

PMR2560 – Visão Computacional

Introdução

Prof. Eduardo L. L. Cabral



Objetivos

- O que é visão computacional?
- Aplicações
- Desafios da visão computacional
- Aspectos críticos
- Como os computadores vêm
- Reconhecimento

O que é visão computacional?

- Entendimento de forma automática de imagens e vídeos por meio de:
 - Cálculo de propriedades do mundo real (3D) a partir de dados visuais
 - Uso de algoritmos e representações que permitem uma máquina reconhecer objetos, cenas, pessoas e atividades (percepção e interpretação)

Para que serve visão computacional?

- Número de imagens disponíveis aumenta a cada dia e também as aplicações possíveis:
 - Liberar o ser humano de atividades chatas e fáceis;
 - Melhorar habilidades humanas: interação computador-homem, visualização
 - Percepção para robôs e agentes autônomos
 - Organização e acesso facilitado a conteúdo visual
- Objetivos da visão computacional:
 - Prover as máquinas de habilidade de entender cenas
 - Assistência ao entendimento e modelagem da visão humana (dispositivos para deficientes visuais)
 - Automação de operações de visualização

Para que serve visão computacional?



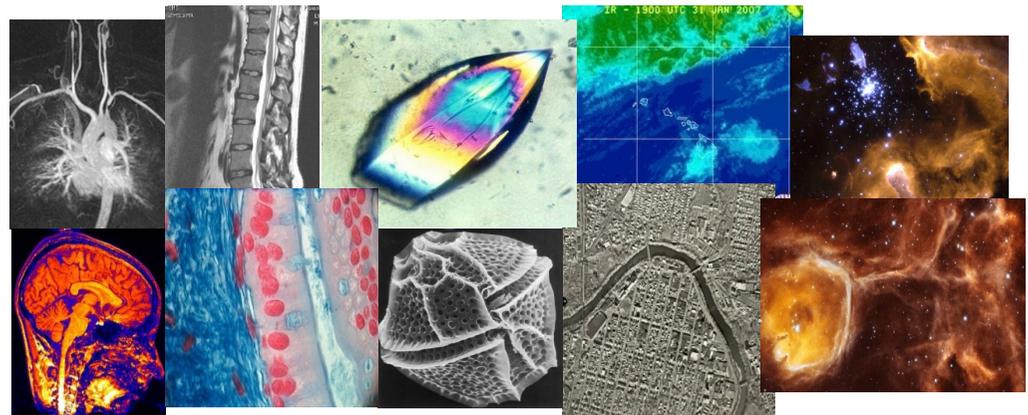
Álbuns de fotos pessoais



Filmes, notícias, esportes



Monitoramento e segurança

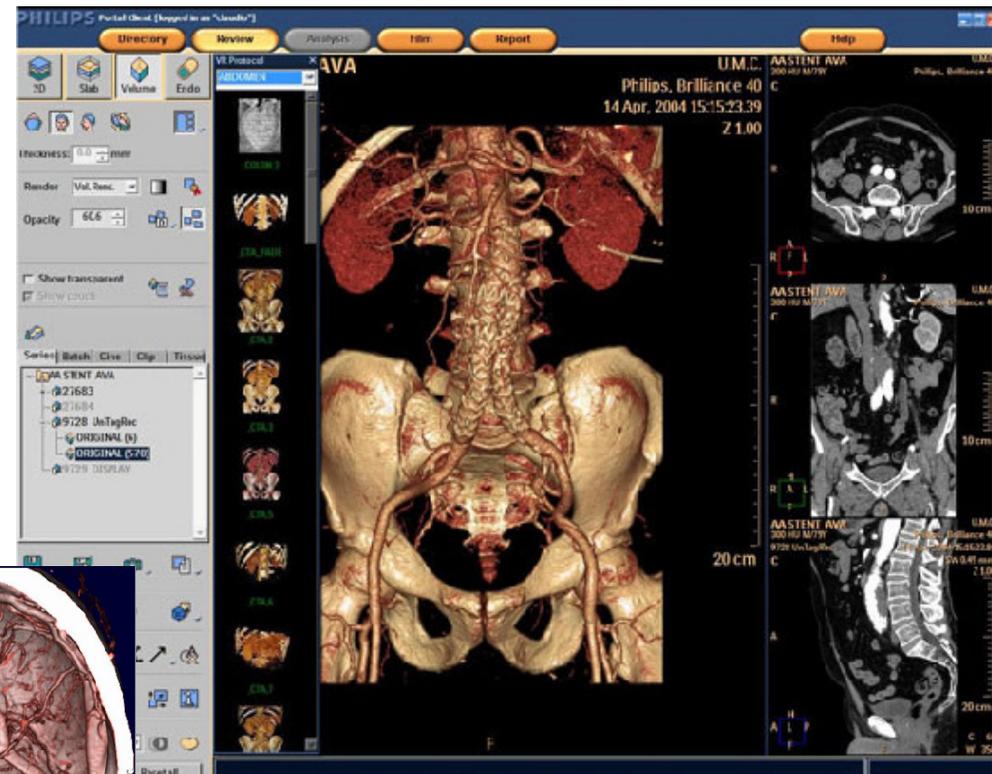
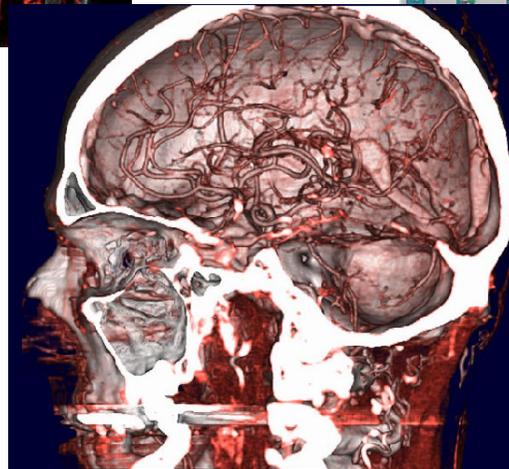


Imagens científicas e médicas

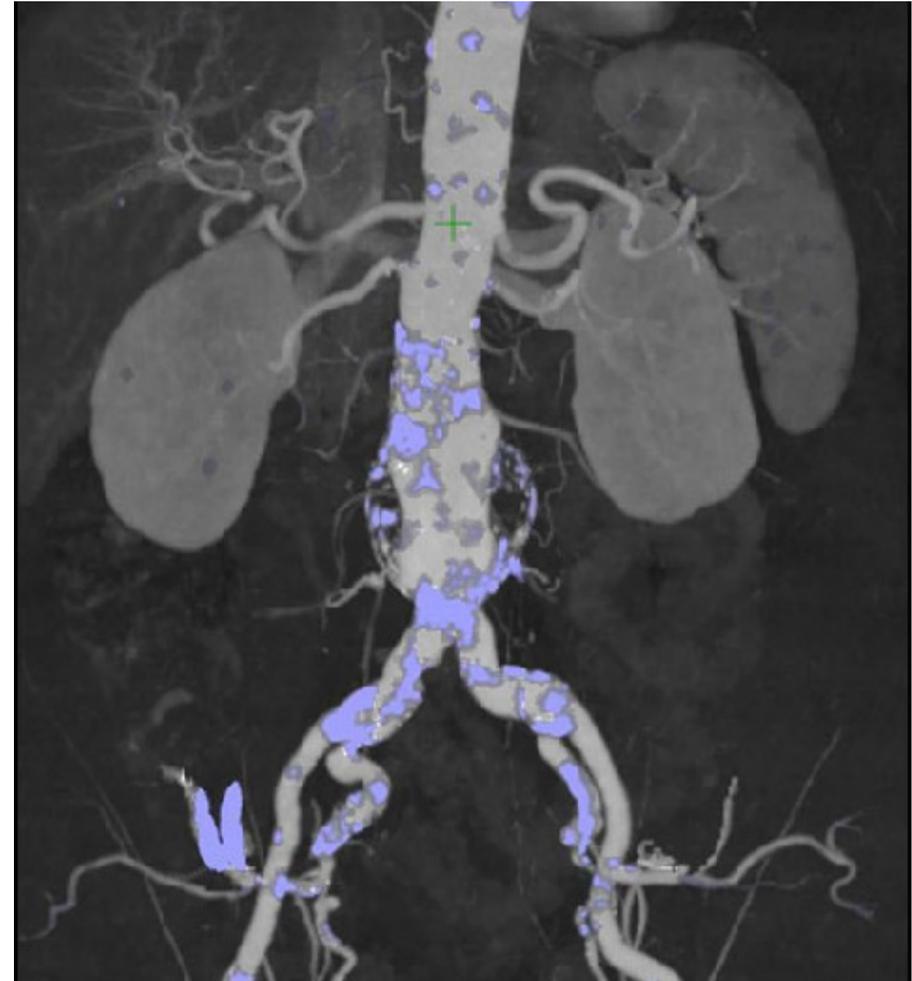
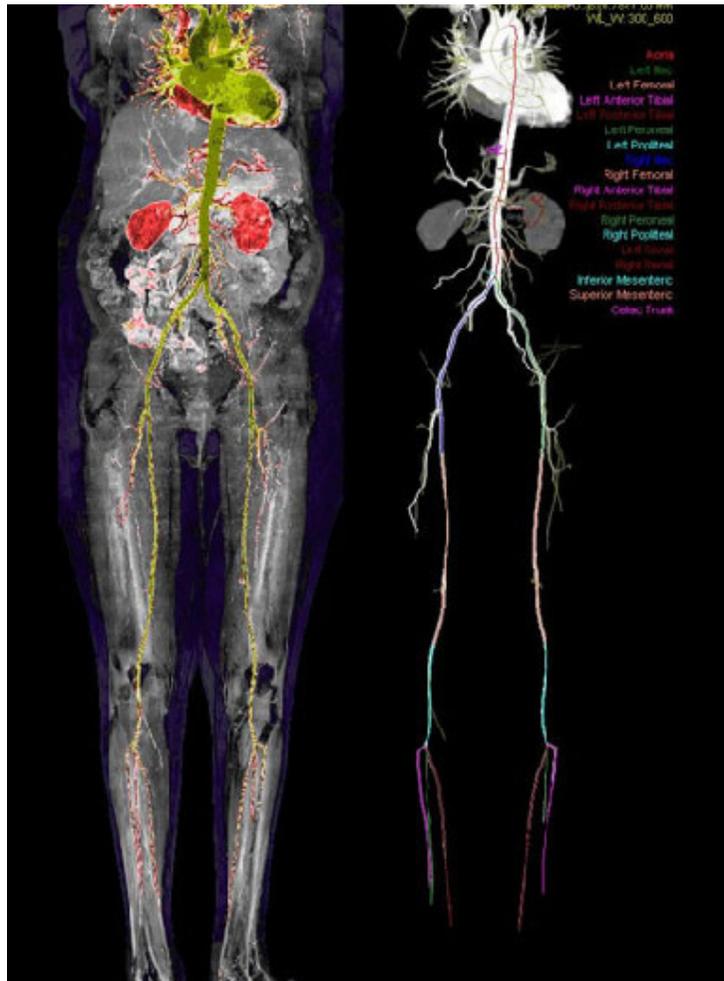
Aplicações

- Análise de imagens médicas
- Segurança: biometria, monitoramento, localização, reconhecimento de alvos
- Sensoriamento remoto
- Robótica
- Inspeção e controle de qualidade industrial
- Análise de documentos
- Multimídia
- Assistência a deficientes
- Interface homem -computador
- ...

