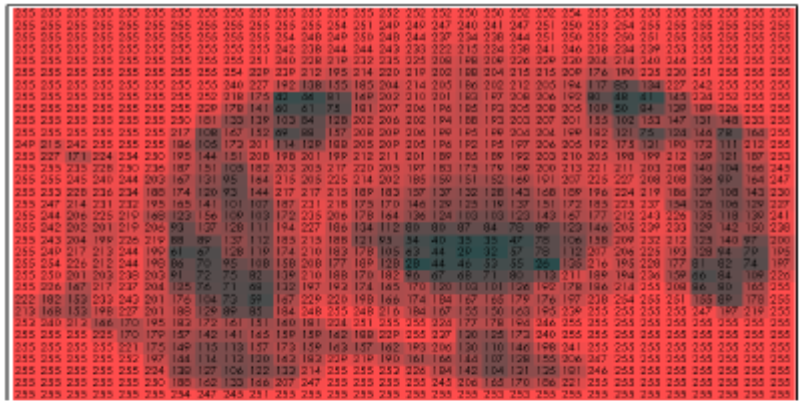
Slike in filtri

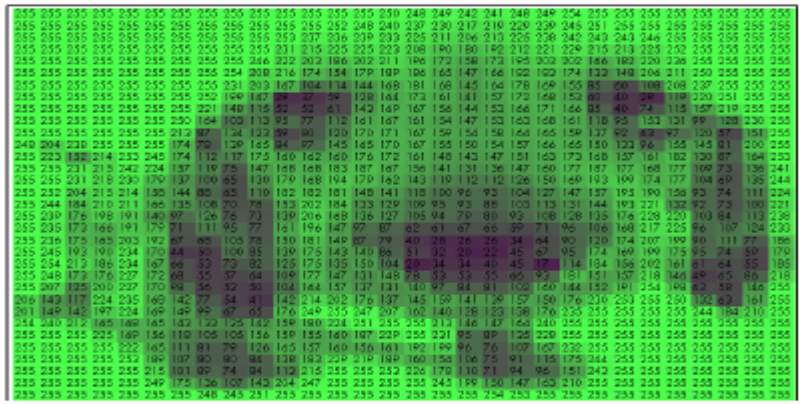
- Kako predstavimo sliko?
  - Slika je lahko sivinska (grayscale)

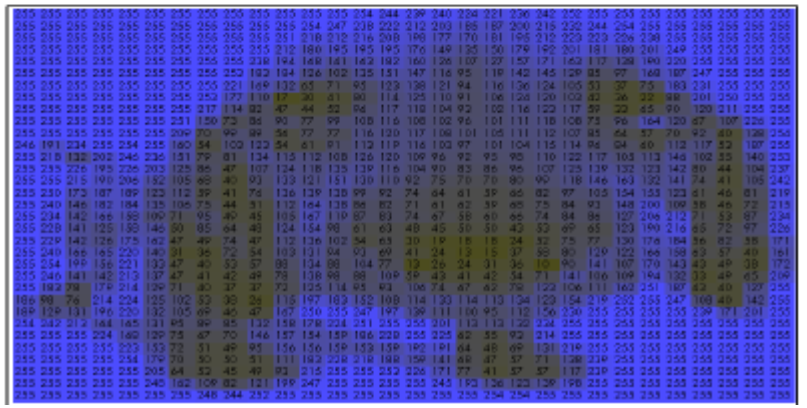


Slike in filtri

- Kako predstavimo sliko?
  - Slika je lahko sivinska (grayscale)
  - Ali barvna. V tem primeru imamo tri matrike – za vsako barvno komponento (rdeča, zelena, modra) po eno matriko







Prosojnice: S. F

Slike in filtri

- Slika lahko razumemo kot funkcijo $I: \mathbb{R}^2 \rightarrow \mathbb{R}$
- Na sliki lahko izvajamo operacije

$I(i,j)$ $J(i,j) = I(i,j) + 50$ $J(i,j) = I(i,-j)$ $I(i,j) \cdot (I(i,j) < 250)$

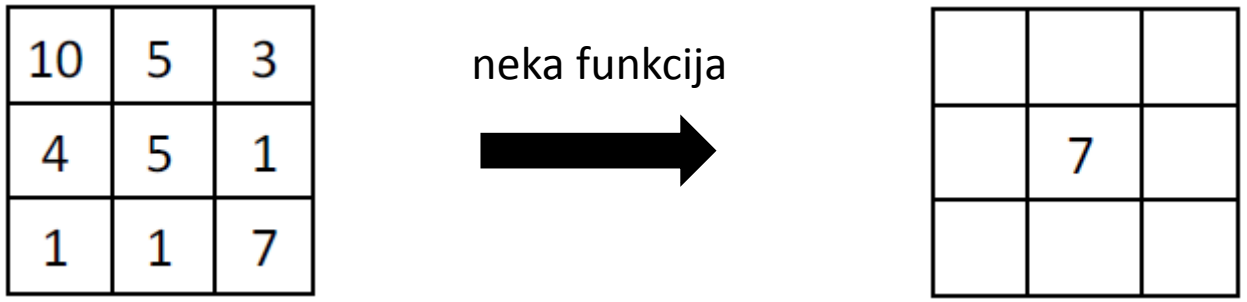
Posebni primeri operacij nad slikami so *korelacije* in *konvolucije* (linearni filtri)

Slike in filtri – motivacija

- Odstranjevanje šuma

Slike in filtri – motivacija

- Spremeni vsak piksel tako, da upoštevaš intenzitete sosednjih pikslov.

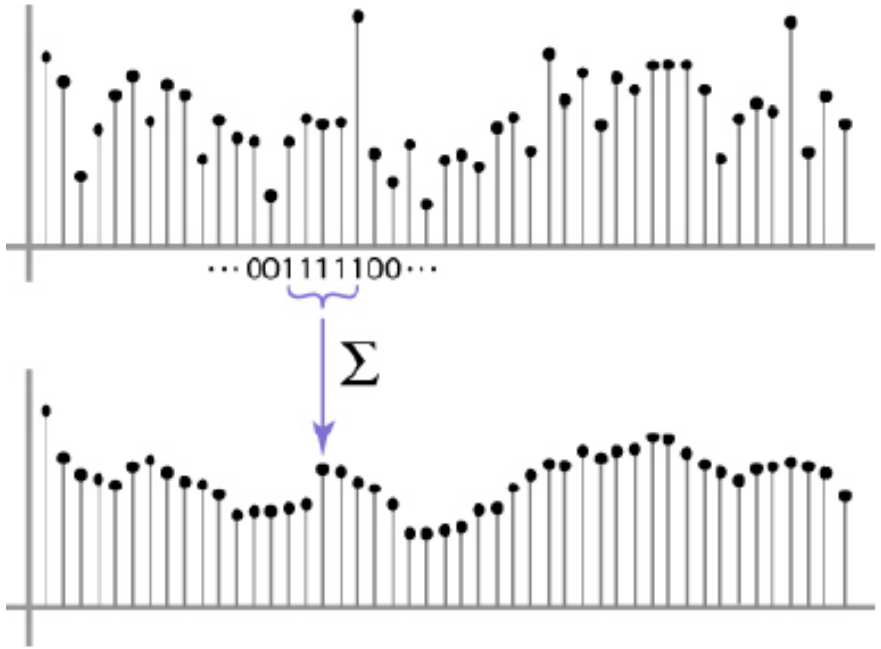


Slike in filtri

- Filtriranje uporabljamo za
  - iskanje vzorcev
  - spreminjanje slike (odstranjevanje šuma, ostrenje, spreminjanje velikosti, ...)
  - iskanje značilnih področij (iskanje teksure, robov, ...)

Slike in filtri – primer

- Odstranjevanje šuma
  - Najenostavnejši način: zamenjaj vsak pixel (intenziteto) s povprečjem intenzitet njegovih sosedov.
  - V 1D: $[1, 1, 1, 1, 1]/5$

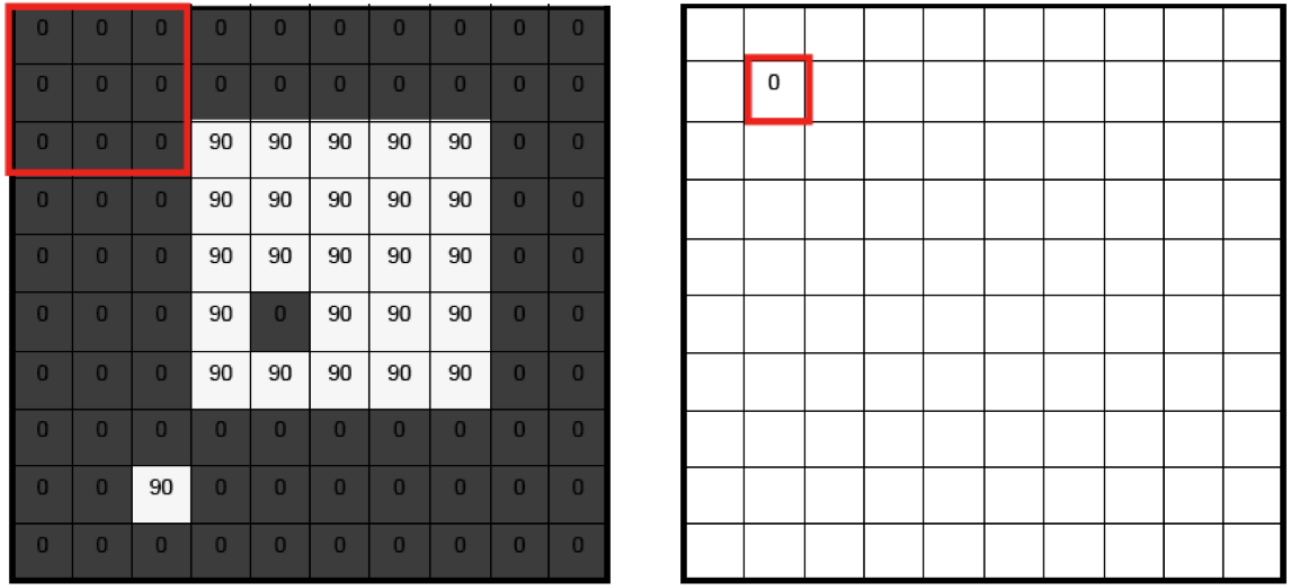


Vir: S. Marschner

Slike in filtri – primer

- Povprečje sosedov v 2D: $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} /9$

