

Digitális képfeldolgozás

Oktató:

Palágyi Kálmán
SZTE, Képfeldolgozás és Számítógépes Grafika Tanszék
Árpád tér 2., 214-es szoba (tetőtér)
Tel: (62) 546 197
E-mail: palagyi@inf.u-szeged.hu

Kurzusanyagok:

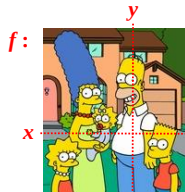
/pub/Digitalis_kepfeldolgozas

1. Bevezetés

Kép

Egy 2D képet a $f(x,y)=v$ függvény ír le, ahol:

- (x,y) egy síkbeli pont koordinátája és
- v a kép intenzitása / szürkességi szintje / színe az (x,y) pontban.



$f(x,y) = \text{„fehér”}$

Digitális kép

Az $f(x,y)=v$ egy 2D a digitális kép, ha x, y és v egyaránt véges és diszkrét.

f



$x = 1, 2, \dots, 320$

$y = 1, 2, \dots, 240$

$v = 0, 1, \dots, 128, \dots, 255$

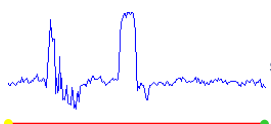
↑
fekete

↑
közép-
szürke

↑
fehér



Digitális kép



255

szürkességi szint/érték

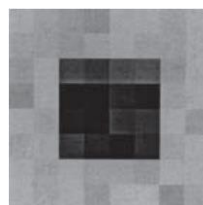
0

Digitális kép

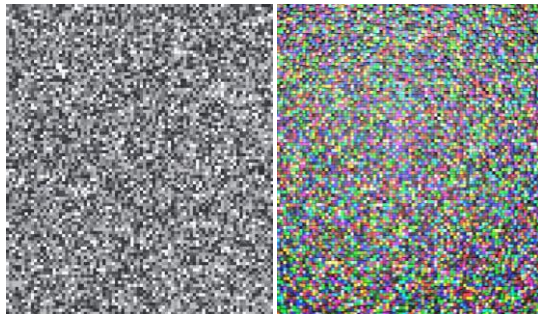
A digitális kép egy pontja, elemének elnevezése

- **pixel** (picture element, 2D) vagy
- **voxel** (volume element, 3D).

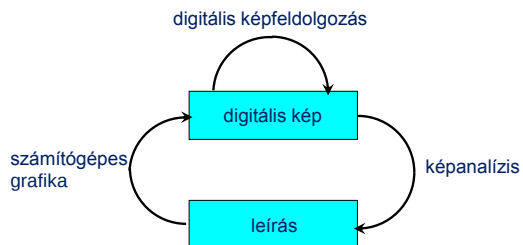
44	43	42	41	44	44	43	44
43	43	42	43	44	43	43	44
42	44	7	6	8	9	42	44
41	44	0	1	4	3	43	44
44	43	2	1	5	6	44	42
43	44	6	4	3	5	43	44
44	43	44	43	43	42	44	44
44	43	44	42	44	44	41	43



A digitális kép ábrázoljon valamit



A képfeldolgozás és társterületei



A képfeldolgozás és társterületei

Számítógépes képfeldolgozás:

- képi információ feldolgozása,
- az eljárások/problémák bemenete és kimenete is kép.

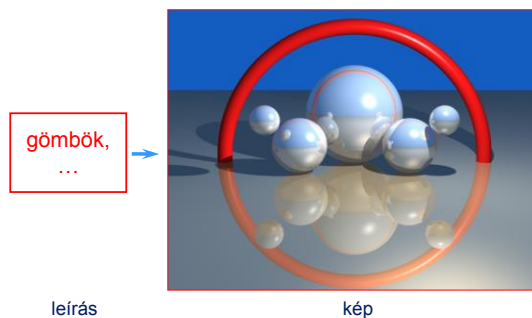
Képi alakfelismerés, képmegértés, számítógépes látás:

- mit ábrázol a kép,
- osztályozás.

Számítógépes grafika:

- képek előállítása, megjelenítése.