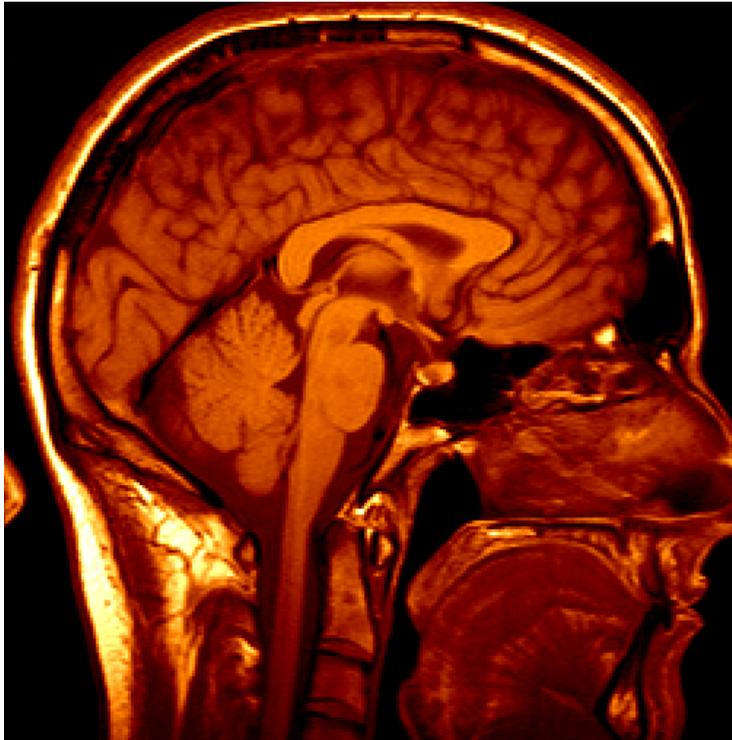
Análise de imagens médicas



Análise de imagens médicas



Análise de imagens médicas

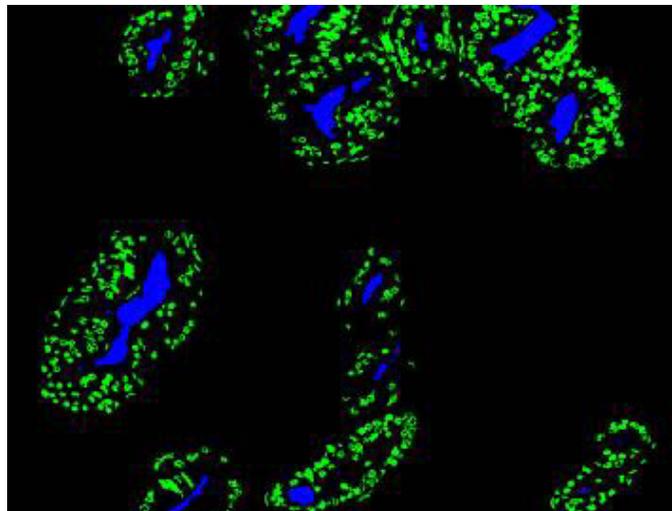
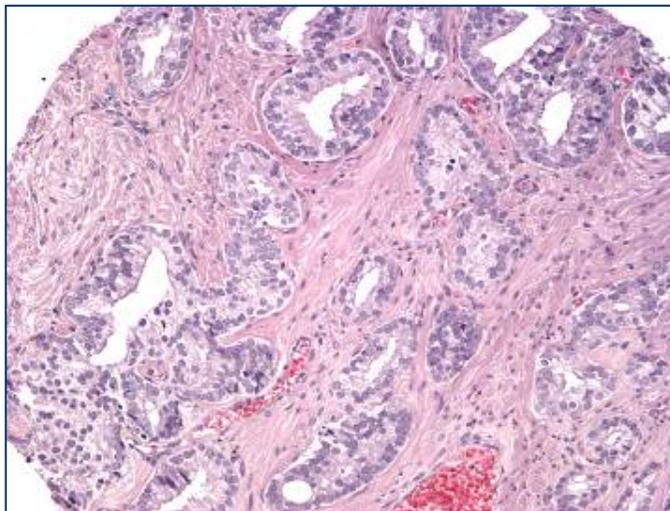
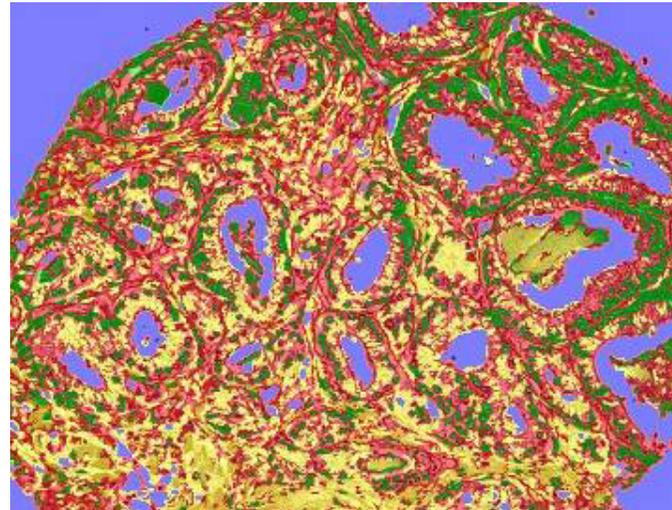
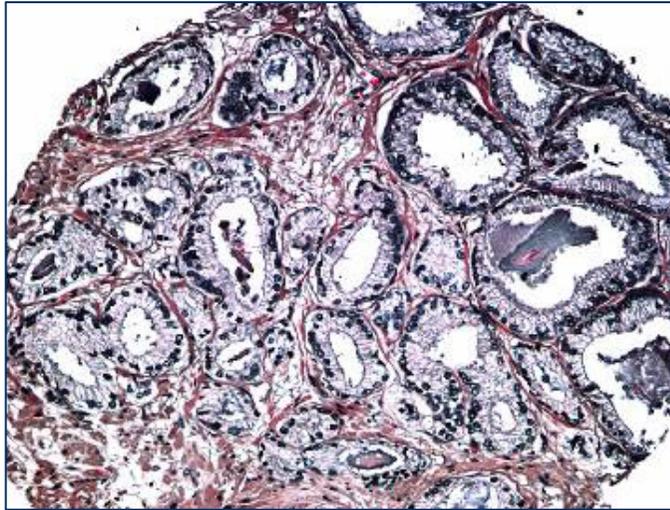


Imagens 3D: MRI, CT



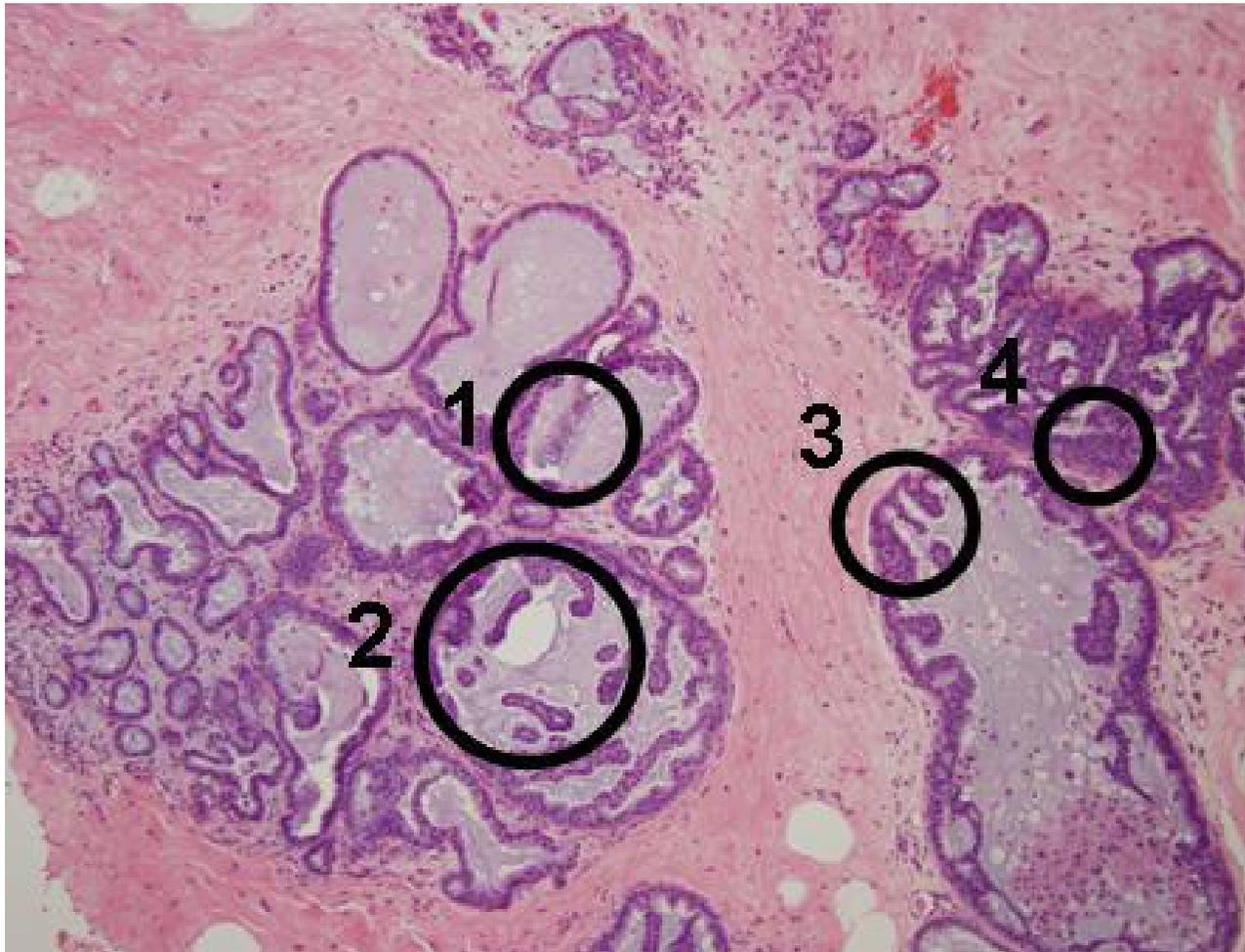
Cirurgia guiada por imagem
(Grimson et al., MIT)