$I(i,j)$ $G(i,j)$



Vir: S. Seitz

Slike in filtri – primer

- Povprečje sosedov

$I(i,j)$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$G(i,j)$

	0	10							

Vir: S. Seitz

Slike in filtri – primer

- Povprečje sosedov

$I(i,j)$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$G(i,j)$

	0	10	20						

Vir: S. Seitz

Slike in filtri – primer

- Povprečje sosedov

$I(i,j)$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$G(i,j)$

	0	10	20	30					

Vir: S. Seitz

Slike in filtri – primer

- Povprečje sosedov

$I(i,j)$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$G(i,j)$

	0	10	20	30	30	30	20	10	
	0	20	40	60	60	60	40	20	
	0	30	60	90	90	90	60	30	
	0	30	50	80	80	90	60	30	
	0	30	50	80	80	90	60	30	
	0	20	30	50	50	60	40	20	
	10	20	30	30	30	30	20	10	
	10	10	10	0	0	0	0	0	

Vir: S. Seitz

Slike in filtri – korelacija

- Povprečje sosedov (formalno)

$$G(i,j) = \frac{1}{(2k+1)^2} \sum_{u=-k}^{k} \sum_{v=-k}^{k} I(i+u, j+v)$$

- V splošnem rečemo uteženi vsoti intezitet sosednjih pikslov *korelacija*, torej

$$G(i,j) = \sum_{u=-k}^{k} \sum_{v=-k}^{k} F(u,v) \cdot I(i+u, j+v)$$

- F rečemo jedro ali maska; elementi $F(u,v)$ so koeficienti filtra
- Označimo $G = F \otimes I$

Slike in filtri – korelacija – primer

I

0	0	0
0	1	0
0	0	0

F

?

Slike in filtri – korelacija – primer

I

0	0	0
0	1	0
0	0	0

F

$F \otimes I$

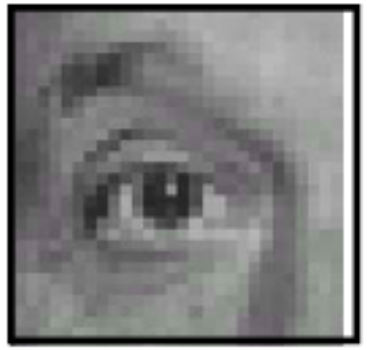
Slike in filtri – korelacija – primer

0	0	0
0	0	1
0	0	0

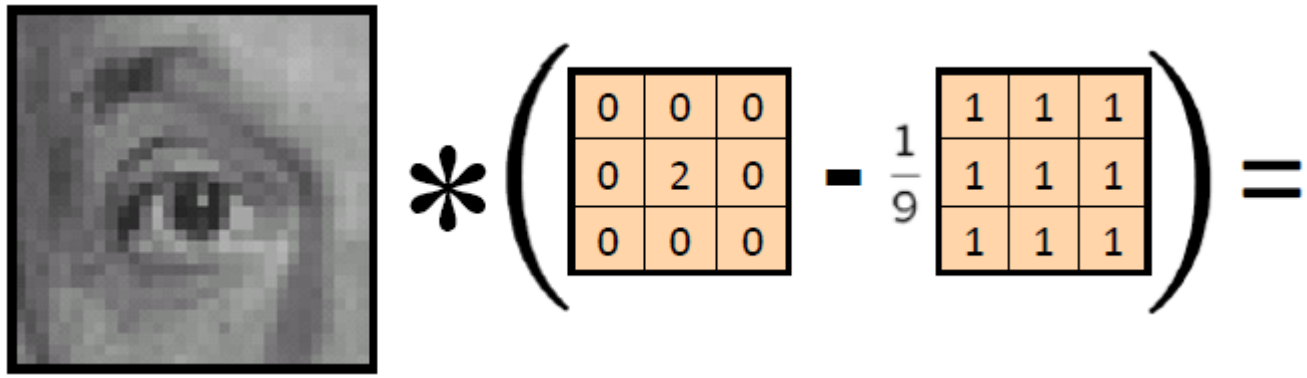
?

Slike in filtri – korelacija – primer

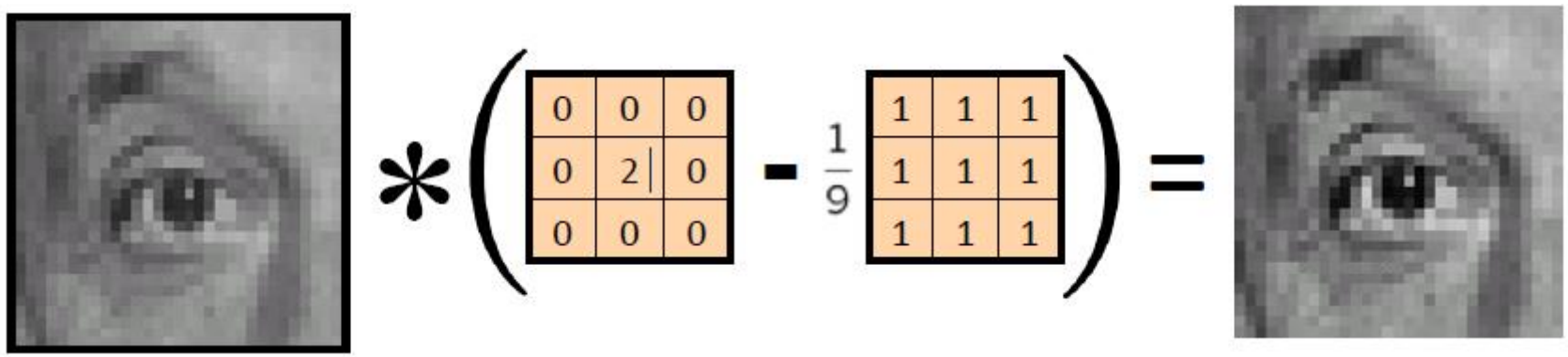
0	0	0
0	0	1
0	0	0



Slike in filtri – korelacija – primer



Slike in filtri – korelacija – primer

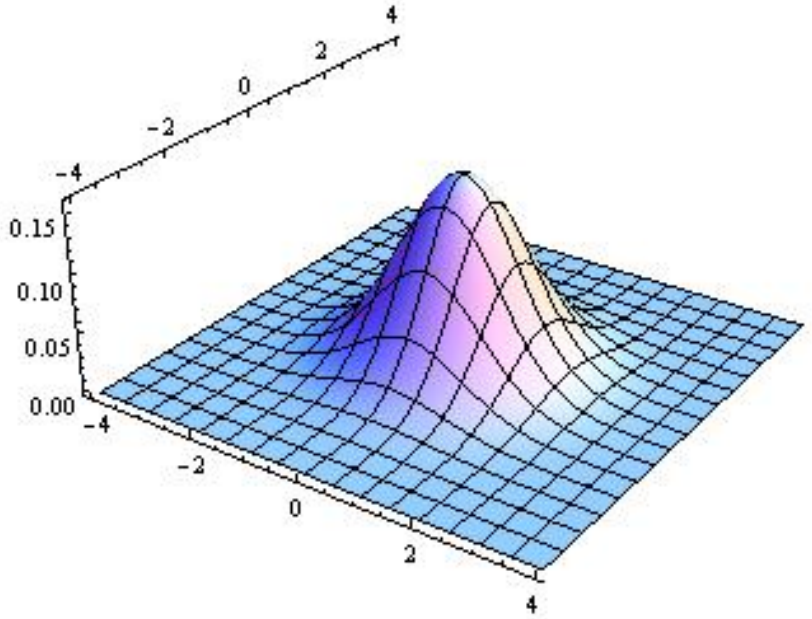


Prosojnice: S. Fidler

Slike in filtri – Gaussov filter

- Jedro Gaussovega filtra je aproksimacija 2D Gaussove funkcije

$$h(u,v) = \frac{1}{2\pi\sigma^2} e^{-(u^2+v^2)/2\sigma^2}$$

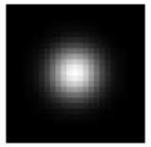


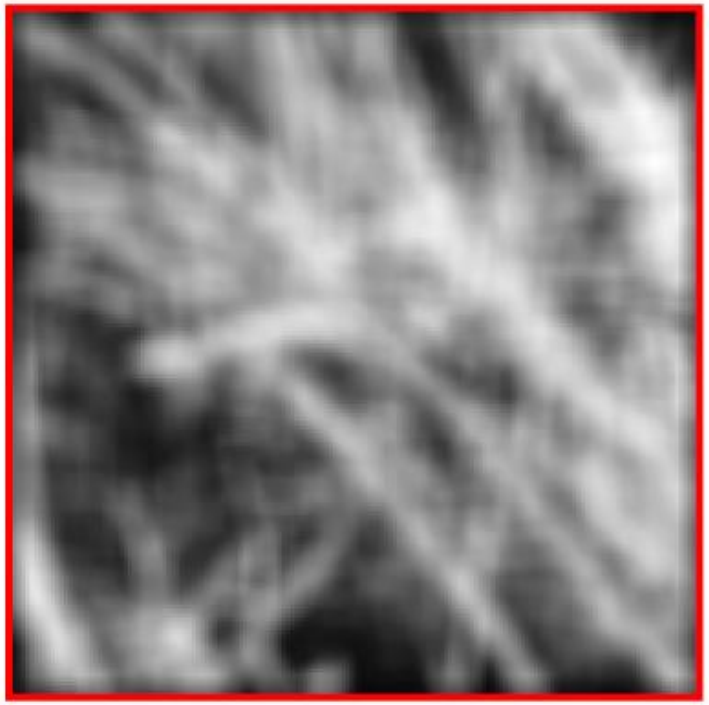
$$\frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