Számítógépes grafika



Virtuális / kiterjesztett / kevert valóság (Virtual / Augmented / Mixed Reality)



Virtuális / kiterjesztett / kevert valóság (Virtual / Augmented / Mixed Reality)



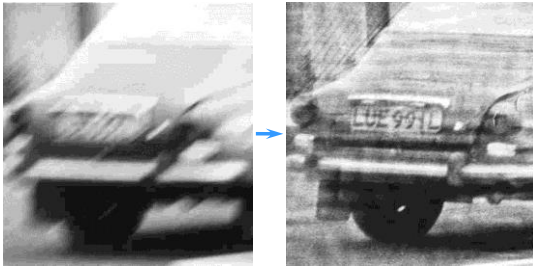
Képanalízis



El Greco
(Domenikos Theotokopoulos):
Toledo látképe
(1597-99)

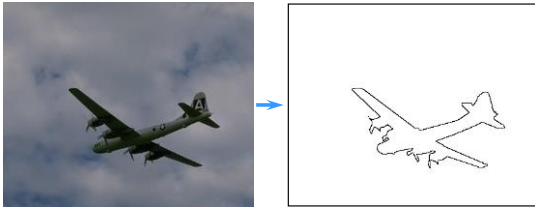
kép leírás

Számítógépes képfeldolgozás



kép helyreállítás kép

Számítógépes képfeldolgozás



kép szegmentálás kép

A számítógépes képfeldolgozás alapvető céljai

- A vizuális információ minőségének javítása az emberi értelmezés számára.
- A képi adatok feldolgozása az autonóm számítógépes látás elősegítéséhez.

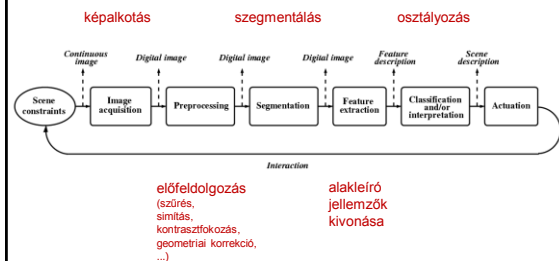
Támaszkodás más tudományterületekre

- függvénytranszformációk
- differenciál- és integrálszámítás
- valószínűségszámítás és matematikai statisztika
- geometria
- topológia
- algebra
- halmazelmélet
- gráfelmélet
- operációkutatás
- mesterséges intelligencia
- szoftverfejlesztés
- algoritmus-tervezés
- számítástudomány
- fizika
- ...

Megtámasztott tudományterületek

- biológia
- orvostudomány
- fizika
- csillagászat
- űrkutatás
- anyagtudomány
- agrártudományok
- archeológia
- geológia és geográfia
- ...

A képanalízis (gépi látás) moduláris rendszere



A képanalízis területei

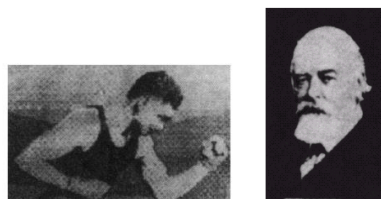
Specific topics of interest include, but are not limited to:

- * 3D from Multi-view and Sensors
- * 3D Point Clouds
- * 3D from Single Images
- * 3D Reconstruction
- * Action Recognition, Understanding
- * Adversarial Learning
- * Biologically Inspired Vision
- * Biomedical Image Processing
- * Biometrics
- * Computational Photography
- * Computer Vision for General Medical, Biological and Cell Microscopy
- * Computer Vision Theory
- * Datasets and Evaluation
- * Deep Learning: Applications, Methodology, and Theory
- * Document Analysis RGBD Sensors and Analytics
- * Face, Gesture, and Body Pose
- * Human Computer Interaction
- * Image and Video Synthesis
- * Large Scale Methods
- * Low-level Vision
- * Machine Learning,
- * Motion and Tracking

A képanalízis területei

- * Optimization Methods
- * Physics-based Vision and Shape-from-X Recognition: Detection, Categorization, Retrieval
- * Pose Estimation
- * Recognition: Detection, Categorization, Indexing and Matching
- * Remote Sensing and Hyperspectral Imaging
- * Representation Learning
- * Robotics and Driving Scene Analysis
- * Scene Understanding
- * Security/Surveillance
- * Semi- and weakly-supervised Learning
- * Segmentation, Grouping and Shape
- * Statistical Learning
- * Stereo/Depth Estimation
- * Tracking
- * Transfer Learning
- * Unsupervised Learning
- * Video Analytics Vision + Graphics Vision + Language
- * Virtual and Augmented Reality
- * Vision and Language
- * Vision for Robotics, Graphics
- * Visual Reasoning Vision Applications and Systems

Az első digitális képek



1921

1922

A korai Bartlane-rendszer 5 árnyalattal kódolt.
A fotókat távíró lyukszalagról rekonstruálták.