Análise de imagens médicas



Detecção e
classificação
de cancer

Análise de imagens médicas

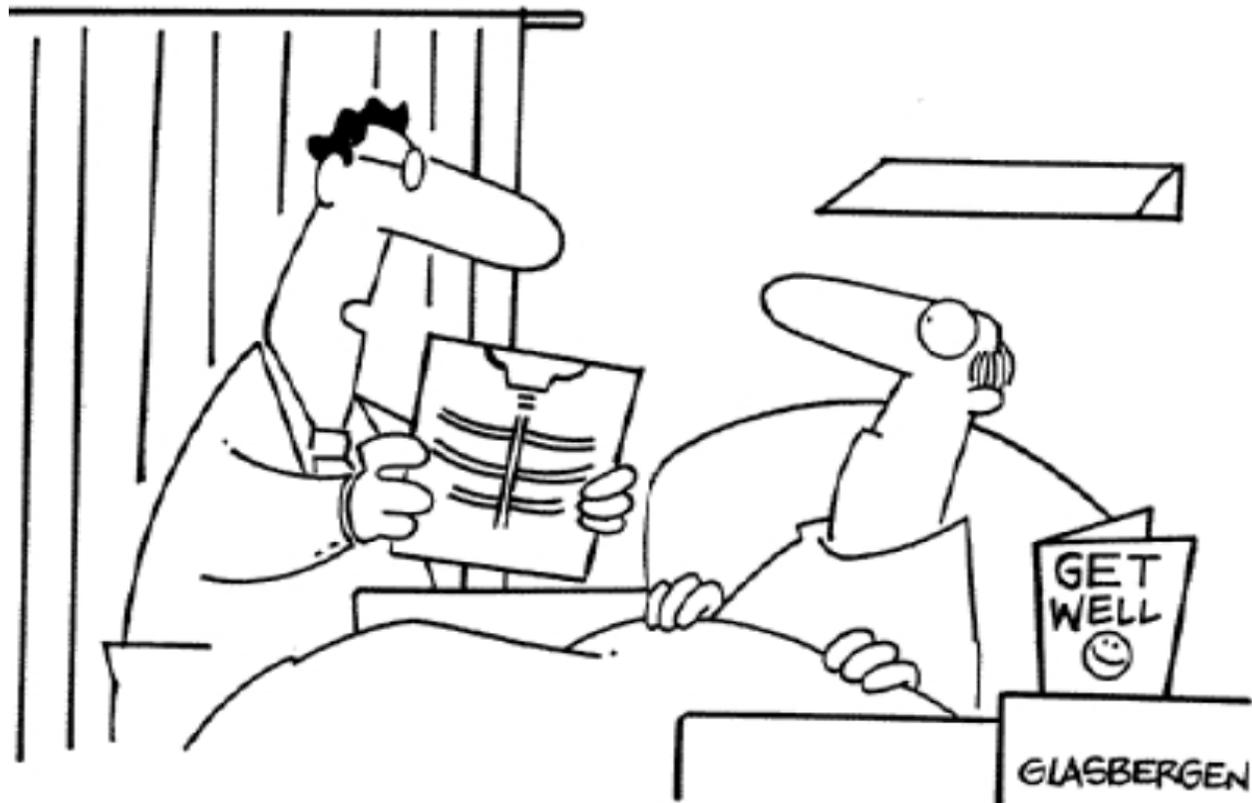


Imagens de biópsias: detecção de regiões de interesse

De Linda Shapiro, U of Washington

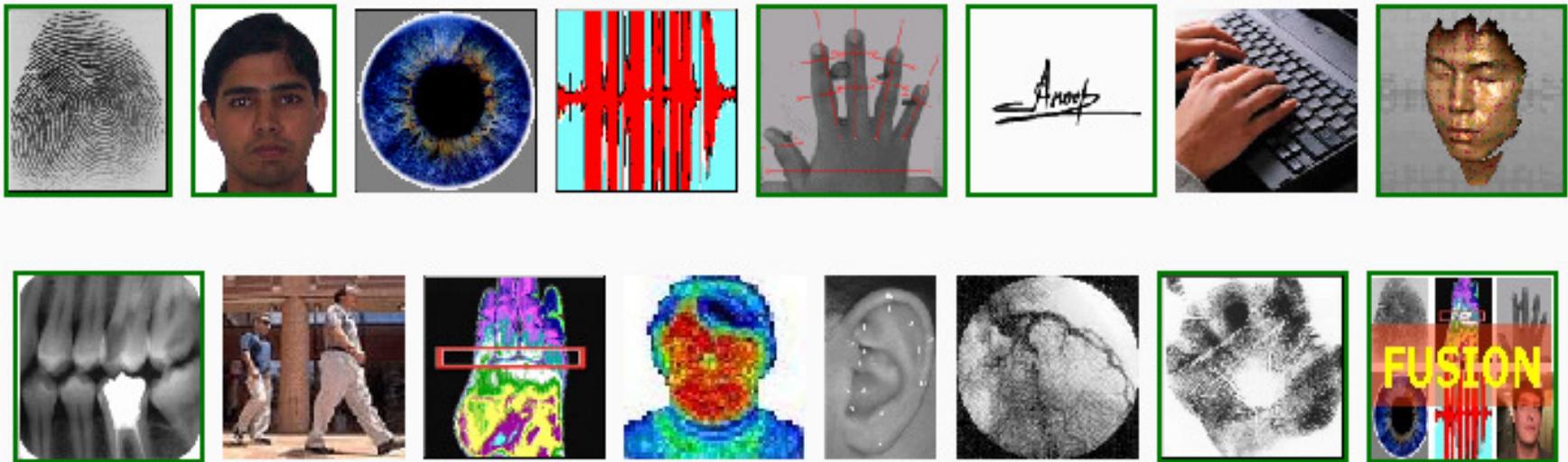
Análise de imagens médicas

© 2000 Randy Glasbergen. www.glasbergen.com



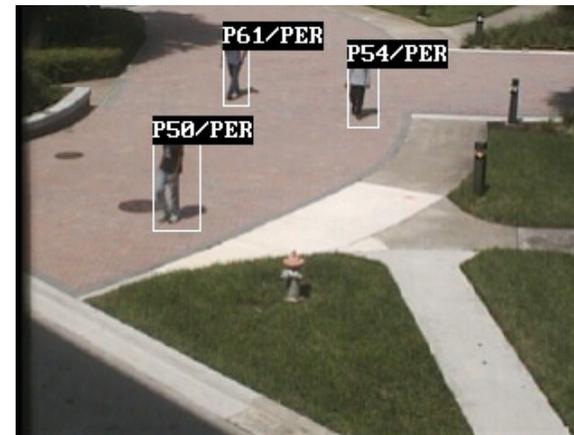
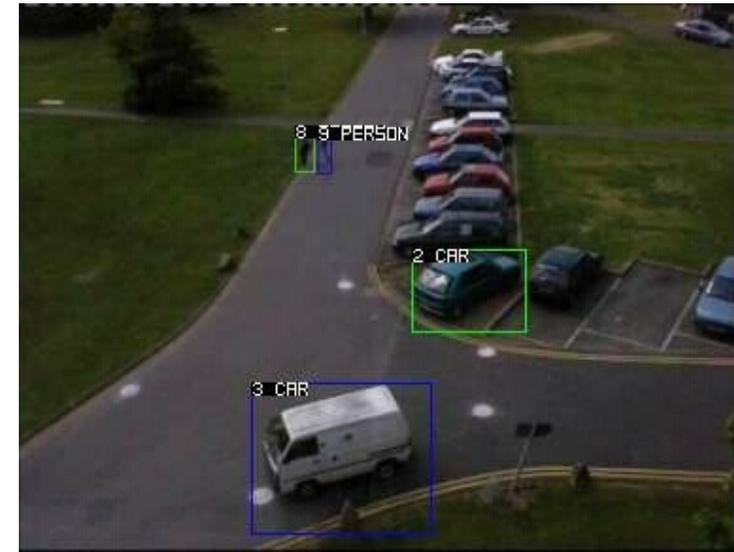
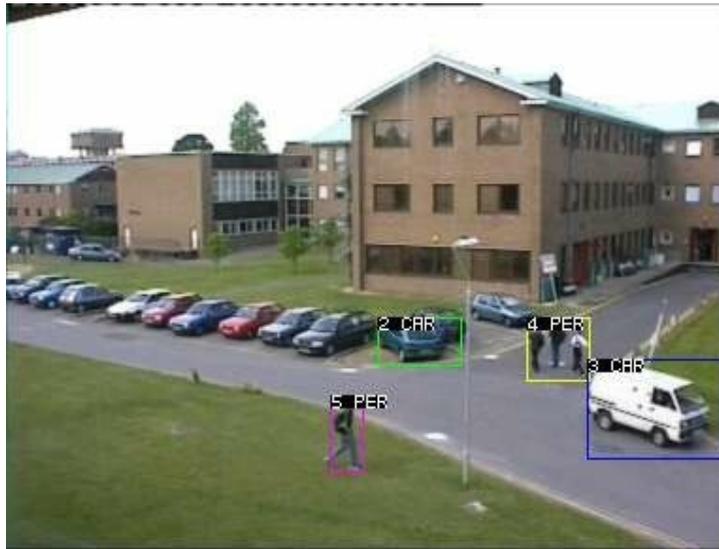
**“Your x-ray showed a broken rib,
but we fixed it with Photoshop.”**

Biometria



De Anil Jain, Michigan State

Monitoramento e localização

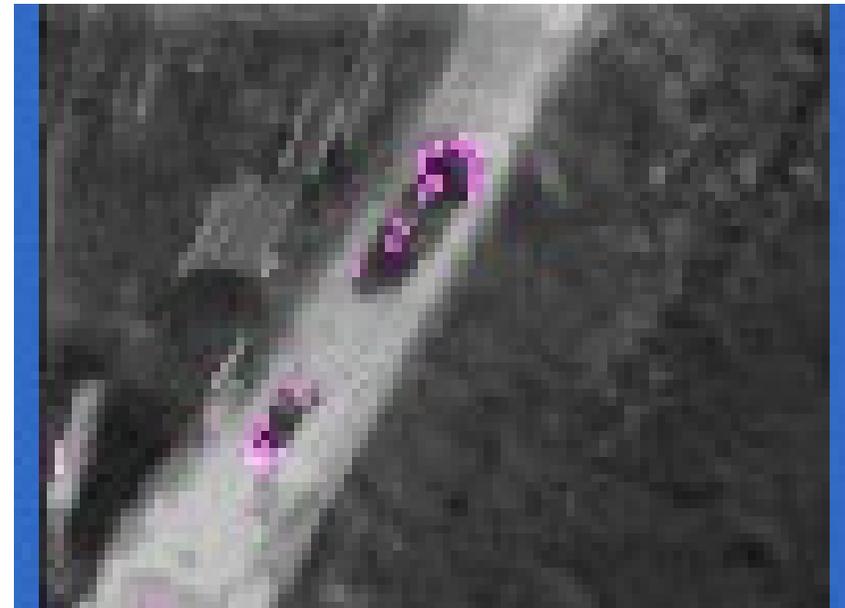


Monitoramento e localização



De Martial Hebert, CMU

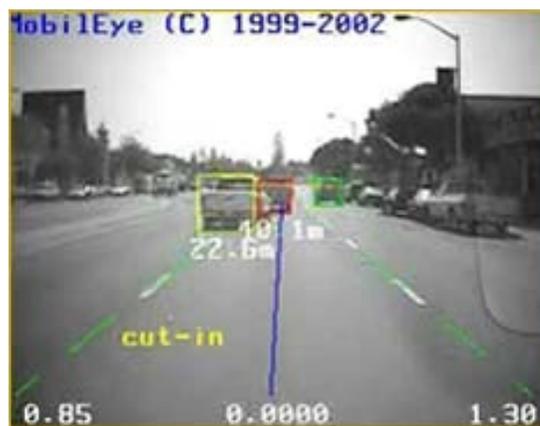
Monitoramento e localização



Localização por meio de vídeos de UAV

De Martial Hebert, CMU, and Masaharu Kobashi, U of Washington

Segurança de pedestres e veículos



Aviso de saída da faixa, aviso de colisão, reconhecimento de sinal de trânsito, reconhecimento de pedestre, aviso de ponto cego

Carros inteligentes

▶ manufacturer products consumer products ◀◀

Our Vision. Your Safety.

rear looking camera forward looking camera
side looking camera

▶ **EyeQ** Vision on a Chip

▶ read more

▶ **Vision Applications**

Road, Vehicle, Pedestrian Protection and more

▶ read more

▶ **AWS** Advance Warning System

▶ read more

News

- ▶ [Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System](#)
- ▶ [Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end](#)

▶ all news

Events

- ▶ [Mobileye at Equip Auto, Paris, France](#)
- ▶ [Mobileye at SEMA, Las Vegas, NV](#)

▶ read more

Carros inteligentes



<http://www.darpa.mil/grandchallenge/index.asp>
http://en.wikipedia.org/wiki/DARPA_Grand_Challenge

Carros inteligentes

Under the bonnet

How a self-driving car works

Signals from **GPS (global positioning system)** satellites are combined with readings from tachometers, altimeters and gyroscopes to provide more accurate positioning than is possible with GPS alone

Lidar (light detection and ranging) sensors bounce pulses of light off the surroundings. These are analysed to identify lane markings and the edges of roads

Video cameras detect traffic lights, read road signs, keep track of the position of other vehicles and look out for pedestrians and obstacles on the road

Radar sensor

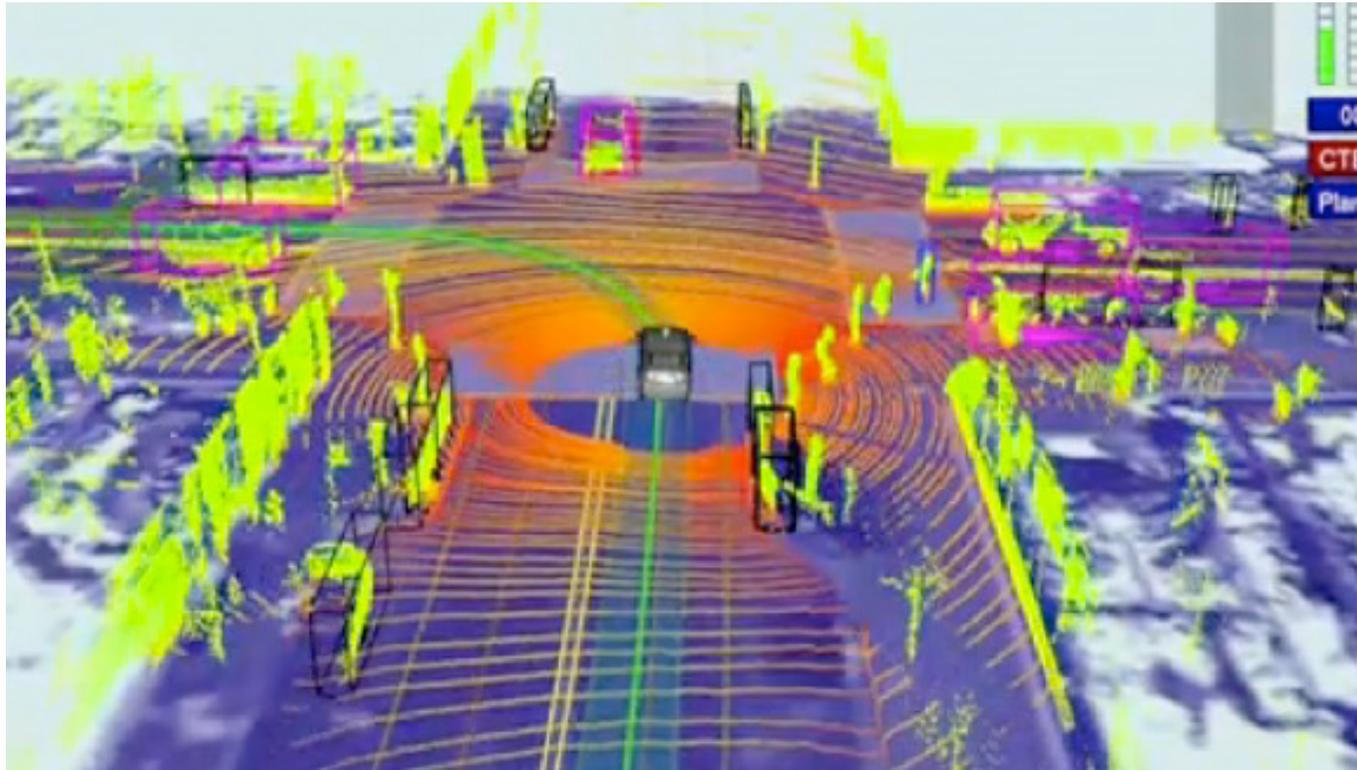
Ultrasonic sensors may be used to measure the position of objects very close to the vehicle, such as curbs and other vehicles when parking

The information from all of the sensors is analysed by a **central computer** that manipulates the steering, accelerator and brakes. Its software must understand the rules of the road, both formal and informal

Radar sensors monitor the position of other vehicles nearby. Such sensors are already used in adaptive cruise-control systems

Source: *The Economist*

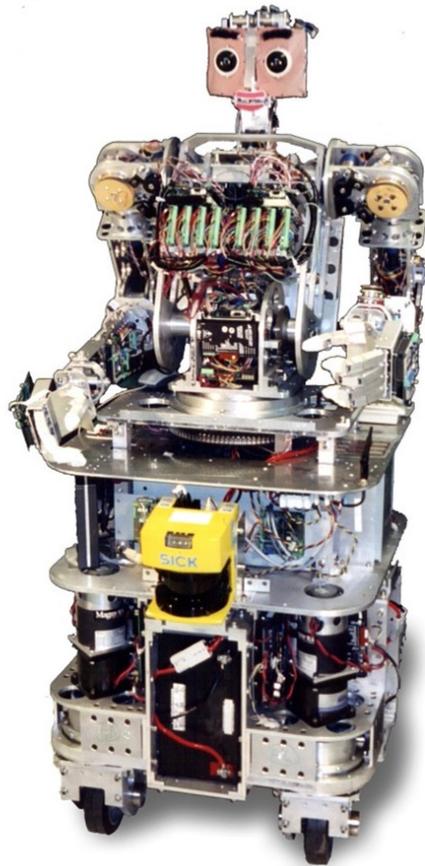
Carros inteligentes



“Our self-driving cars have now traveled nearly 200,000 miles on public highways in California and Nevada, 100 percent safely. They have driven from San Francisco to Los Angeles and around Lake Tahoe, and have even descended crooked Lombard Street in San Francisco. They drive anywhere a car can legally drive.”

- Sebastian Thrun, Google

Navegação autônoma



Michigan State University



General Dynamics Robotics Systems
<http://www.gdrs.com>

Monitoramento de incêndios florestais

Orman Yangını Erken Uyarı Sistemi

24 August 2008
12:22:38

Acil Durum! Duman Tespit Edildi.