$$F(i,j)$$

Slike in filtri – Gaussov filter

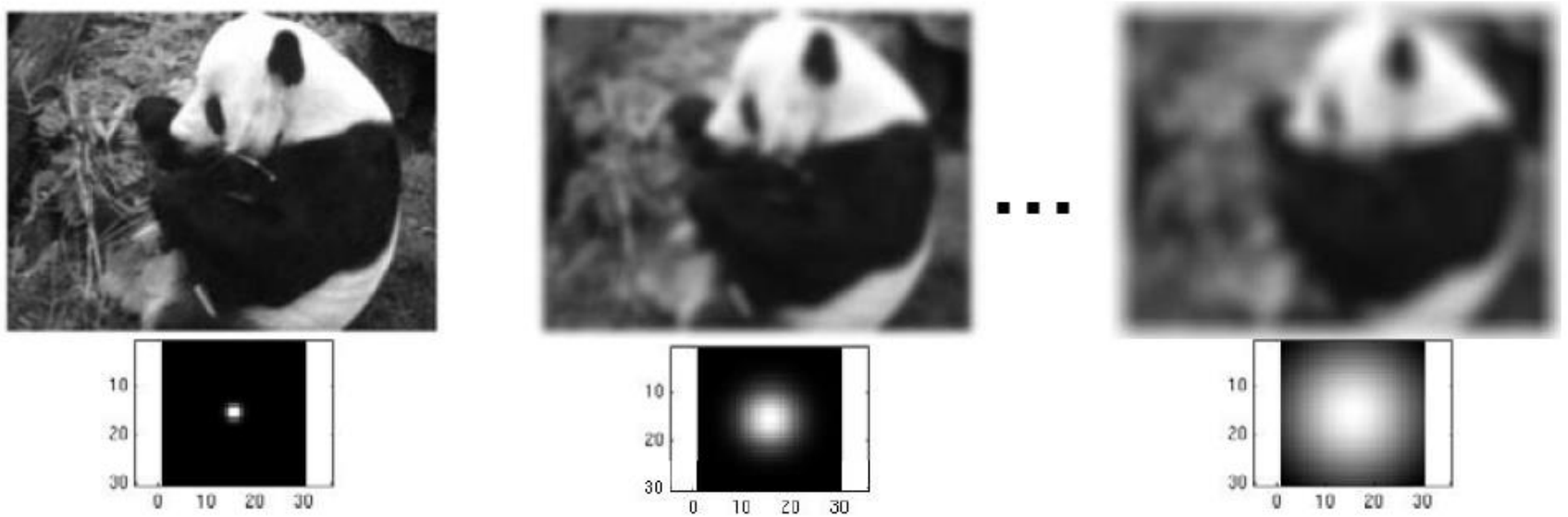
- Glajenje z Gaussovim filtrom





Slike in filtri – Gaussov filter

- Glajenje z Gaussovim filtrom




```
for sigma=1:3:10
   h = fspecial('gaussian', fsize, sigma);
   out = imfilter(im, h);
   imshow(out);
   pause;
end
```

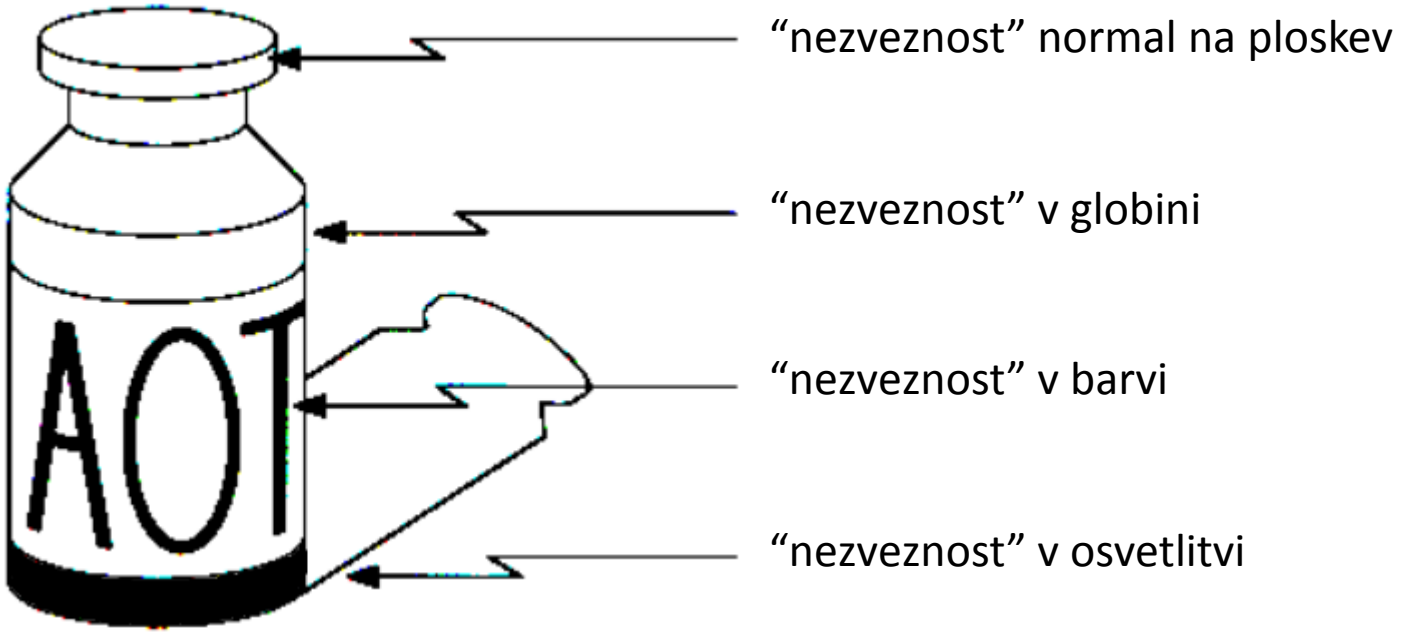
Prosojnice: S. Fidler

Iskanje robov

Prosojnice: S. Fidler

Iskanje robov

- Rob nastane zaradi nenadne spremembe v itenziteti



N. Snavely

Iskanje robov

- Matematično: odvod

$$\frac{\partial f}{\partial x}(x,y) = \lim_{\varepsilon \to 0} \frac{f(x+\varepsilon, y) - f(x,y)}{\varepsilon}$$

- ki ga aproksimiramo z diferenco

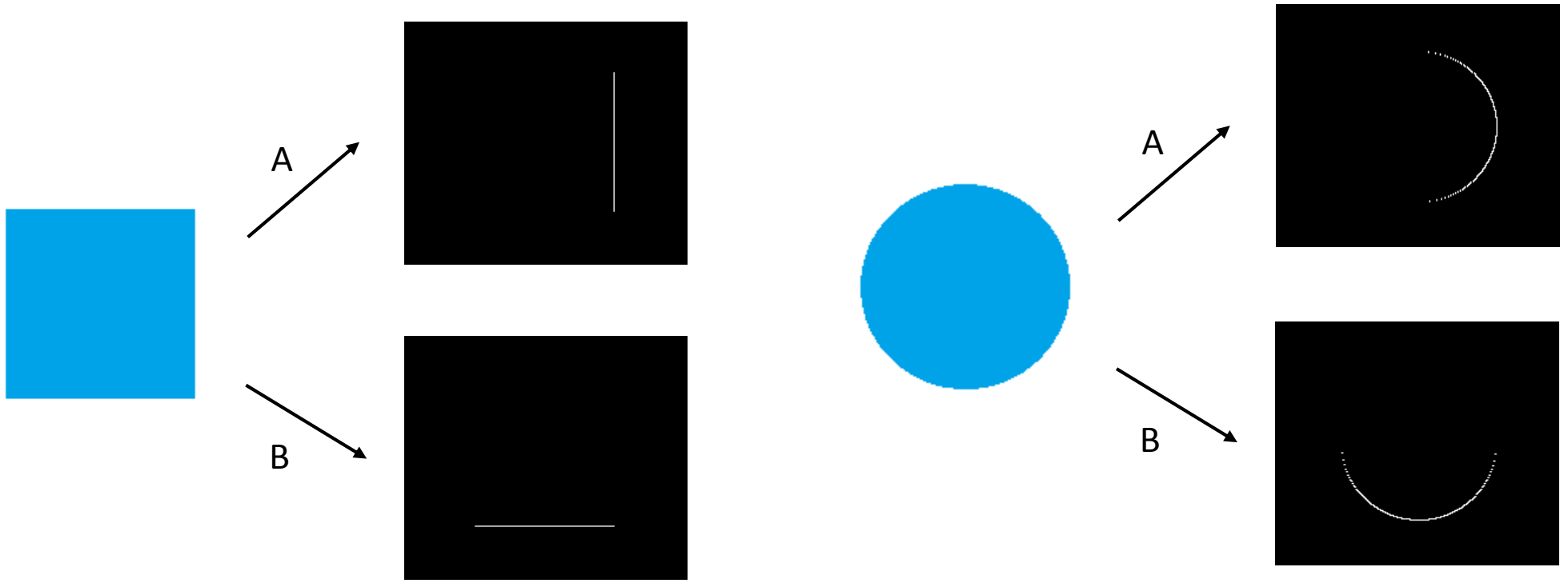
$$\frac{\partial f}{\partial x}(x,y) \approx \frac{f(x+1, y) - f(x,y)}{1}$$

- Ustrezen filter je:

$$\text{A}=\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \text{oziroma B} = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \text{za odvod po } y$$

Prosojnice: S. Fidler

Iskanje robov



Prosojnice: S. Fidler

Iskanje robov

- Filtri za iskanje robov

Prewitt: $M_x =$

-1	0	1
-1	0	1
-1	0	1

; $M_y =$

1	1	1
0	0	0
-1	-1	-1

Sobel: $M_x =$

-1	0	1
-2	0	2
-1	0	1

; $M_y =$

1	2	1
0	0	0
-1	-2	-1

Roberts: $M_x =$

0	1
-1	0

; $M_y =$

1	0
0	-1

Prosojnice: S. Fidler

Značilke

- Značilka (angl. feature) je informacija (manjši del slike), ki je relevantna za določen tip uporabe (recimo, za detekcijo objektov, stikanje slik, ...). Na primer:
  - robovi
  - koti
  - ukrivljenosti
  - …

Prosojnice: S. Fidler

Deskriptorji

- Deskriptor je množica (vektor) značilk ki opisujejo del slike. Želimo, da je:
  - invarianten glede na skalo (velikost), rotacijo, ...
  - ni visoko dimenzionalen (iskanje mora biti hitro)
  - ga lahko učinkovito izračunamo

Prosojnice: S. Fidler

Deskriptorji

- SIFT
- PCA-SIFT
- GLOH
- HOG
- SURF
- DAISY
- LBP
- Shape Context
- Barvni Histogrami

Prosojnice: S. Fidler

Deskriptorji – kako poiščemo človeka?