Az első digitális képek



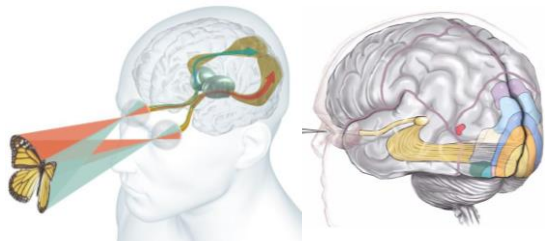
1929,
15 árnyalat,
(London – New York)
tengeralatti kábel
továbbították

Az első digitális képek

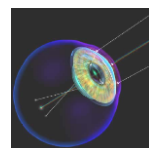


Ranger 7 (1964)

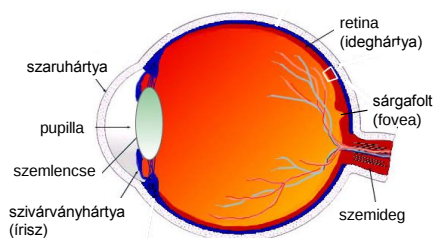
Az emberi látás



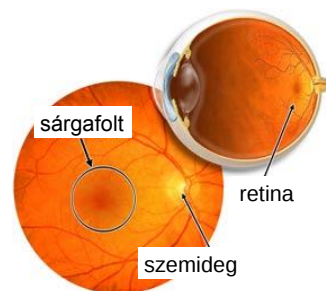
A kameránk



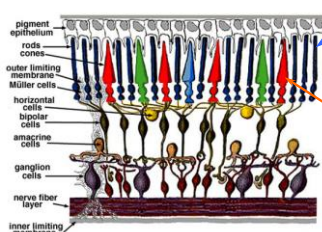
A kamera alkatrészei



Szenzortömb - retina



Szenzorok

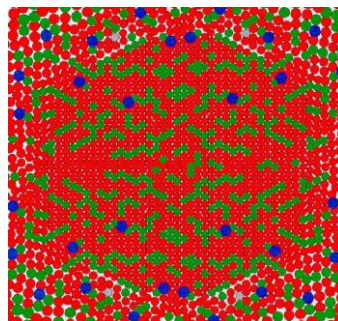


pálcika
(120-130 millió,
az esti látás
érzékeny elemei)

csap
(5-7 millió,
a színlátás kevésbé
érzékeny elemei, 3
különböző pigmentet
tartalmaznak – RGB)

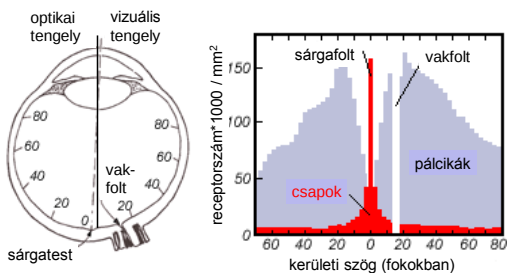
A csapok színérzékenysége: 65%, 33%, 2%

A csapok színérzékenysége



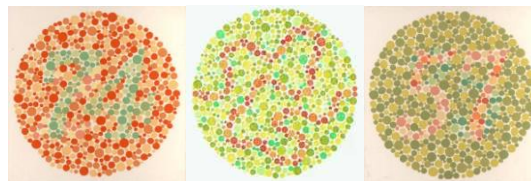
65%
33%
2%

A csapok és a pálcikák eloszlása



https://index.hu/tudomany/til/2018/10/13/hany_megapixeles_a_szemunk

A színvakság, szintévesztés tesztképei

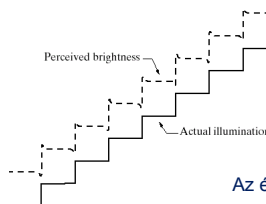


Európában a férfiak 8%-a, a nők 0.5%-a szintévesztő.

Képmozaik

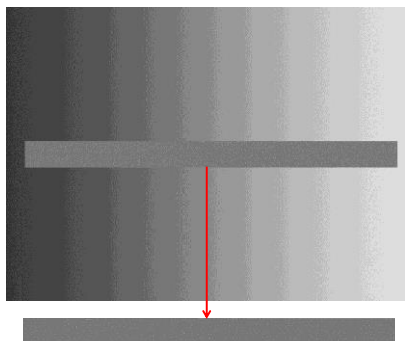


Mach-féle hatás



Az éleket eltűzve érzékeljük.

Mach-féle hatás



Előtér vagy háttér?