Camera List

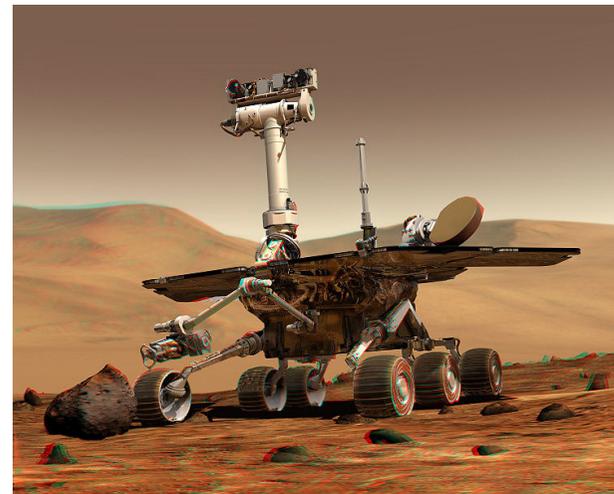
- Marmaris Tansac
- Marmaris AltinSivri
- Marmaris Palamut Bah
- Marmaris Kovalica
- Marmaris Palamut Bah
- Manavgat Gelin
- Manavgat Sorgun
- Manavgat Turkkeleni
- Manavgat Kartal
- Manavgat Sivri
- Aksehir Tekke Daresi - 1
- Aksehir Tekke Daresi - 2
- Aksehir Tekke Daresi - 3
- Aksehir Tekke Daresi - 4
- Aksehir Tekke Daresi - 5
- Aksehir Tekke Daresi - 6
- Aksehir Tekke Daresi - 7
- Aksehir Tekke Daresi - 8
- Aksehir Tekke Daresi - 9
- Aksehir Tekke Daresi - 10
- Aksehir Tekke Daresi - 11
- Aksehir Tekke Daresi - 12
- Aksehir Tekke Daresi - 13
- Aksehir Tekke Daresi - 14
- Aksehir Tekke Daresi - 15
- Aksehir Tekke Daresi - 16

OGM

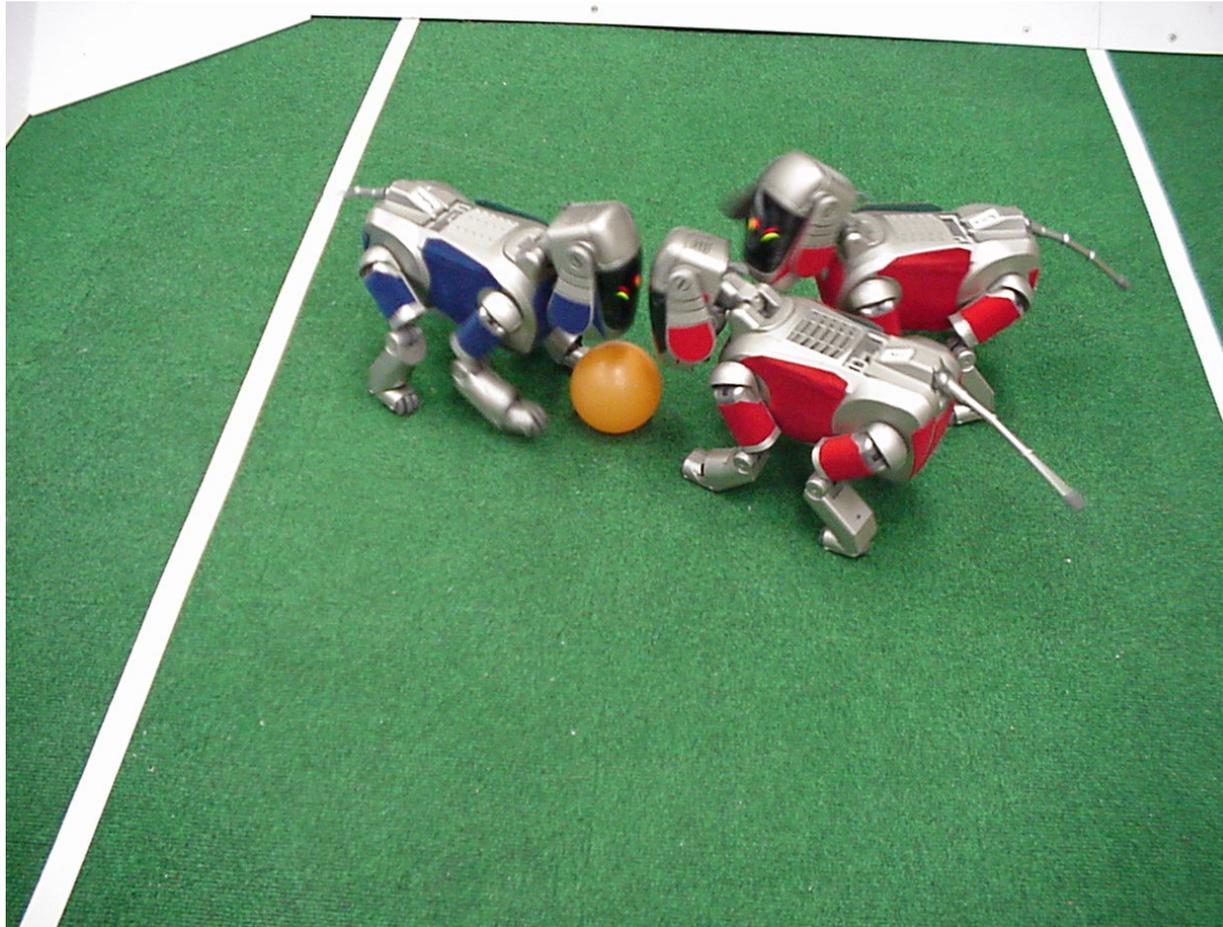
Alerta precoce de incêndios

De Enis Cetin, Bilkent University

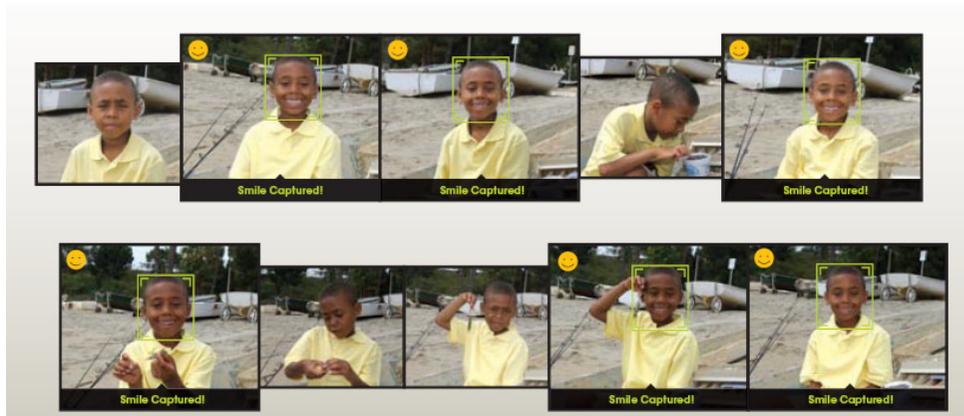
Robótica



Robótica



Detecção de face



Automação industrial



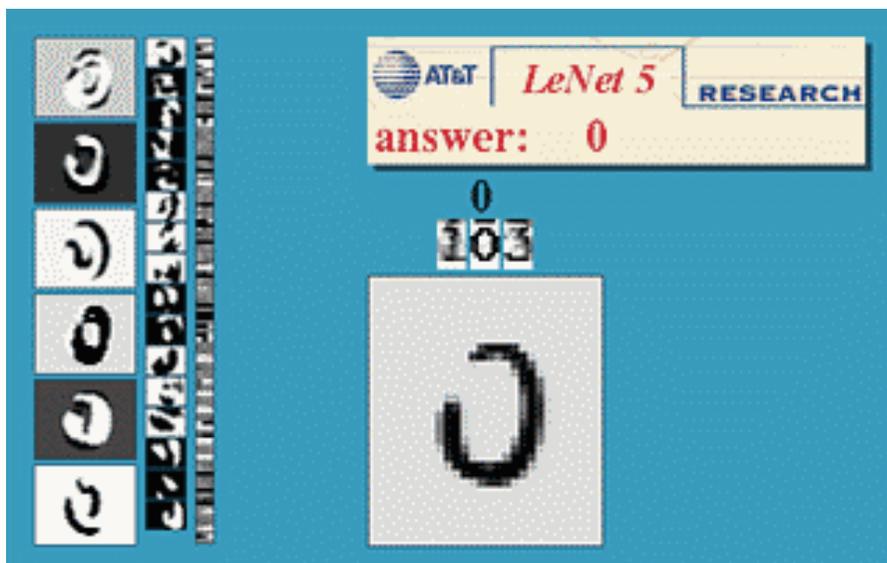
Seleção automática de frutas

Automação industrial



Robô industrial
realizando tarefa de
“pick and place”
guiado por câmera

Reconhecimento de carácter



Reconhecimento de letras, AT&T labs
<http://www.research.att.com/~yann>



De Steven Seitz, U of Washington

Reconhecimento de placas

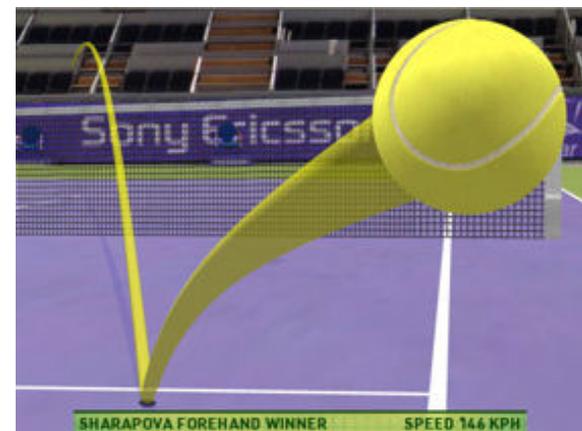
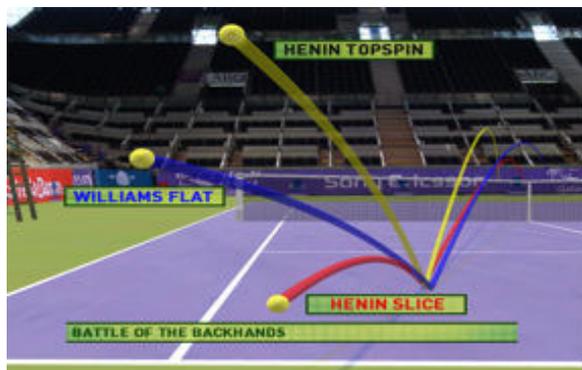
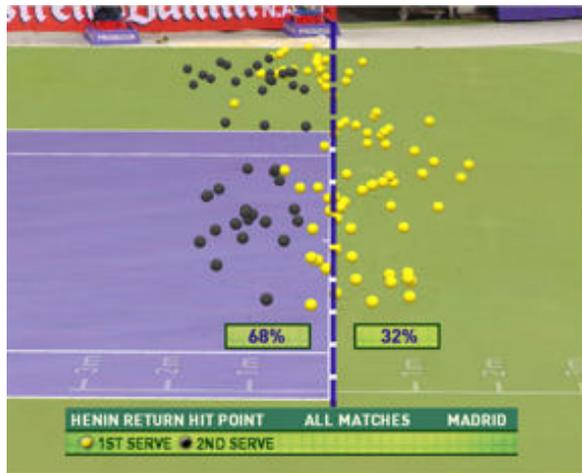
Análise de documentos

儘眼望遠極，
佰程無窮哩。
壹物明域現，
以迺吾後脊！

I looked as hard as I could see,
beyond 100 plus infinity
an object of bright intensity
– it was the back of me!

Figure 1.5: (Left) Chinese characters and (right) English equivalent. Is it possible that a machine could automatically translate one into the other? Chinese characters and poem courtesy of John Weng.