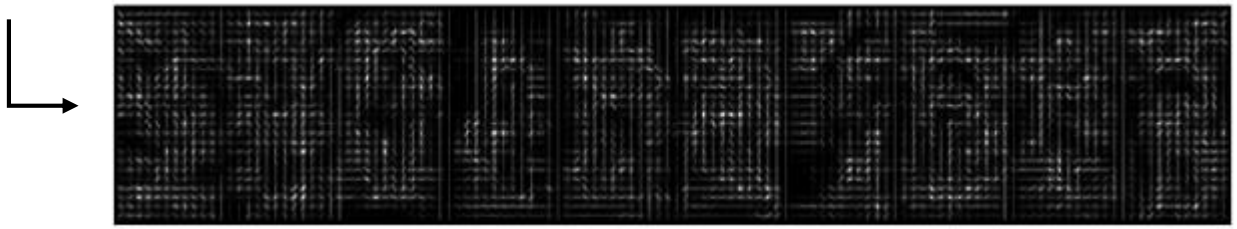
- HOG (zelooo na hitro)
  - 15 x 7 “celic”
  - vsak piksel v celici prispeva utežen “glas” glede na orientacijo roba v tisti točki (9 orientacij)
  - → dobimo 3780 dimenzionalen vektor

Prosojnice: S. Fidler

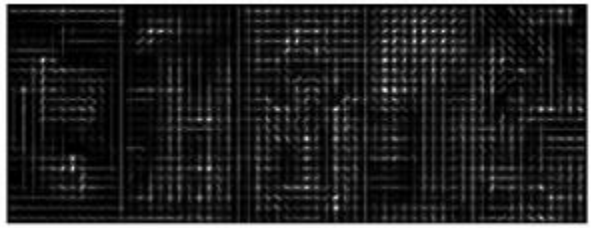
Deskriptorji – kako poiščemo človeka?

- HOG (zelooo na hitro)

Pozitivni primeri:



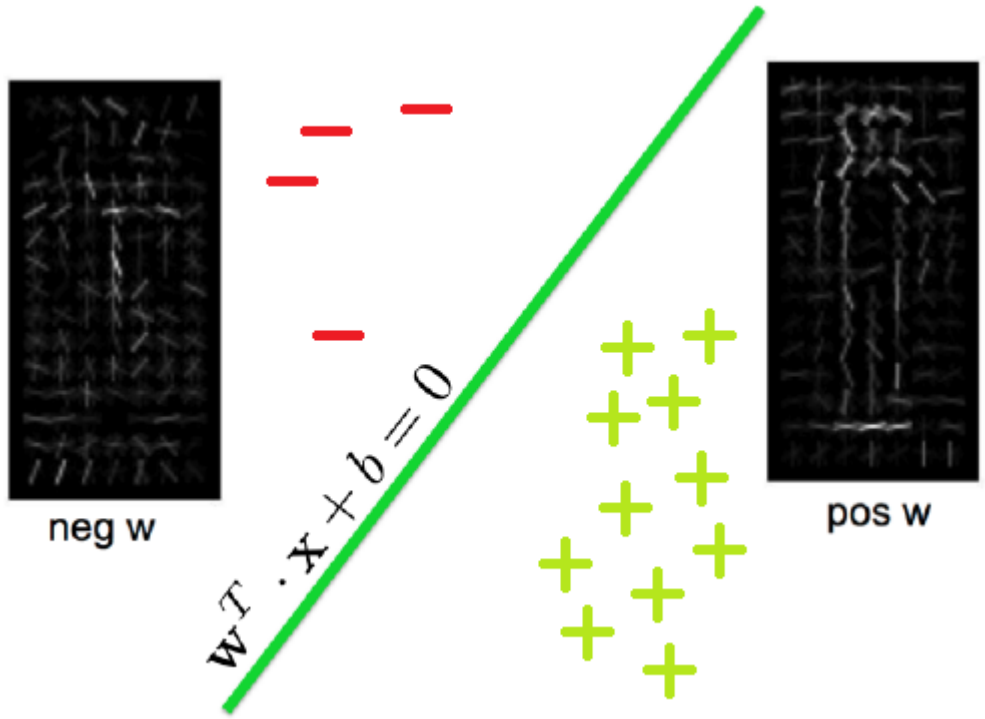
Negativni primeri:



Prosojnice: S. Fidler

Deskriptorji – kako poiščemo človeka?

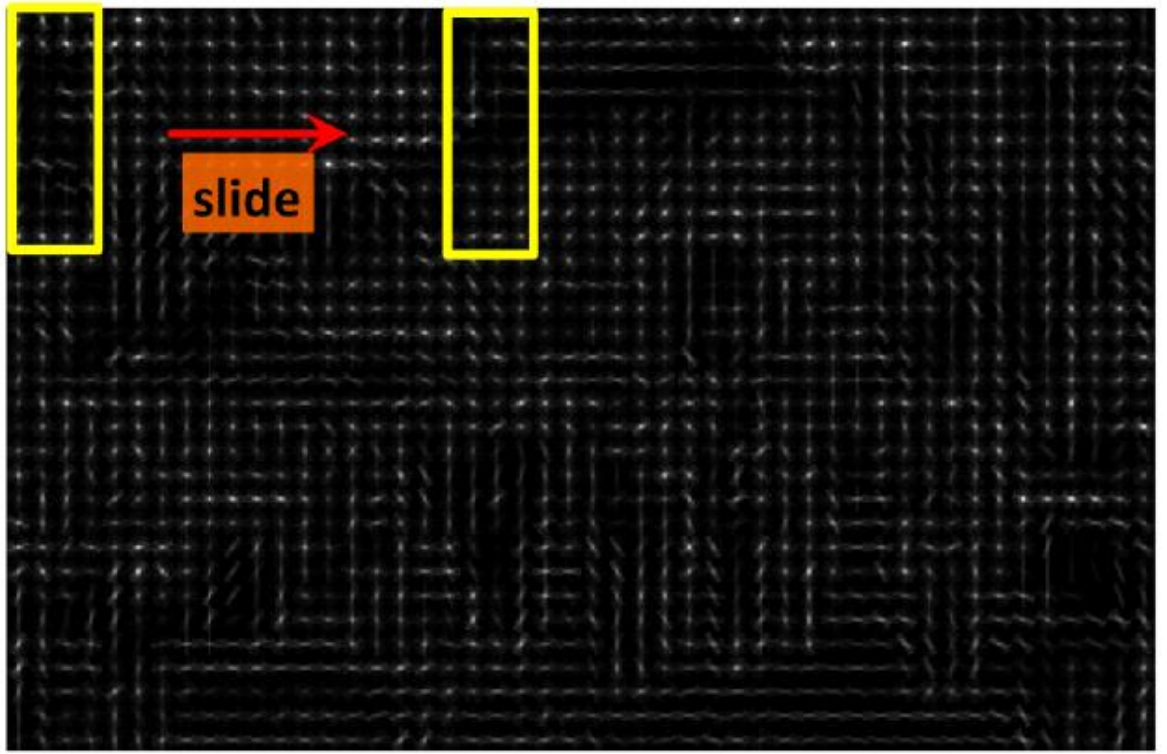
- HOG (zelooo na hitro)
  - Na podlagi vektorjev, ki pripadajo pozitivnim in negativnim primerom naučimo binarni klasifikator
  - Tipično uporabimo SVM (support vector machine)



Prosojnice: S. Fidler

Deskriptorji – kako poiščemo človeka?

- HOG (zelooo na hitro)



Za vsako “okno” (rumeno) izračunamo vektor x in izračunamo vrednost $w^T \cdot x + b$,
t.j. ugotovimo ali pade na pozitivno ali negativno stran (hiper)ravnine.

Prosojnice: S. Fidler

Deskriptorji – kako poiščemo človeka?

- HOG (zelooo na hitro)

Okna, pri katerih so vrednosti $w^T \cdot x + b$ nad nekim pragom (threshold).

Prosojnice: S. Fidler

Deskriptorji – kako poiščemo človeka?

- HOG (zelooo na hitro)

Od prekrivajočih se oken ohranimo tisto z največjo vrednostjo in ...