René Magritte:
La firma in bianco
(The blank signature)
1965

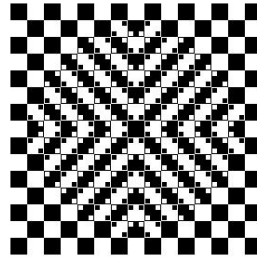
Relatív képtartalom



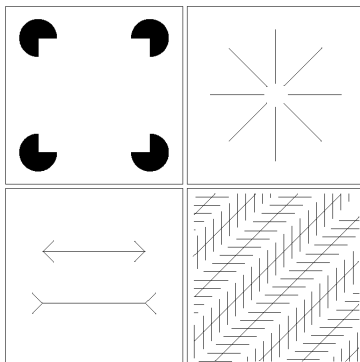
orr vagy bibircsók ?



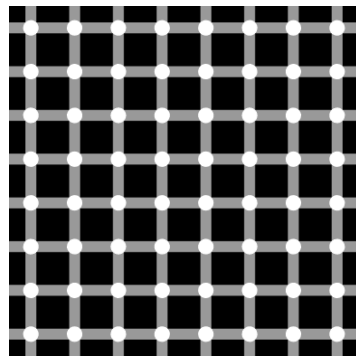
Optikai csalódások



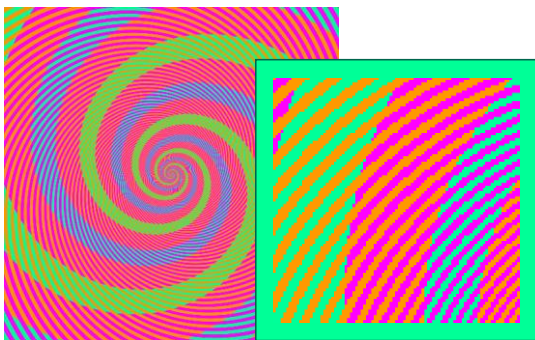
Optikai csalódások



Optikai csalódások

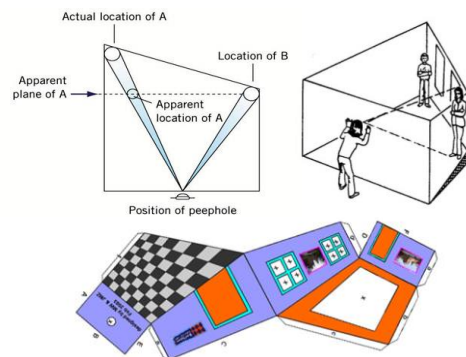


Optikai csalódások

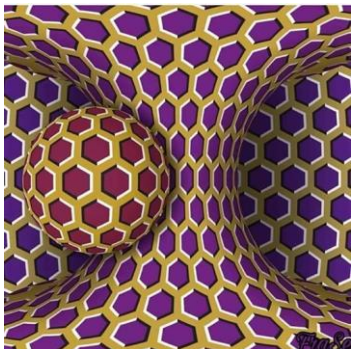


<http://www.ritsumeai.ac.jp/~akitaoka/>

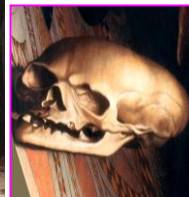
Ames room



Optikai csalódások

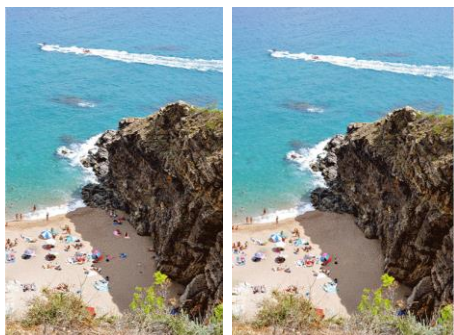


Optikai csalódások



Anamorfikus optikai illúzió – Hans Holbein (1533)

Változásvakság (change blindness)