Análise de video de esportes



Sistema de replay em jogo de tênis

<http://www.hawkeyeinnovations.co.uk>

Realidade aumentada



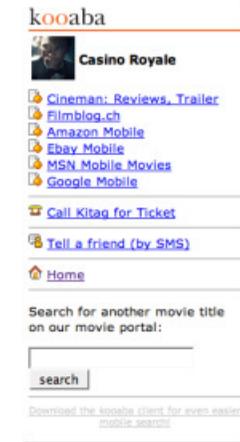
Classificação de terras/áreas



Reconhecimento de objetos



Lincoln, Microsoft Research



kooaba



Pesquisa de local
Yeh et al., MIT



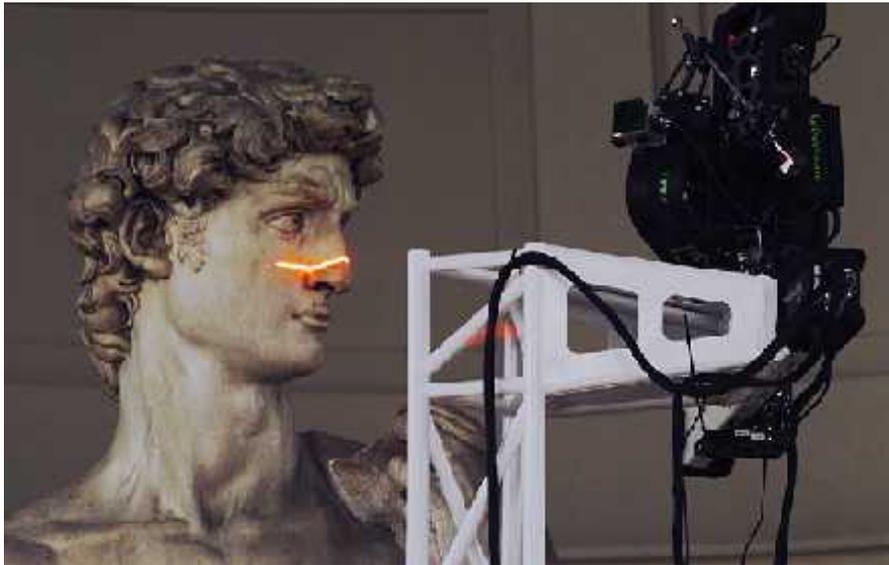
Google Goggles
Bing Vision

Reconhecimento de objetos

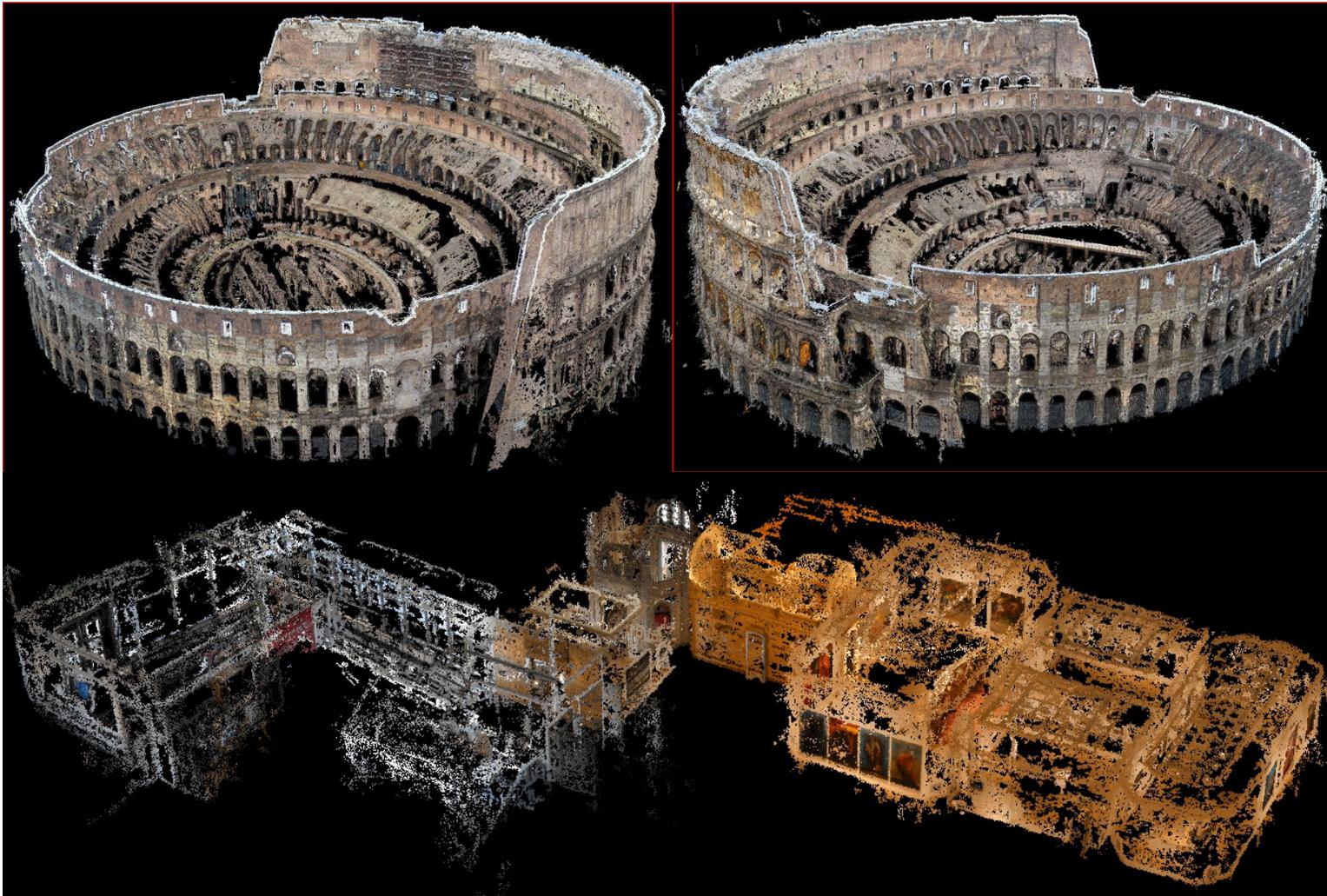


Reconhecimento de prédios e grupo de prédios

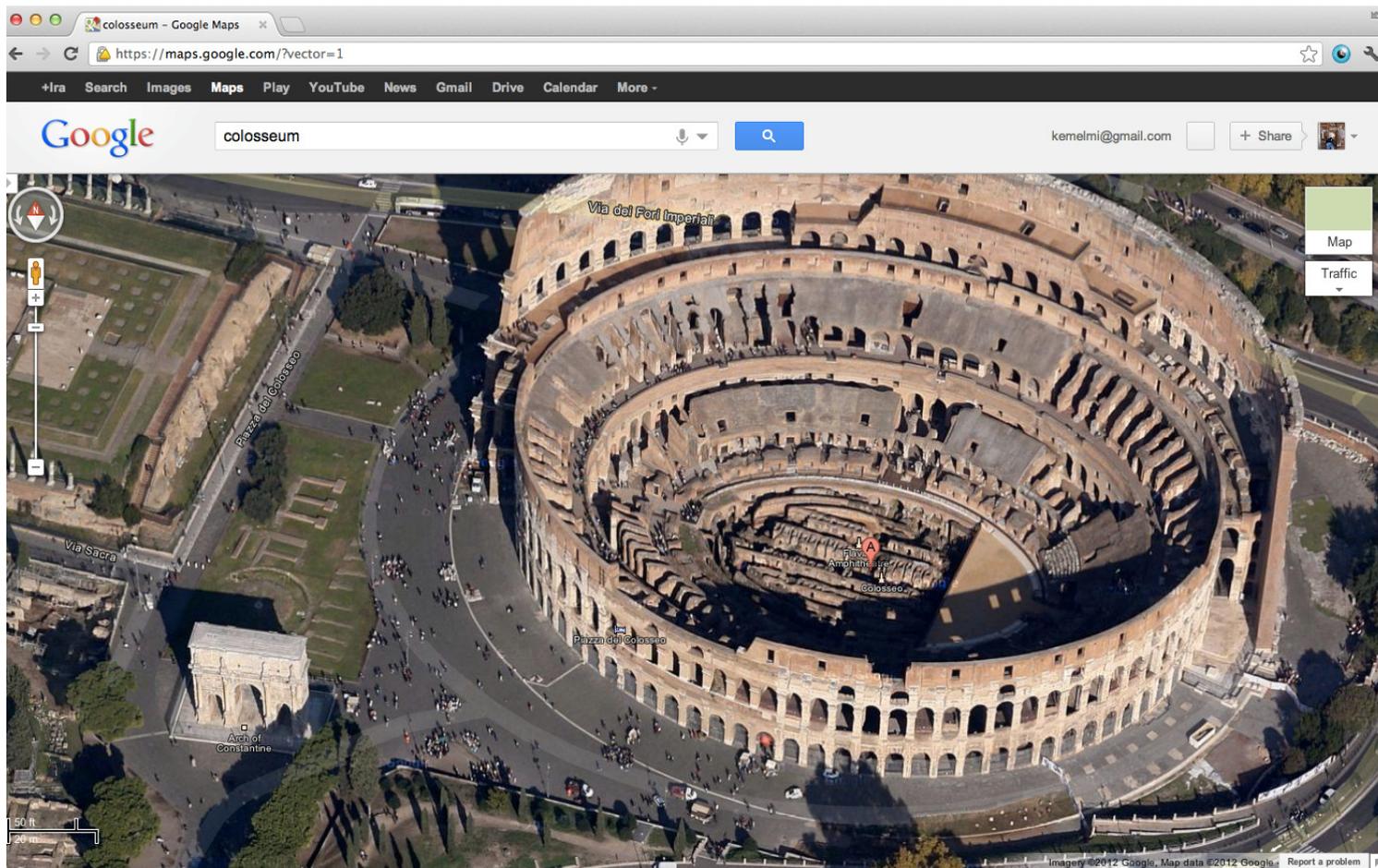
3D scanning and reconstruction



Modelagem 3D



Visão da superfície terrestre



Visão da superfície terrestre



Mapa 3D da Apple

Efeitos visuais



Captura de movimento



De Linda Shapiro, U of Washington

Captura de movimento



Kinect do Xbox da Microsoft

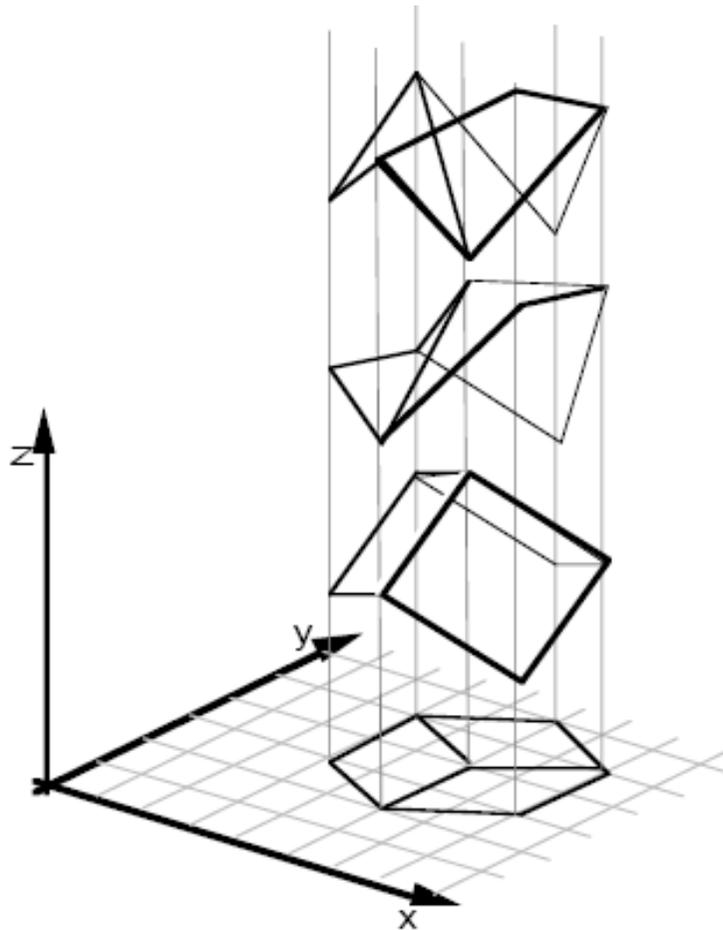
Mosaico – imagens panorâmicas



Mosaico – imagens panorâmicas



Porque visão é tão difícil?



Problema mal-posto

[Sinha and Adelson 1993]