Prosojnice: S. Fidler

Razpoznavanje objektov – hierarhični pristop

- Sanja Fidler, Marko Boben, Aleš Leonardis: Learning a Hierarchical Compositional Shape Vocabulary for Multi-class Object Representation
- Predstavljeno na različnih konferencah od 2008 do 2012.
- Že rahlo zastarelo ☹

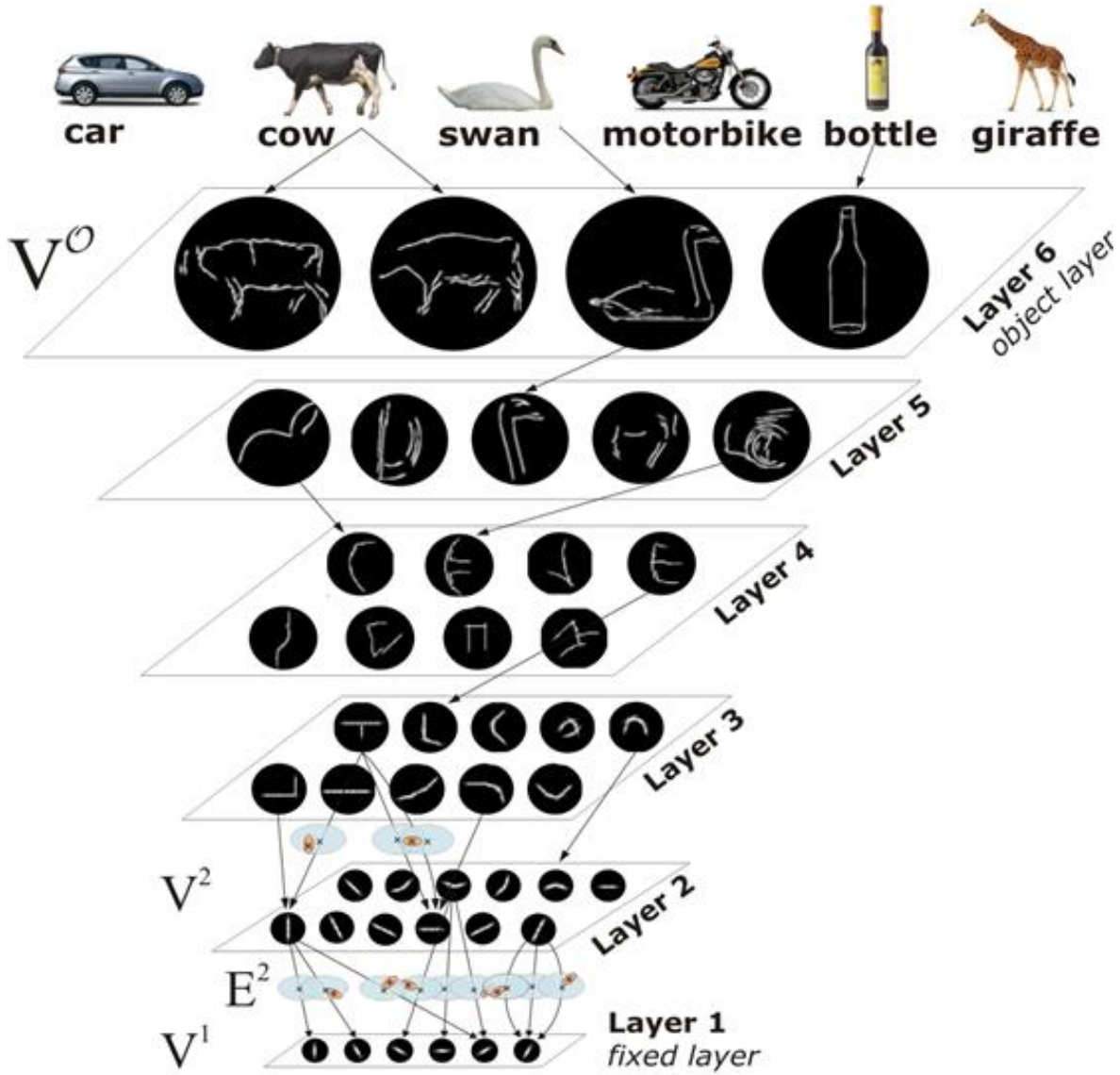
Prosojnice: S. Fidler

Razpoznavanje objektov – hierarhični pristop

- Zgradimo “slovar oblik”
- Najnižji nivo: orientirani robovi (Gaborjevi filtri)
- Vsak element naslednjega nivoja je zgrajen iz elementov prejšnjega nivoja.
- Nivoje se “učimo” na podlagi učnih slik
- Končni nivo je oblika objekta, ki ga učimo.

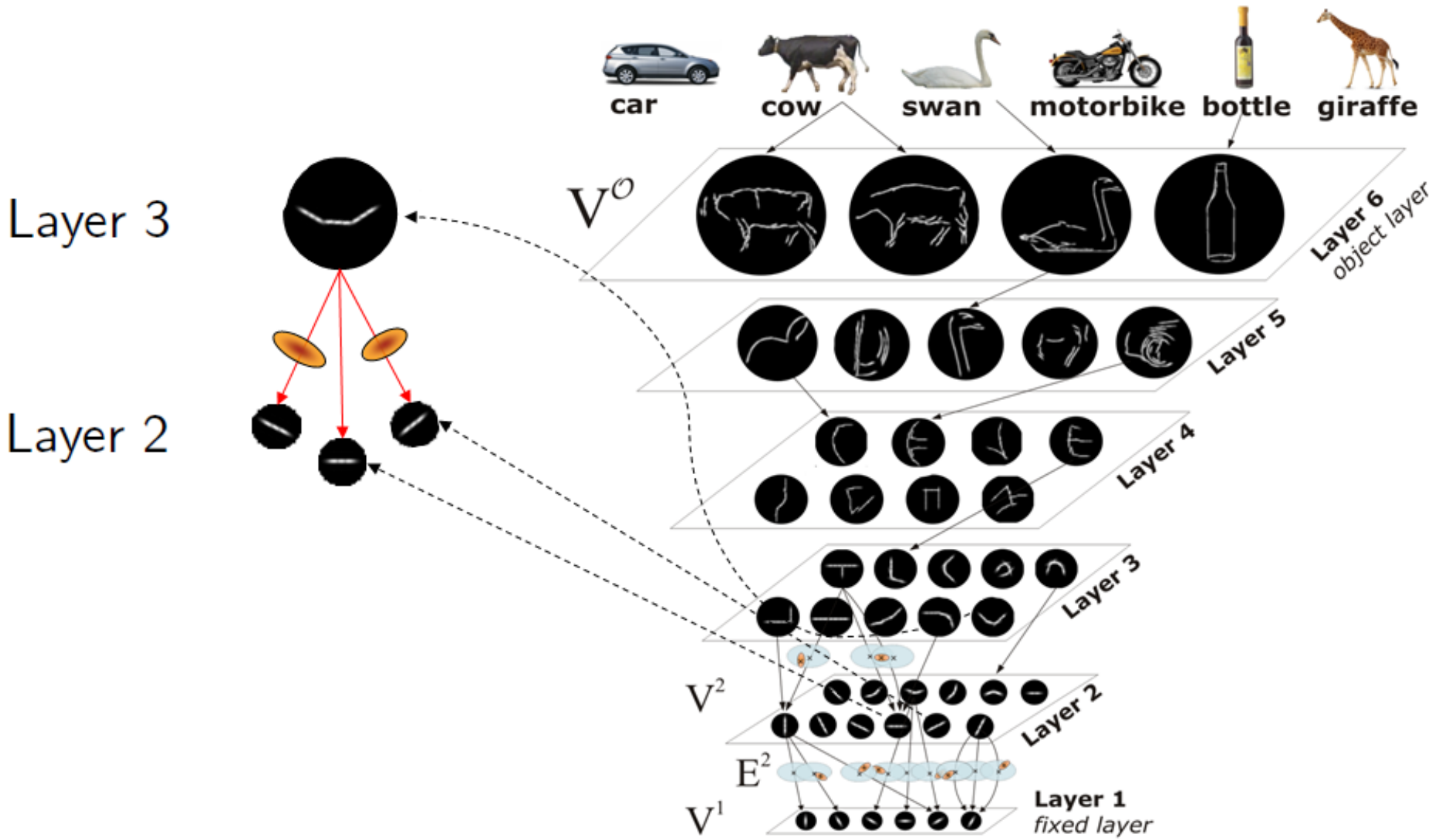
Prosojnice: S. Fidler

Razpoznavanje objektov – hierarhični pristop



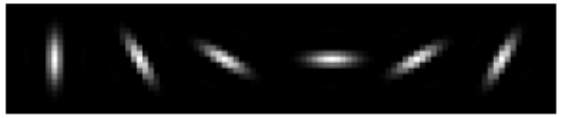
Prosojnice: S. Fidler

Razpoznavanje objektov – hierarhični pristop

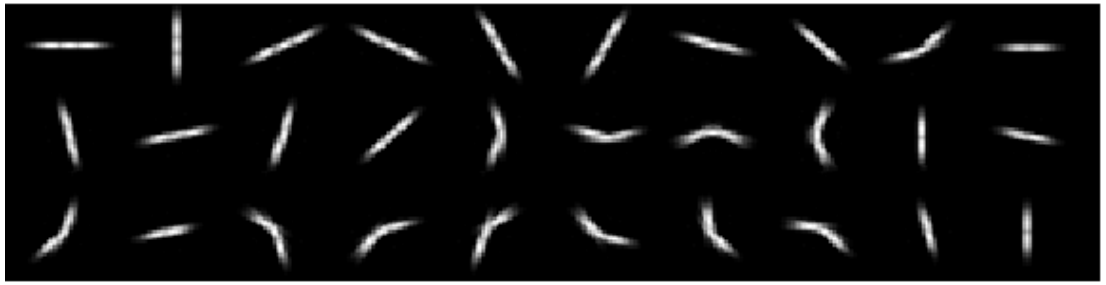


Prosojnice: S. Fidler

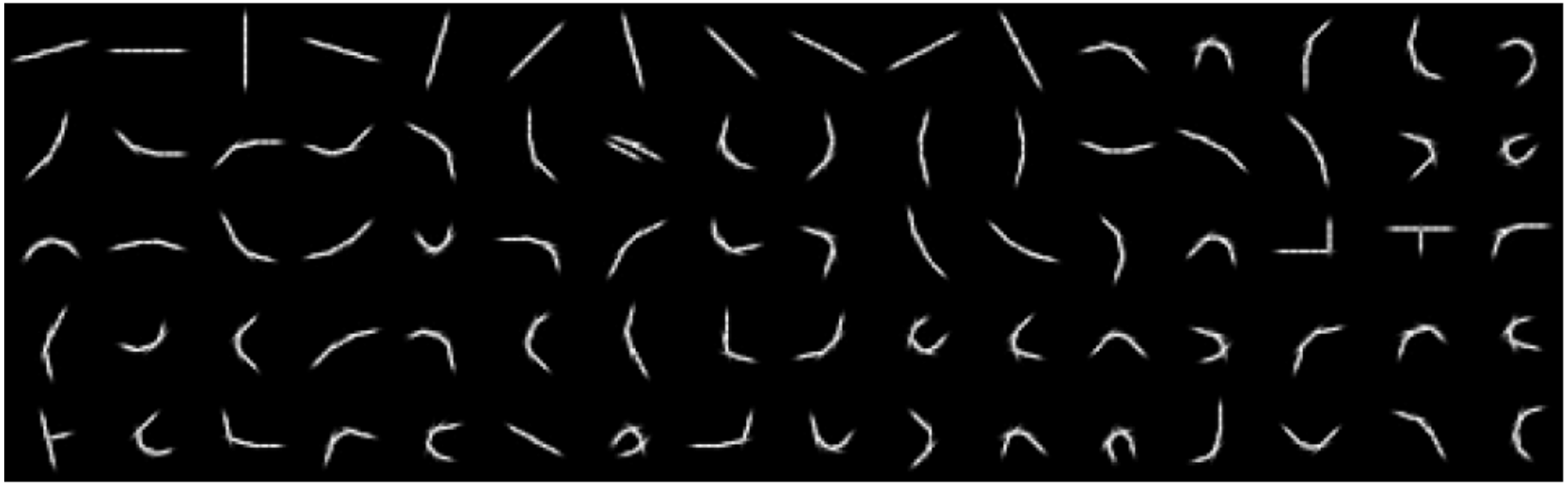
Razpoznavanje objektov – hierarhični pristop