Desafio 1: ponto de vista

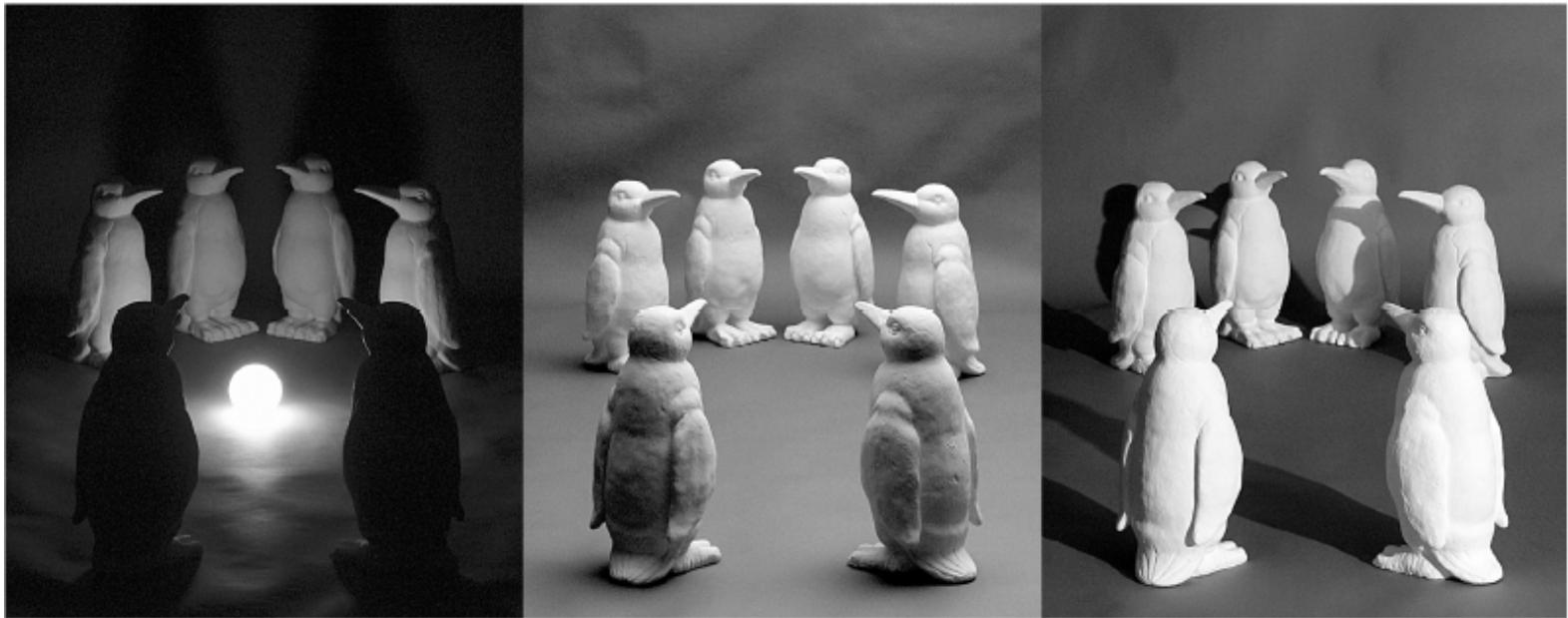


Michelangelo 1475-1564



De L. Fei-Fei,
R. Fergus, A. Torralba

Desafio 2: iluminação



De Fei-Fei Li

Desafio 3: oclusão



Magritte, 1957

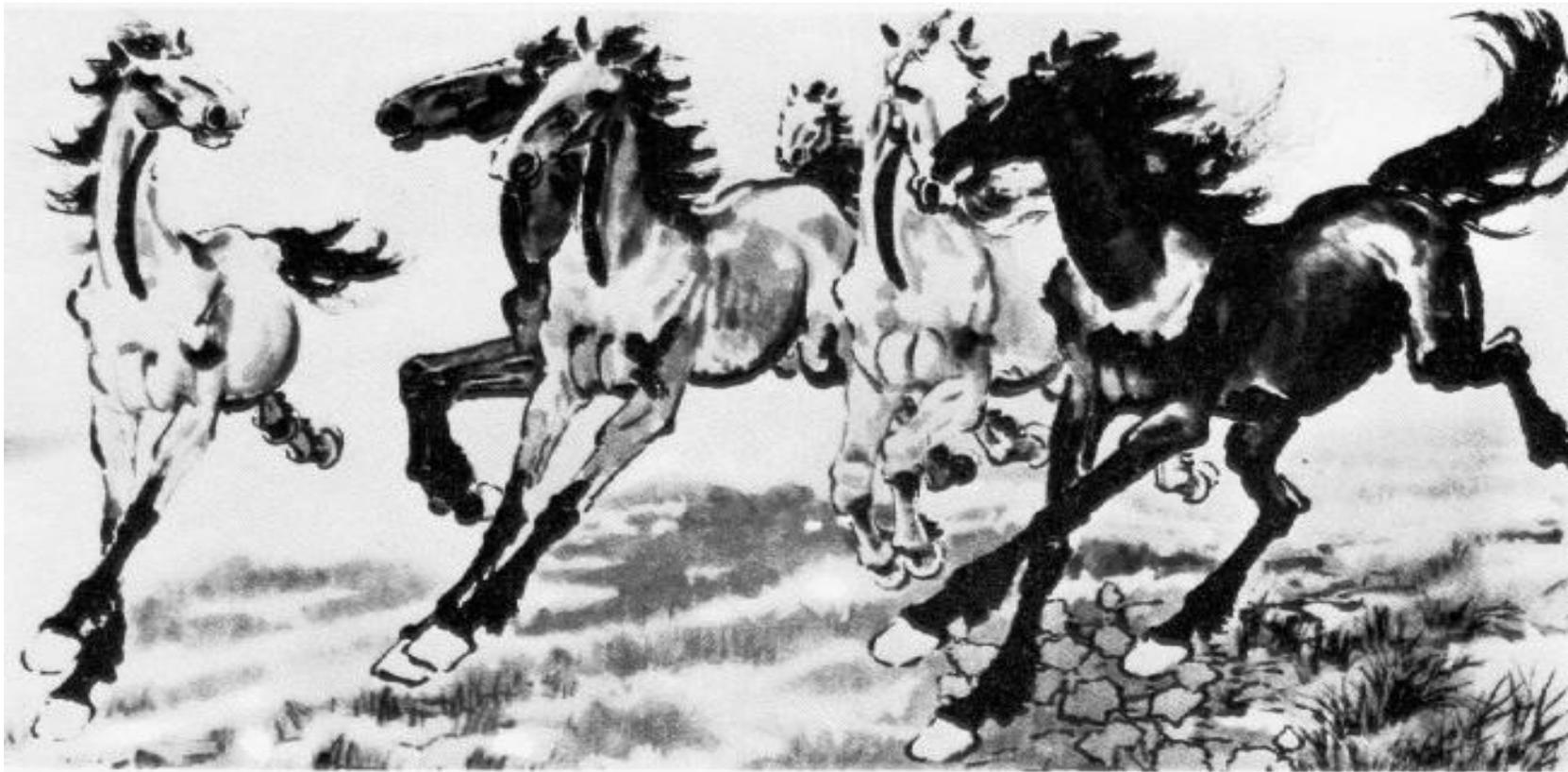
De L. Fei-Fei,
R. Fergus, A. Torralba

Desafio 4: escala



De L. Fei-Fei,
R. Fergus, A. Torralba

Desafio 5: deformação



De L. Fei-Fei, R. Fergus, A. Torralba

Xu, Beihong 1943

Desafio 6: fundo confuso



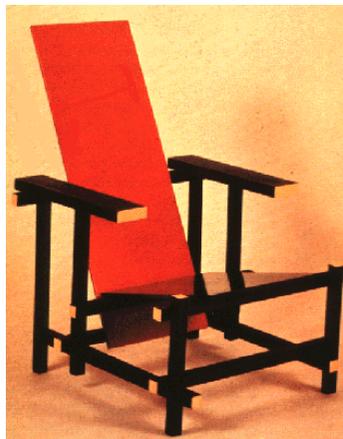
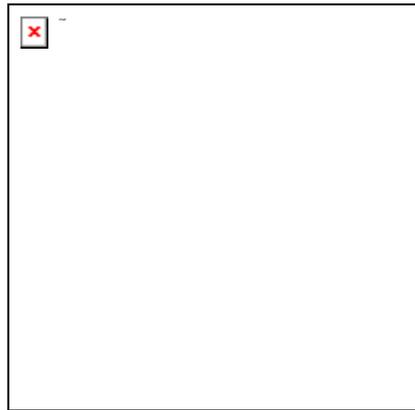
Emperor shrimp and cannibal crab on a sea cucumber in Fiji
Photograph by Tim Laman

NATIONAL
GEOGRAPHIC

© 2007 National Geographic Society. All rights reserved.

De Fei-Fei Li

Desafio 7: variação intra-classe



De L. Fei-Fei, R. Fergus, A. Torralba

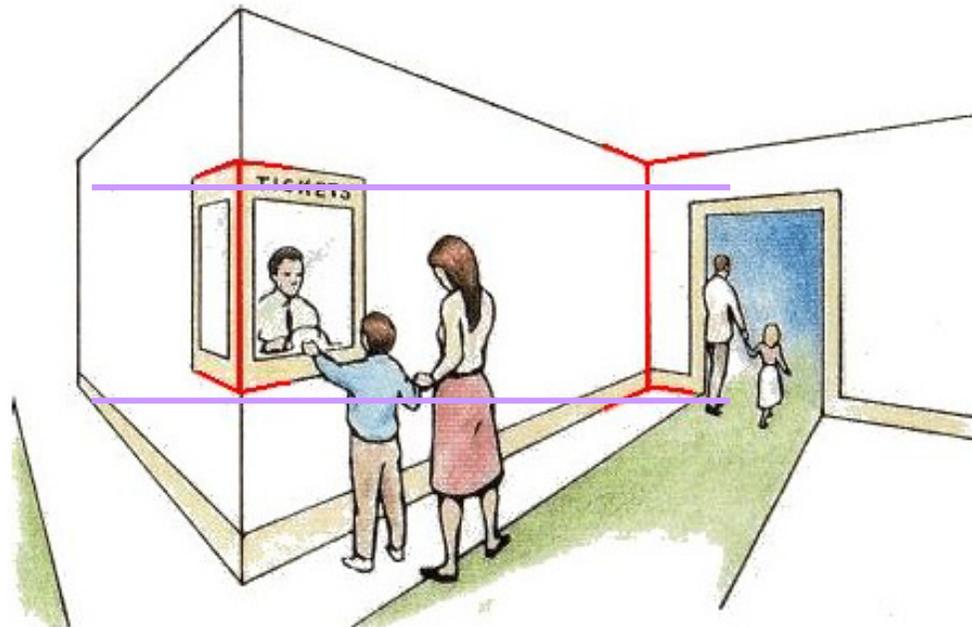
Desafio 8: percepção e agrupamento



Contornos
subjetivos

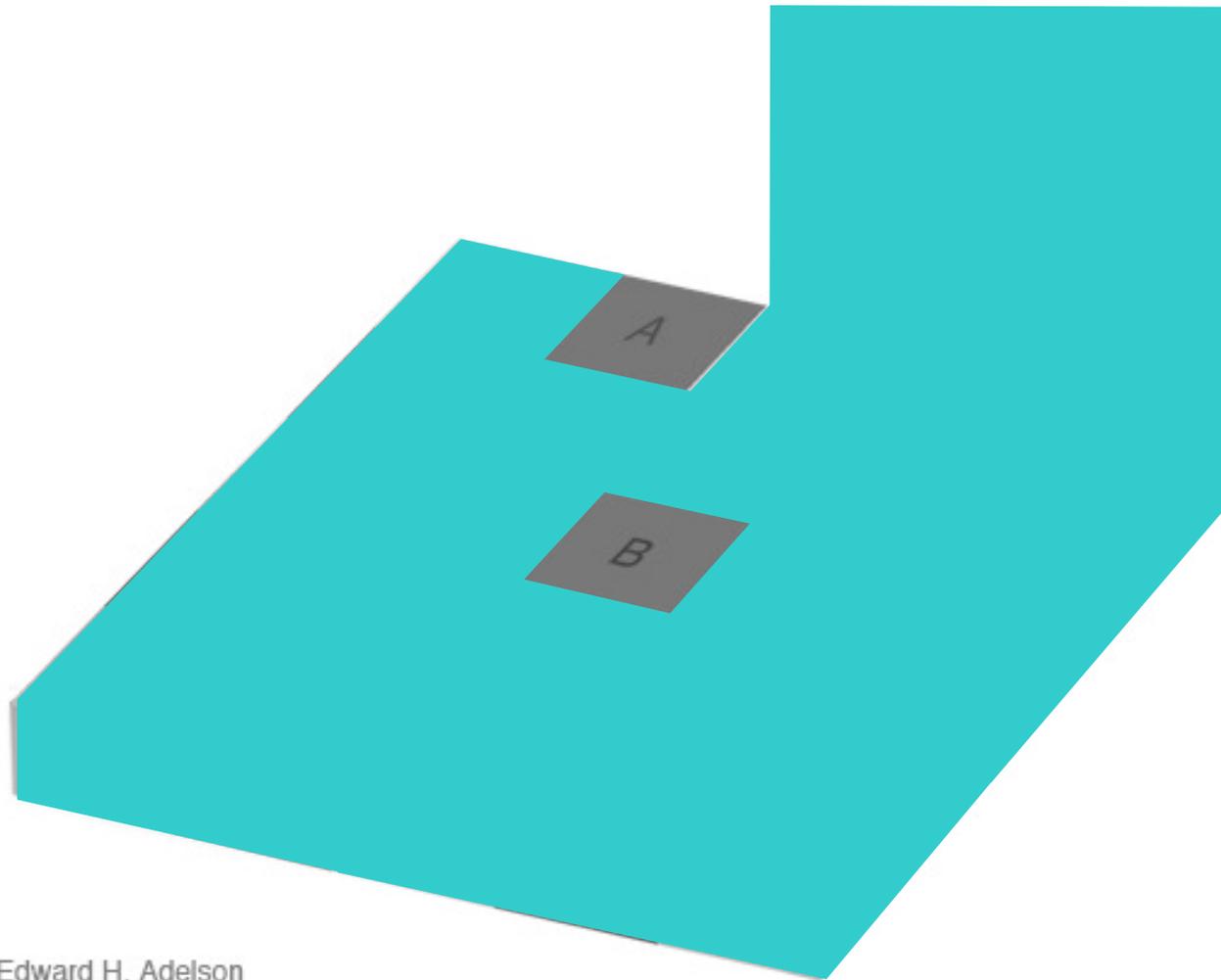
Desafio 8: percepção e agrupamento

Ilusão Müller-Lyer



De Alyosha Efros, Carnegie Mellon

Desafio 8: percepção e agrupamento



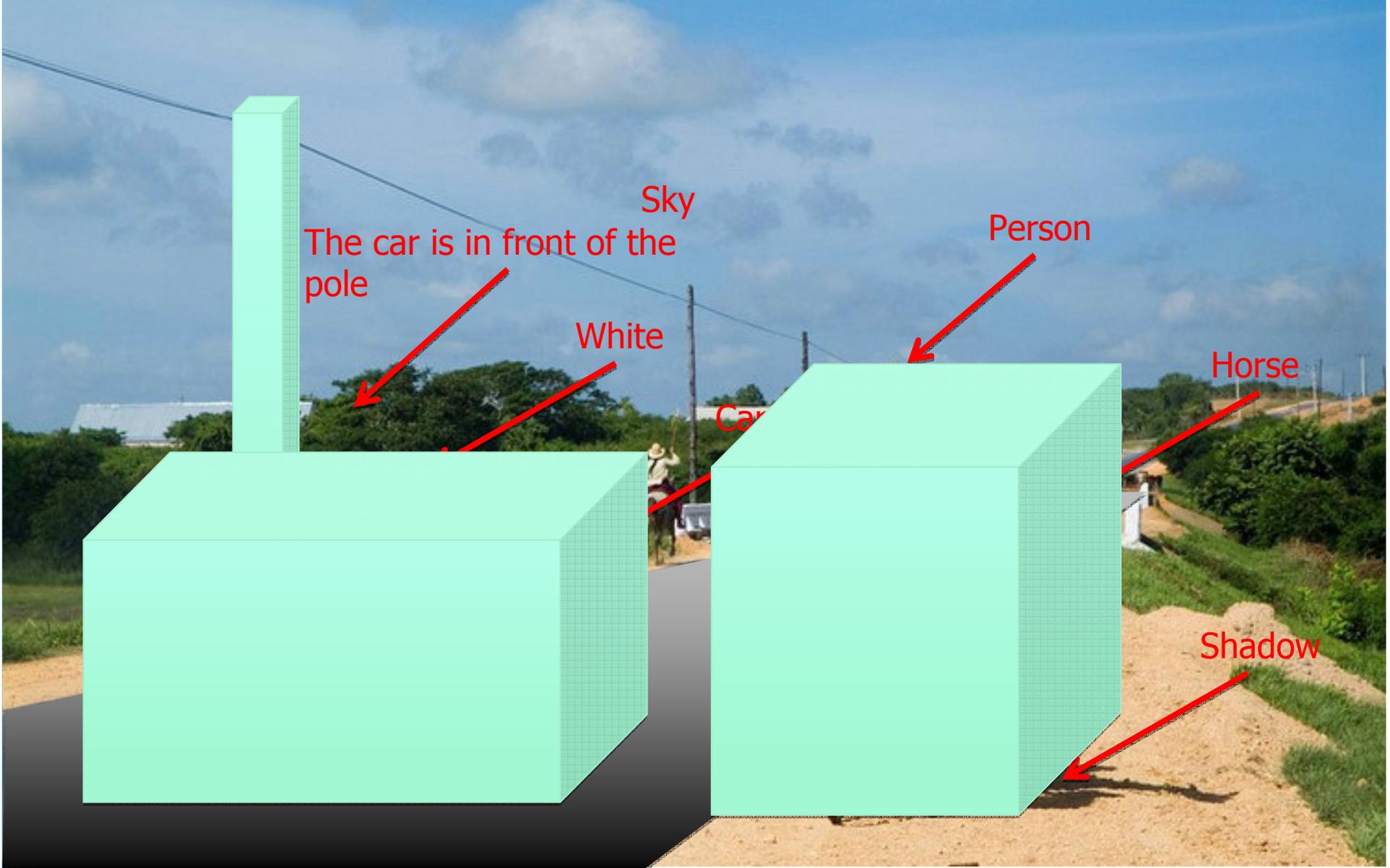
Edward H. Adelson

Questões críticas

- Qual informação deve ser extraída da imagem?
- Como a informação pode ser extraída?
- Como a informação deve ser representada?
- Como a informação pode ser usada para análise e entendimento?