Nivo 1



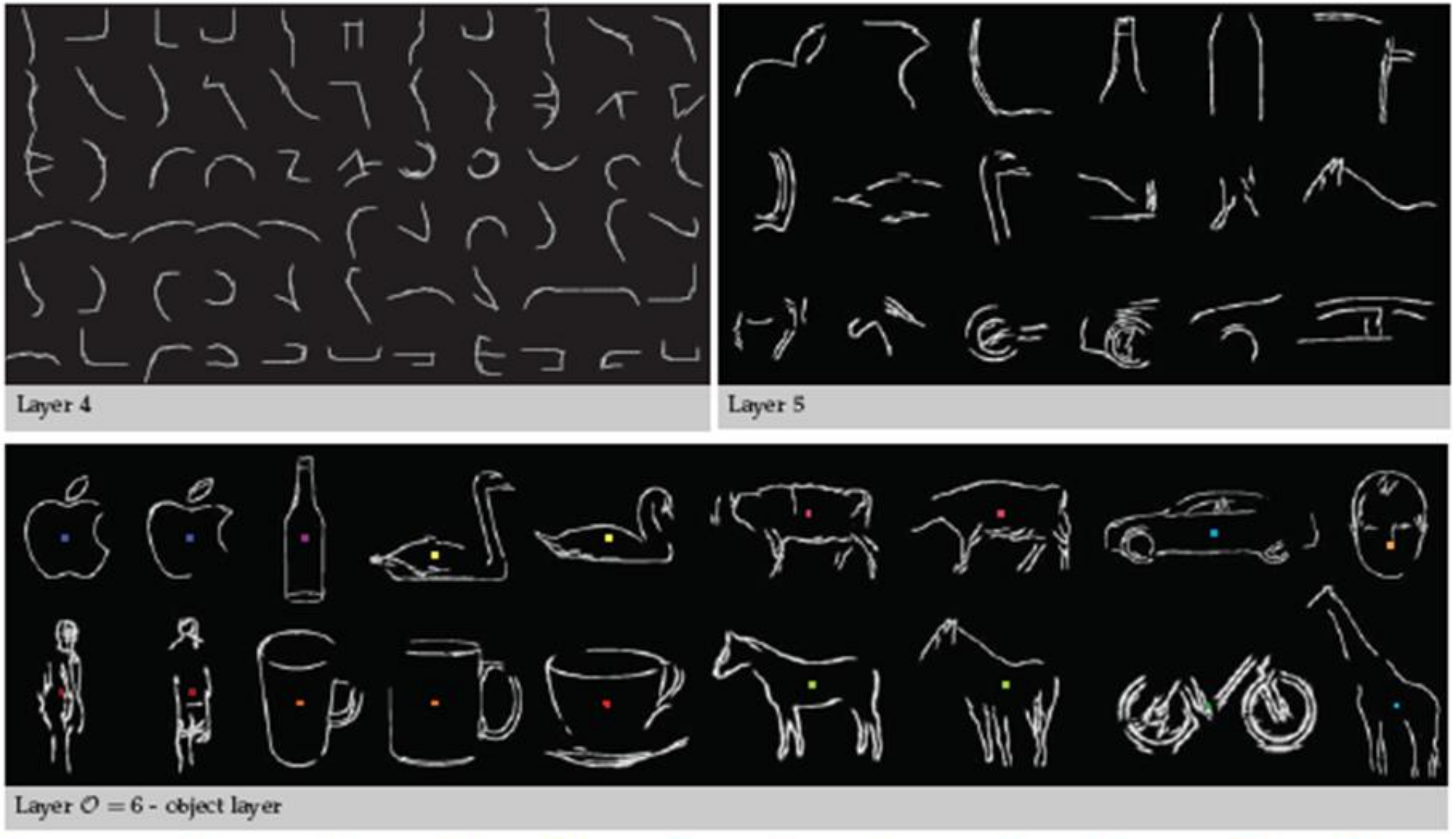
Nivo 2



Nivo 3

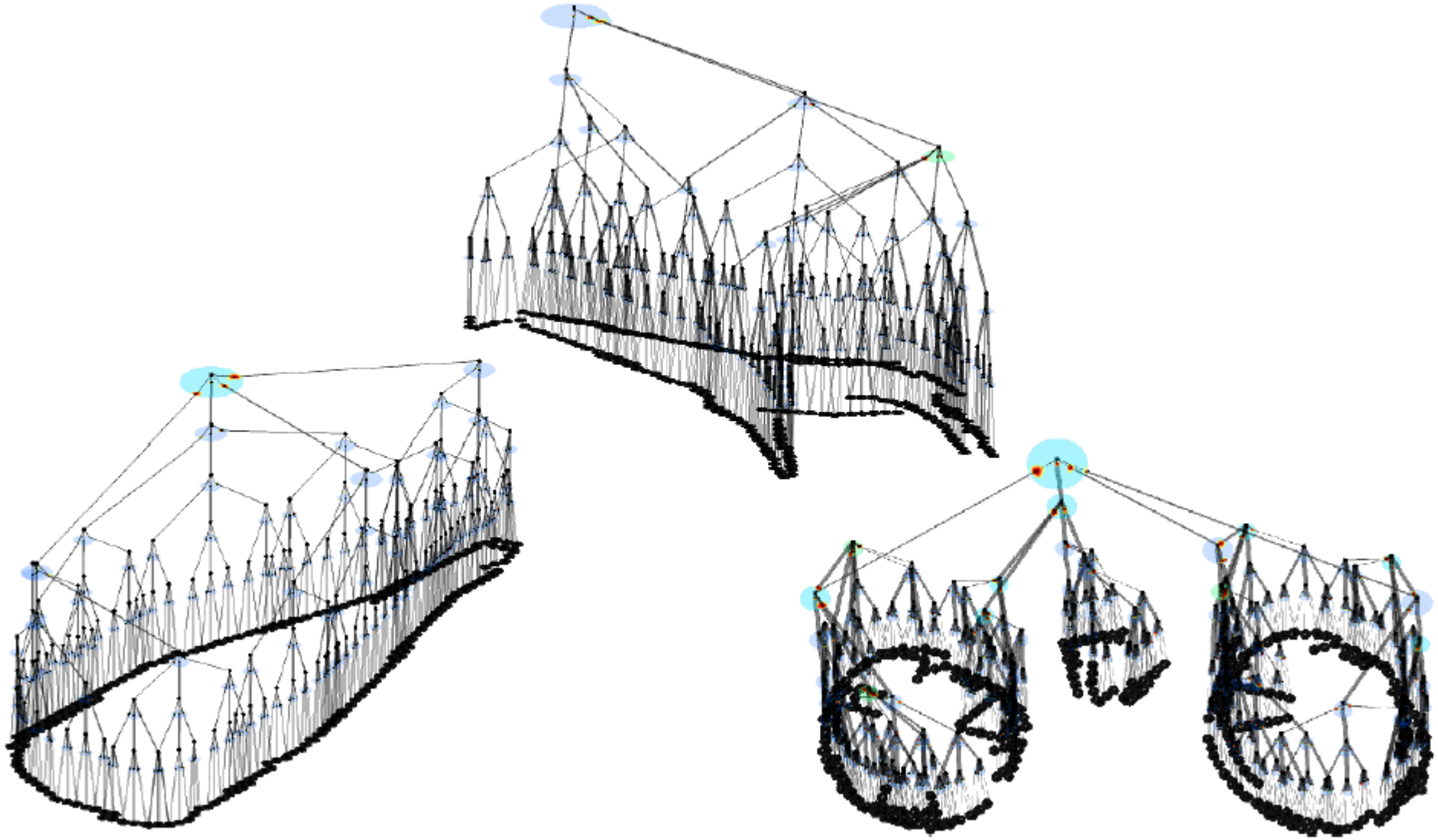
Prosojnice: S. Fidler

Razpoznavanje objektov – hierarhični pristop



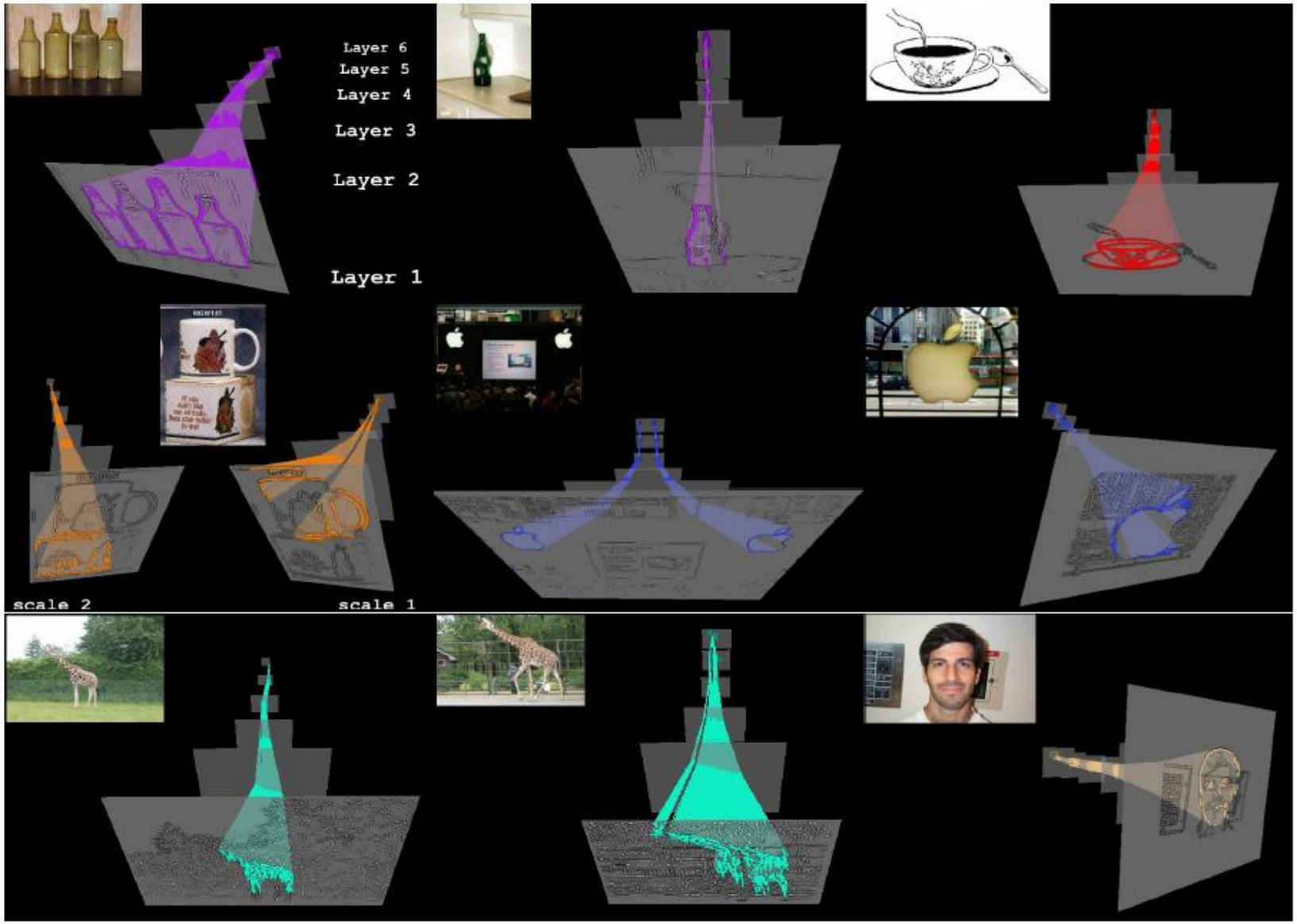
Prosojnice: S. Fidler

Razpoznavanje objektov – hierarhični pristop



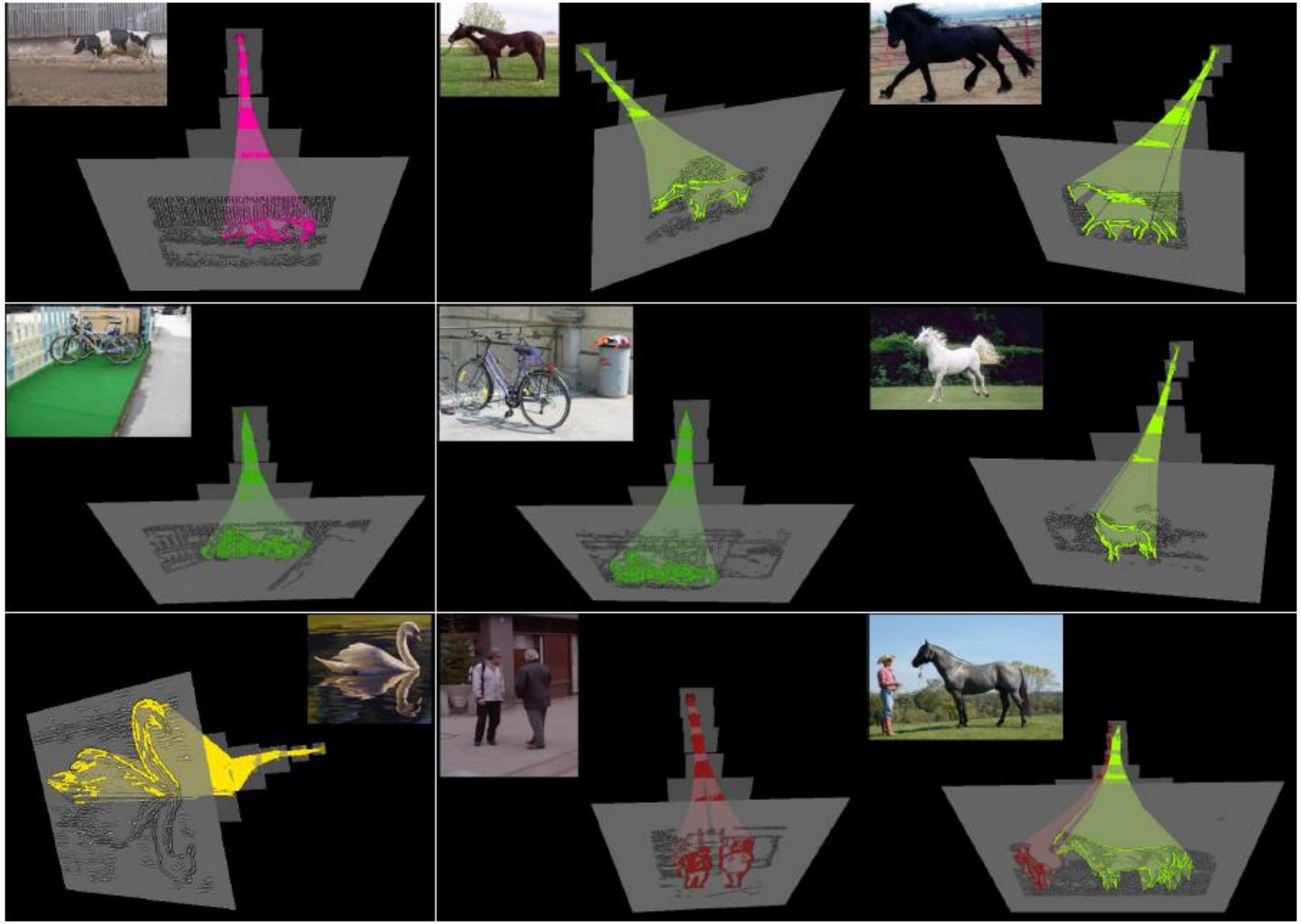
Prosojnice: S. Fidler

Razpoznavanje objektov – hierarhični pristop



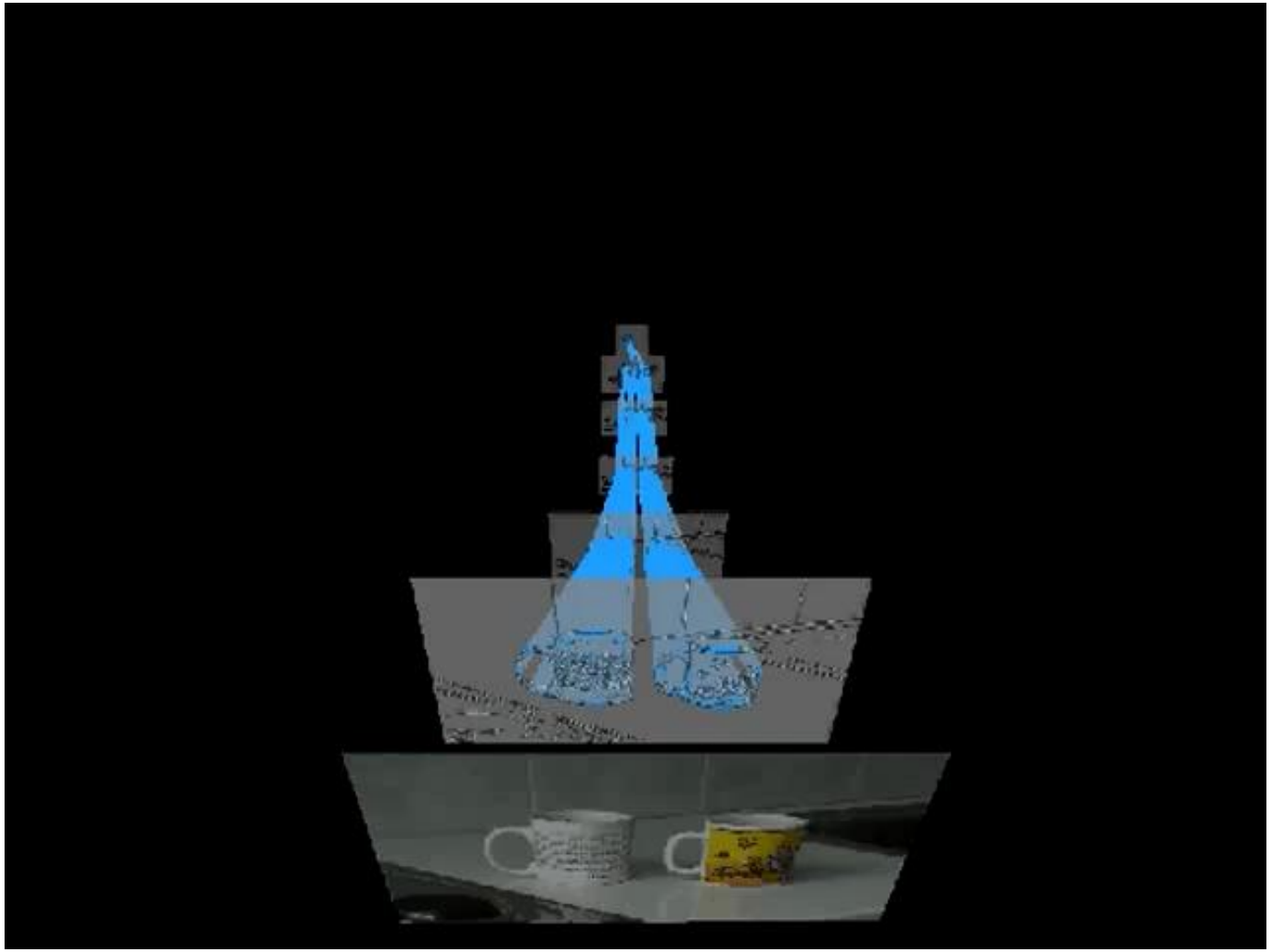
Prosojnice: S. Fidler

Razpoznavanje objektov – hierarhični pristop



Prosojnice: S. Fidler

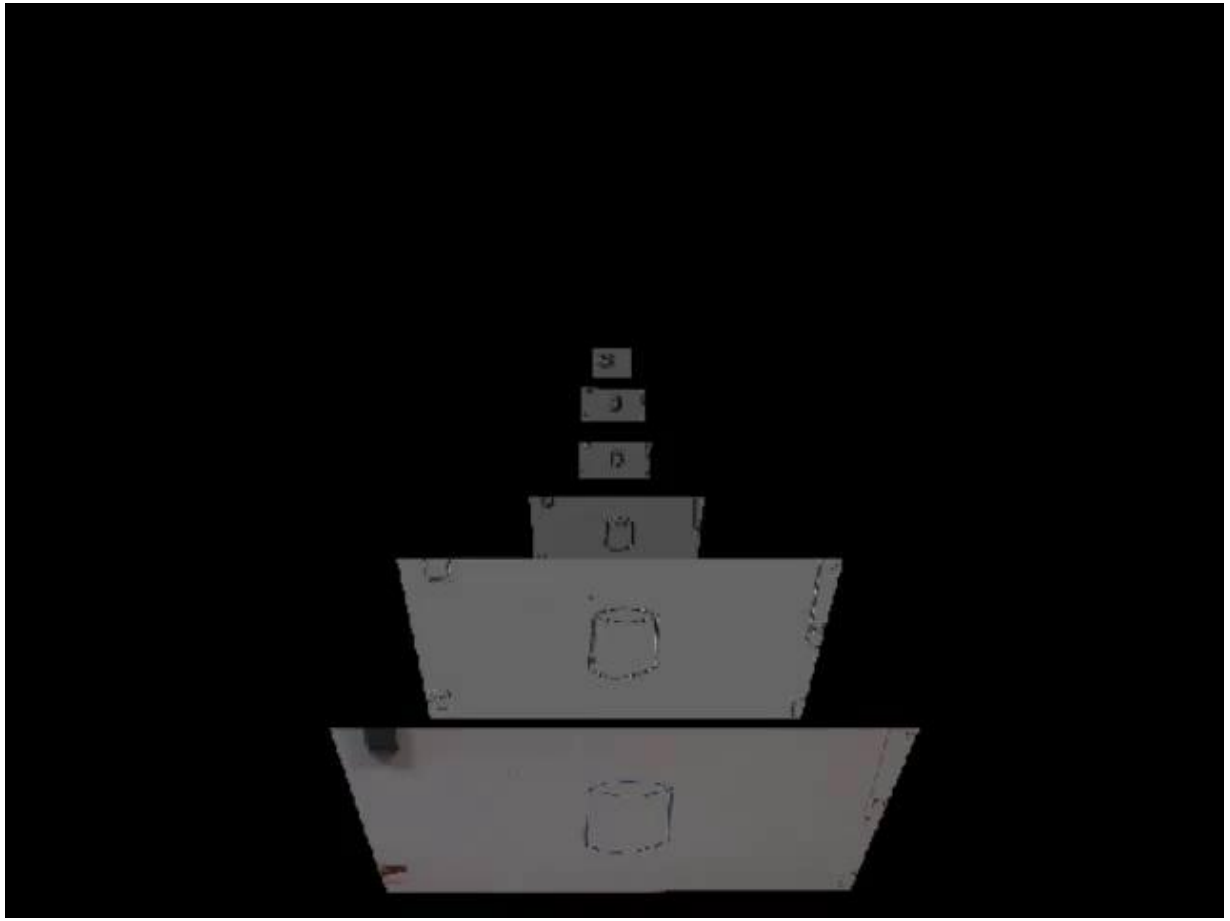