O que os computadores vêm?

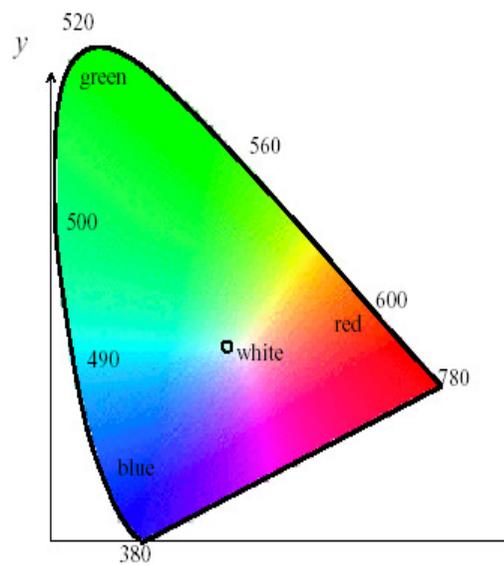
47	49	51	47	41	41	41	38	42	54	66	66	58	56	53	48	43	43	45	47	50	47	47	47
45	44	39	38	37	48	67	95	138	151	156	157	165	157	125	79	36	38	47	48	48	43	38	36
43	35	31	45	64	109	155	179	178	160	142	132	146	187	195	170	133	86	45	46	51	41	36	32
33	24	24	47	88	149	135	136	160	170	166	135	111	153	169	169	109	113	86	57	49	46	40	36
22	19	22	47	122	131	99	120	204	199	185	150	119	152	159	173	110	80	83	82	63	58	45	42
22	20	24	60	114	108	123	191	215	212	198	169	156	169	168	172	151	115	91	77	82	59	53	53
20	19	29	86	127	87	169	223	219	218	212	182	178	190	194	185	169	108	88	85	74	55	52	51
20	20	26	131	138	129	214	228	224	222	221	206	207	208	203	193	177	136	88	87	72	54	44	42
24	23	28	130	125	152	226	224	222	223	217	218	214	201	185	168	164	114	70	39	45	47	39	34
29	26	25	104	92	123	220	226	230	228	218	213	210	193	152	118	136	97	50	26	39	41	36	33
26	24	25	66	95	140	222	223	228	225	218	208	205	181	140	97	101	121	71	35	78	51	40	37
26	30	24	51	149	179	224	221	218	215	205	204	210	191	140	108	107	127	112	43	46	42	39	40
27	34	30	23	142	198	210	226	233	220	205	204	222	210	175	154	134	125	137	51	54	55	44	34
26	32	29	18	124	197	178	174	140	113	182	183	174	112	98	74	34	69	126	54	53	78	59	41
30	27	26	19	114	197	207	138	73	43	167	191	49	29	139	66	33	76	92	60	85	50	42	40
26	25	23	18	91	198	220	221	184	133	210	214	40	112	210	129	120	105	81	62	60	28	22	30
23	19	16	13	53	201	211	227	220	227	226	216	75	72	196	190	130	58	62	58	32	21	24	26
18	14	12	11	13	93	198	220	226	209	219	218	121	34	148	170	53	37	50	25	17	17	23	24
17	15	14	13	15	25	177	203	189	151	223	219	139	59	33	78	30	39	45	26	22	21	16	38
12	14	17	13	15	11	125	201	149	194	223	203	67	19	15	22	33	43	55	37	29	28	31	68
10	13	14	11	16	15	58	196	170	193	213	175	123	34	19	48	37	93	35	32	30	38	93	118
17	19	19	20	31	35	30	145	191	201	215	182	134	47	66	89	45	196	45	16	52	98	141	149
25	28	34	34	28	32	20	105	216	215	213	187	168	130	73	26	148	195	34	12	21	76	121	123
31	36	30	26	29	42	20	77	220	215	221	213	185	131	37	117	201	85	56	11	16	10	22	38
24	20	21	40	43	42	24	106	190	235	212	188	134	85	138	178	45	89	40	13	19	13	19	21
14	21	41	43	42	32	19	131	207	250	239	197	206	236	220	33	18	94	13	16	18	11	12	17
32	36	46	39	40	27	10	157	250	230	190	156	172	216	250	135	149	50	9	18	16	12	13	18
38	38	38	45	40	29	10	140	240	244	151	50	30	118	229	255	187	11	16	26	20	18	25	29
40	34	33	31	36	27	16	117	237	253	169	60	101	217	245	255	93	3	15	25	21	21	24	27
43	34	34	32	31	21	18	80	232	252	147	85	208	247	252	207	18	13	10	13	19	20	20	21
41	33	33	32	31	18	27	64	220	211	62	71	209	246	250	108	5	19	11	13	16	18	21	20
40	33	33	34	30	17	31	50	182	159	49	45	136	248	208	24	11	13	12	17	11	10	15	19



Reconhecimento

- Como diferentes características (cor, textura, formato, movimento etc) podem ser usadas para reconhecimento?
 - Quais partes da imagem devem ser reconhecidas juntas?
 - Como reconhecer objetos sem focalizar em detalhes?
 - Como objetos que tem muitos parâmetros livres podem ser reconhecidos?
 - Como estruturar base de dados de modelos de objetos?

Cores



Texturas



De David Forsyth, UC Berkeley

Cor, textura e proximidade



De Fei-Fei Li

Segmentação

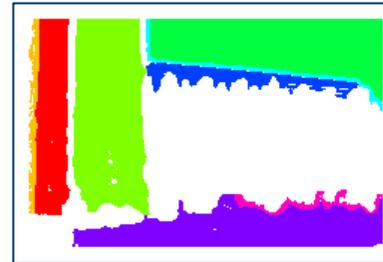
Imagens originais



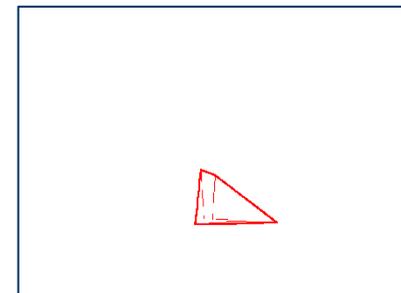
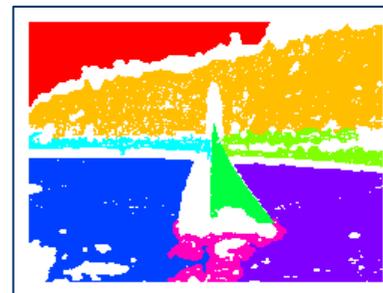
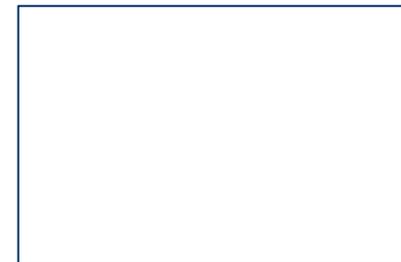
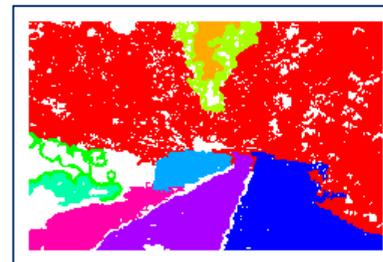
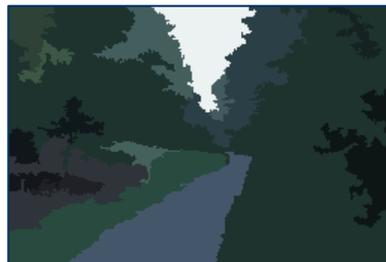
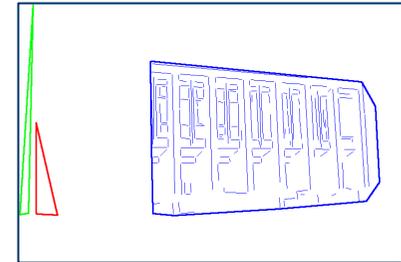
Regiões de cores



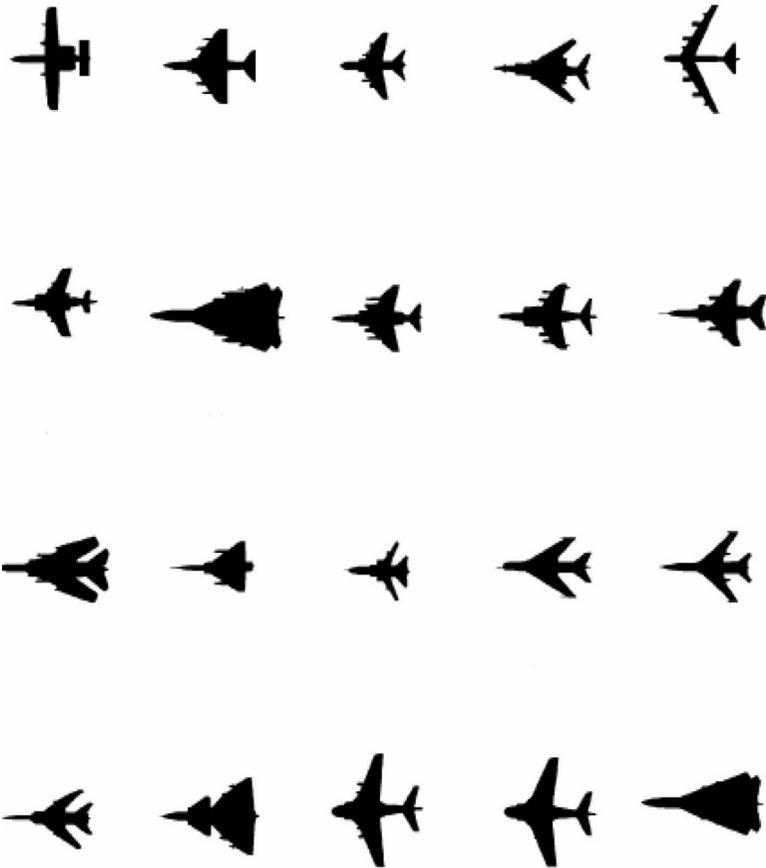
Regiões de texturas



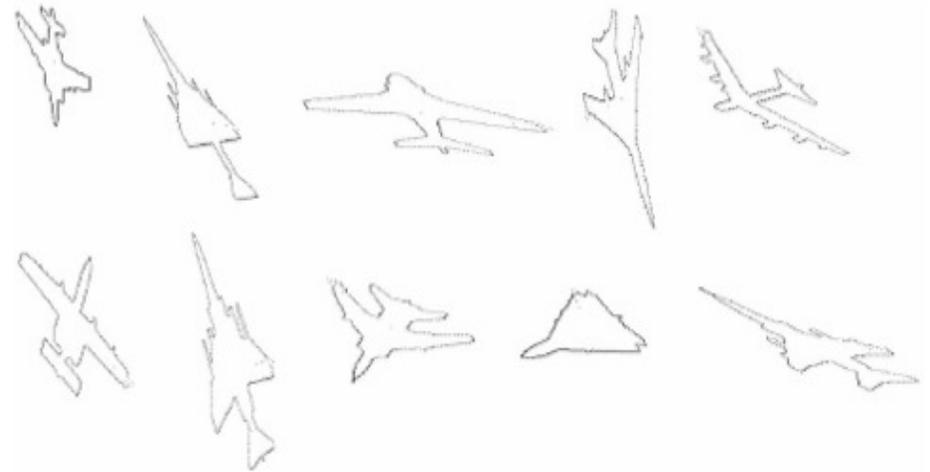
Agrupamentos



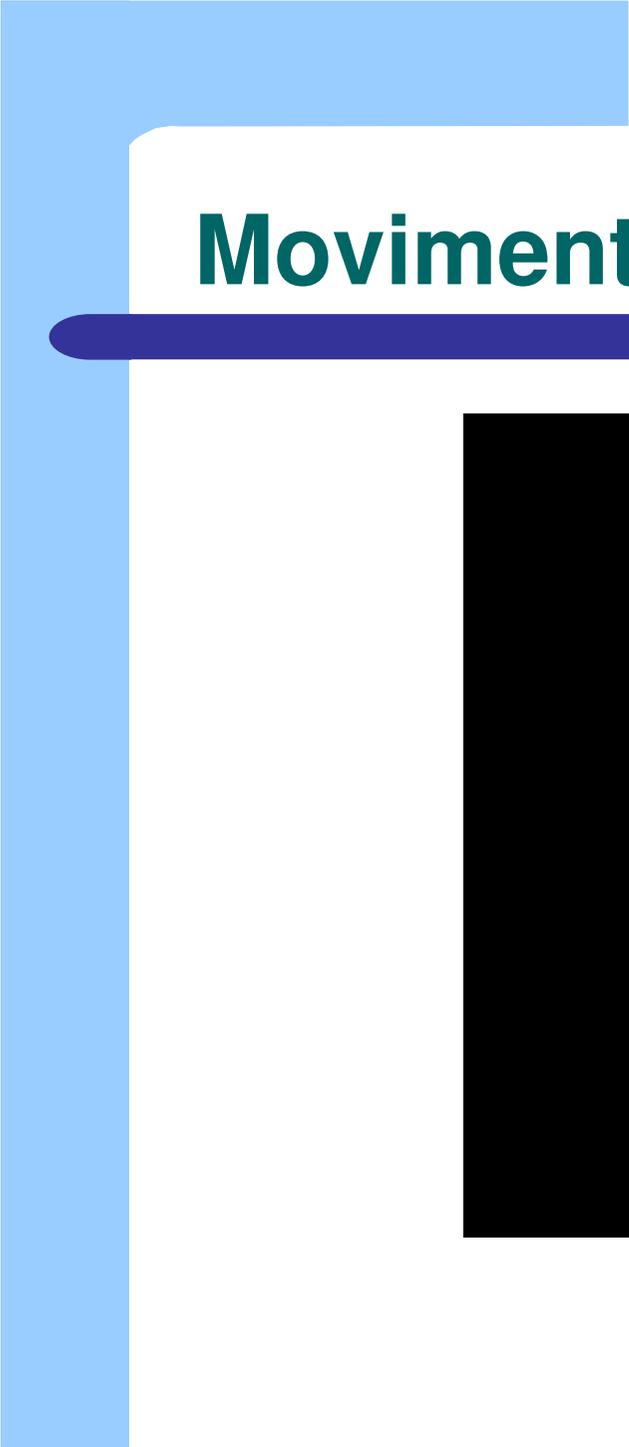
Formato



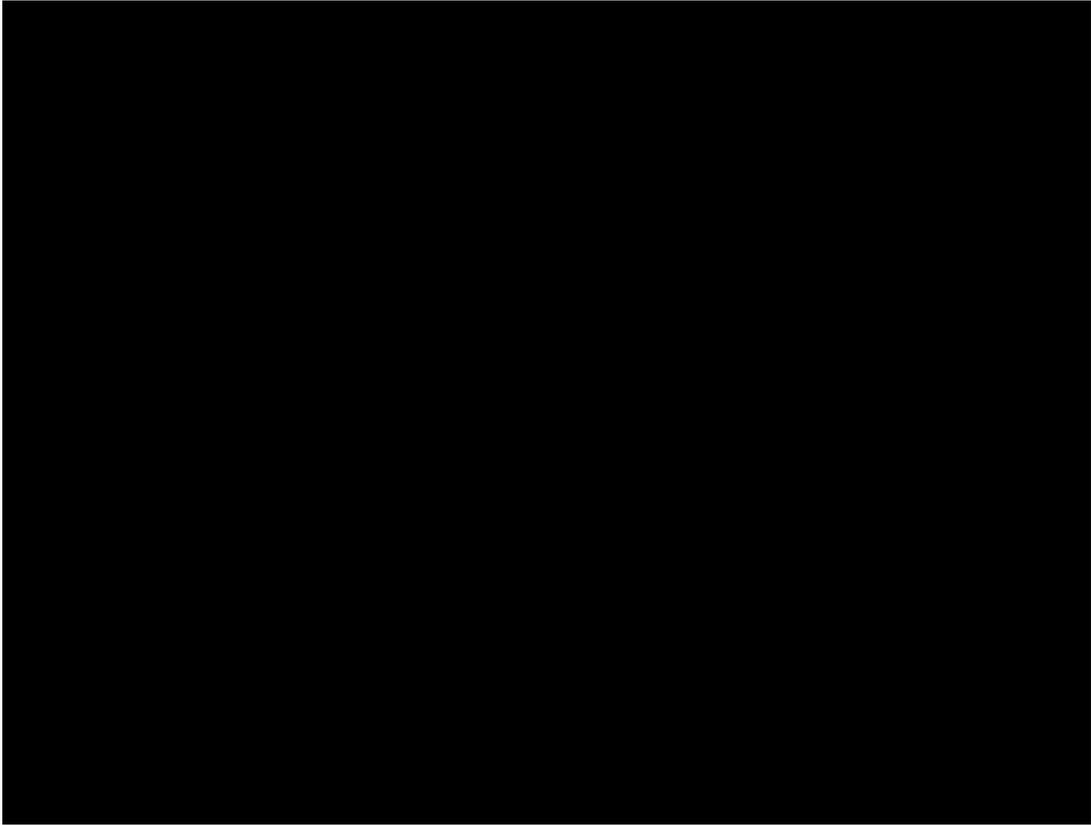
Base de datos de modelos



Objetos reconhecidos



Movimento



Detecção e reconhecimento

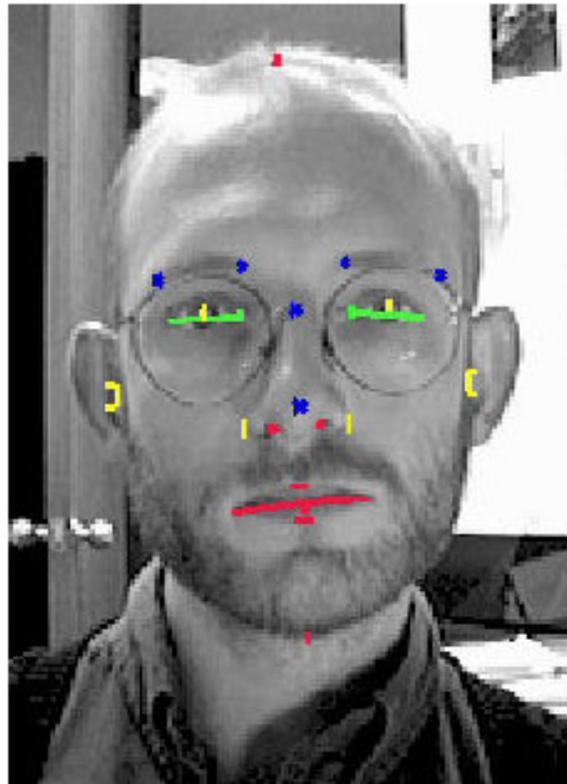
Quais são os modelos usados?

Esses modelos são bons?

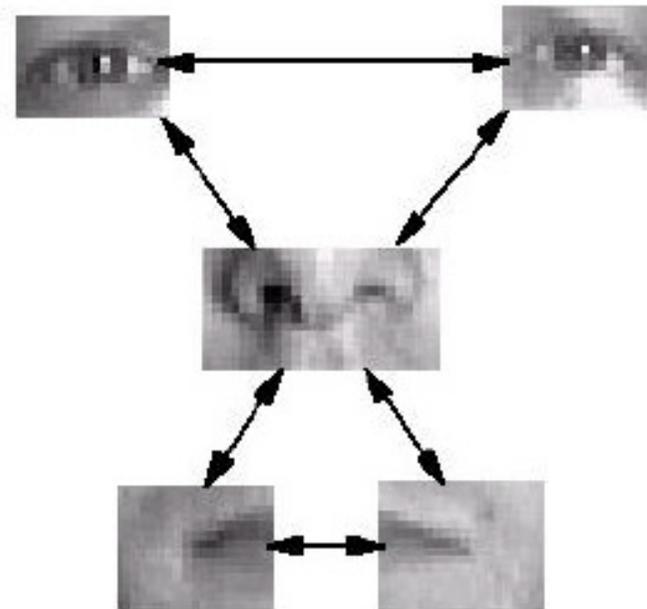


De Michael Black, Brown University

Partes e relações

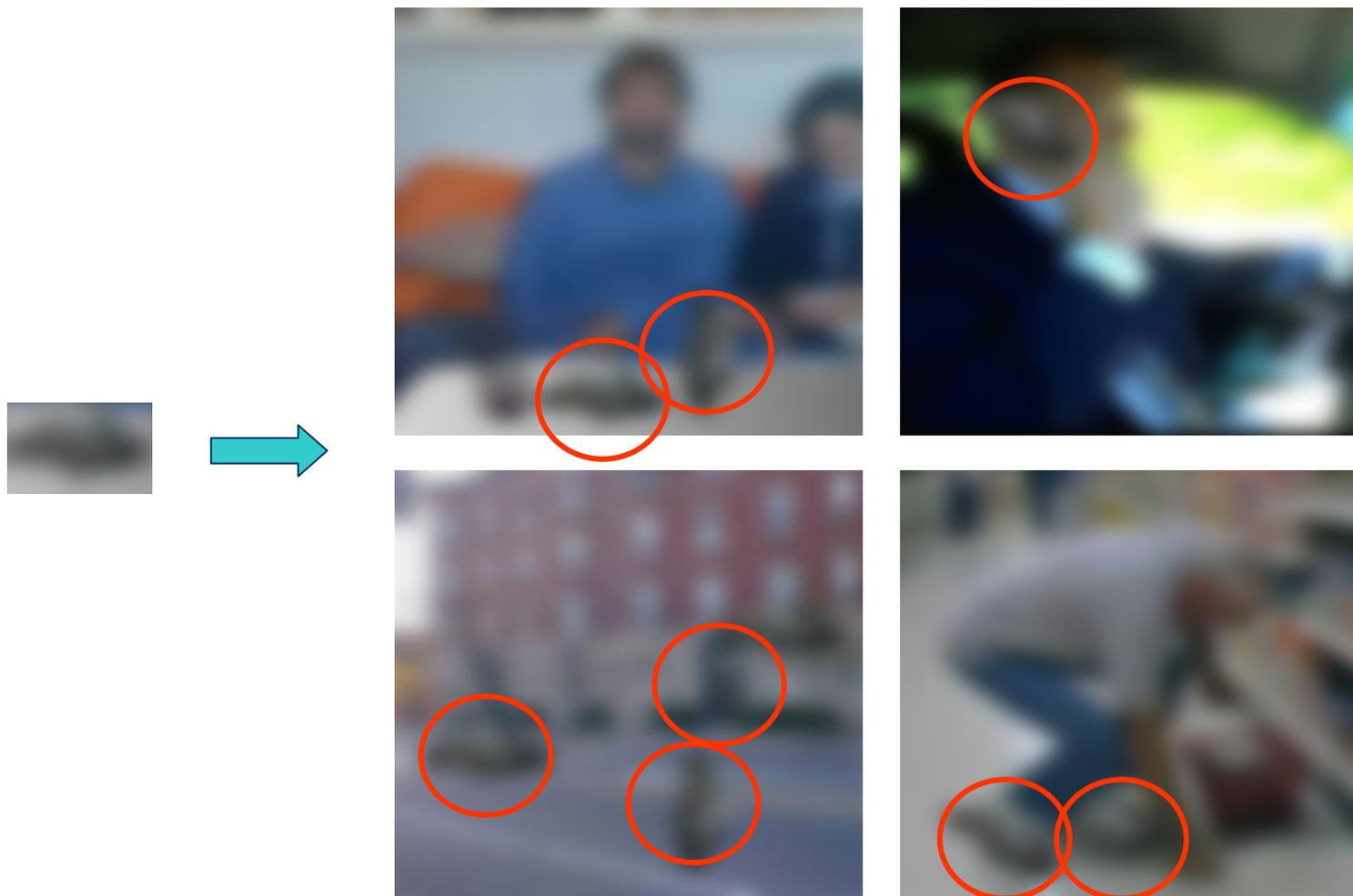


Patch Model



<http://www.research.ibm.com/ecvg/biom/facereco.html>

Contexto e similaridade



Contexto e similaridade



Etapas do processamento de imagens

- Baixo nível:
imagem → imagem
- Nível médio:
imagem → características e atributos
- Alto nível:
características → reconhecimento

Baixo nível

maior nitidez



menor nitidez

Nível médio



imagem original

Detecção
de bordas
→



image de bordas



imagem de bordas

Características
→



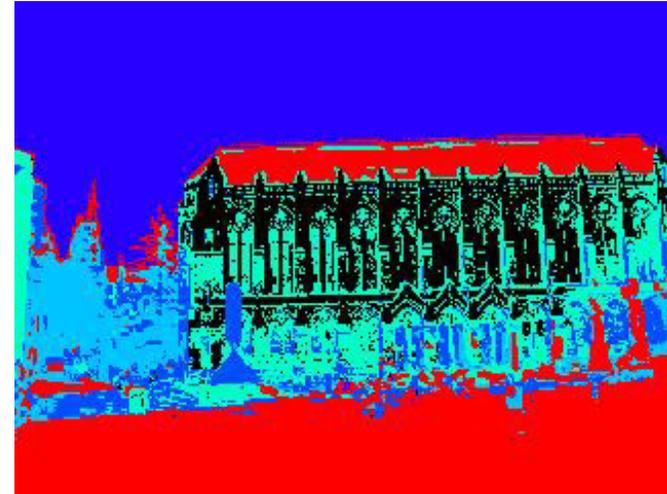
arcos e segmentos de retas

Nível médio



imagem colorida original

Detecção
de manchas
coloridas



regiões homogênea de cor

Alto nivel



Alto nível

Classificação

Existe um carro na imagem?



Alto nível

Detecção

Onde está o carro na imagem?



Alto nivel

Estimativa de pose:



Alto nível

Categorização
de objeto:



Alto nível

Segmentação:

Céu

Árvore

Carro

Pessoa