Razpoznavanje objektov – hierarhični pristop



http://www.cs.utoronto.ca/~fidler/slides/CSC420/videos/mugs.mp4

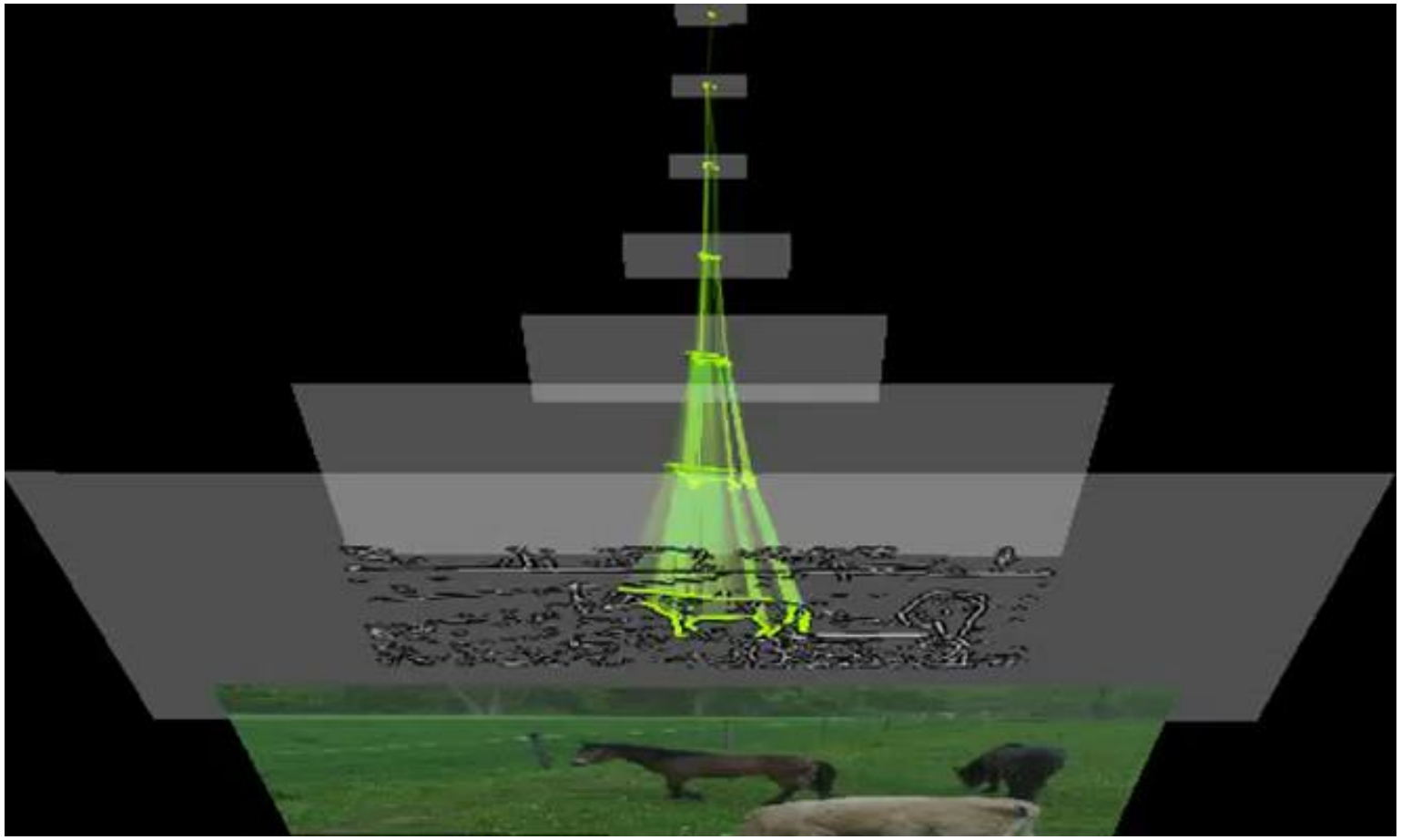
Prosojnice: S. Fidler

Razpoznavanje objektov – hierarhični pristop



http://www.cs.utoronto.ca/~fidler/slides/CSC420/videos/mugs_trick.mp4

Razpoznavanje objektov – hierarhični pristop



http://www.cs.utoronto.ca/~fidler/slides/CSC420/videos/horse.mp4

Prosojnice: S. Fidler

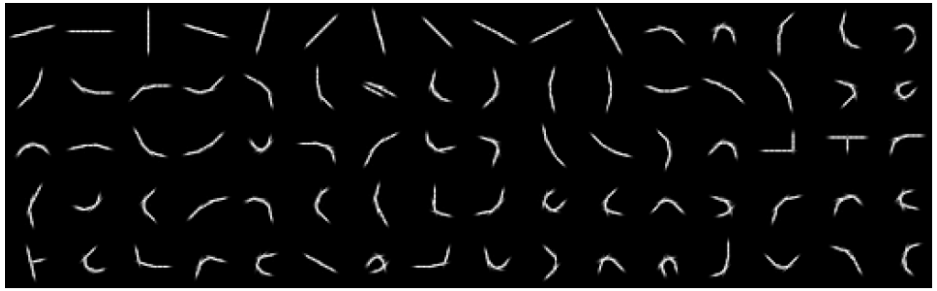
Razpoznavanje objektov – hierarhični pristop

- “Komercialna uporaba”: segmentacija oliv!

Prosojnice: S. Fidler

Razpoznavanje objektov

- Nivoje 1 – 3 lahko uporabimo pri konstrukciji deskriptorja (podobno kot HOG): v celicah preštevamo, kolikokrat se pojavi vsaka izmed naučenih oblik:



- http://www.vicos.si/Research/VicosEye
- http://eye.vicos.si/

ViCOS
Visual Cognitive Systems Lab

Razpoznavanje objektov – hierarhični pristop

- Trenutno najboljša metoda za detekcijo objektov v slikah: “Convolutional Neural Networks” (CNN)
- To so “hierarhično zgrajene nevronske mreže”.

Razpoznavanje objektov

https://www.youtube.com/watch?v=zsVsUvx8ieo

Konec