Margaret Thatcher illusion



Margaret Thatcher illusion



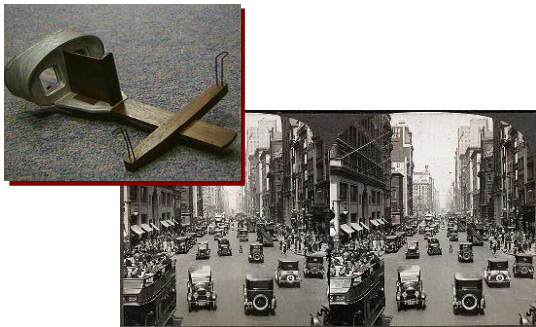
Optikai illúziók



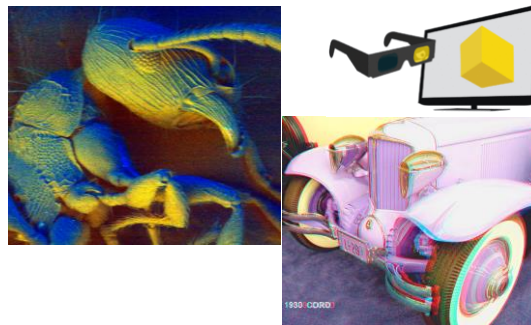
www.moillusions.com

www.eyetricks.com

Sztereoó képek



Sztereoó képek



Sztereogram



www.magiceye.com

Hologramok

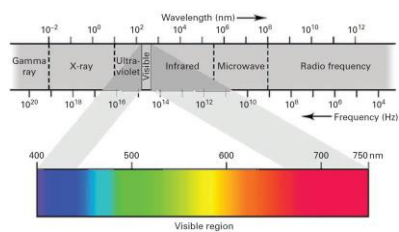


3D TV



Látásunk tulajdonságai

Érzékeljük az elektromágneses színekép 400-800 nm sávját, továbbá (a csapoknak köszönhetően) van színlátásunk.



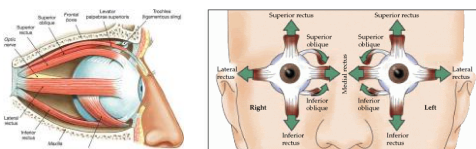
Látásunk tulajdonságai

Szabályozható a bejutó fény erőssége: a pupilla erős fényben összehúzódik, gyenge fényben pedig kitágul.



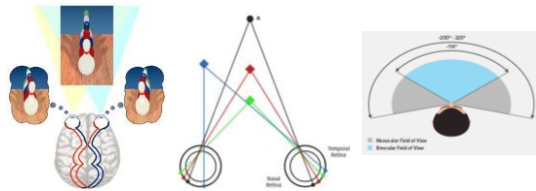
Látásunk tulajdonságai

A szemmozgások lehetővé teszik a fókuszálást, a tárgyaltást és a mozgáskövetést.



Látásunk tulajdonságai

A sajátosan elhelyezkedő két szem segíti a térlátást (a mélység és a távolság észlelését).



A látás fejlődik és (a tárgyak észlelése, felismerése, értelmezése) tanult folyamat.

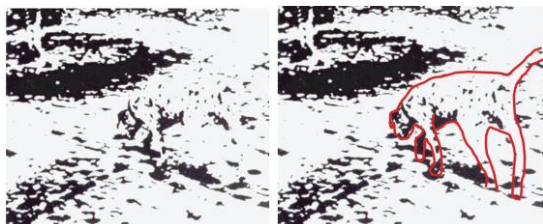
Látásunk tulajdonságai

Változatos bonyolultságúak a feladatok.
Hány X van a képen?



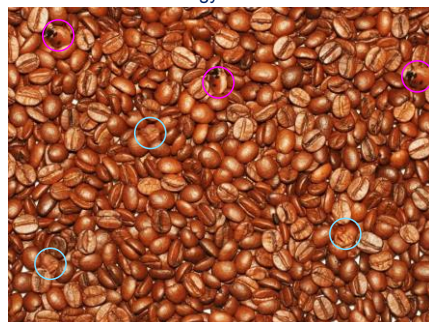
Látásunk tulajdonságai

Képesek vagyunk globális értelmezésre.



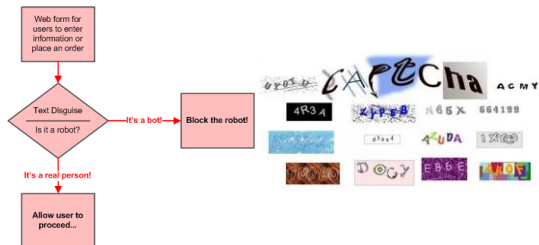
Látásunk tulajdonságai

Szelektálás és koncentráció:
az információáradat egyes elemeinek kiemelése.



Látásunk tulajdonságai

Képesek vagyunk bonyolult absztrakciókra.



A gépi látás fejlődése

Deep Visual-Semantic Alignments for Generating Image Descriptions

Andrej Karpathy Li Fei-Fei
Department of Computer Science, Stanford University
{karpathy, feifei}@cs.stanford.edu

Abstract

We present a model that generates natural language descriptions of images and their regions. Our approach leverages datasets of images and their sentence descriptions to learn about the inter-modal correspondences between language and visual data. Our alignment model is based on a novel combination of Convolutional Neural Networks over image regions, bidirectional Recurrent Neural Networks over sentences, and a structured objective that aligns the two modalities through a multimodal embedding. We then describe a Multimodal Recurrent Neural Network architecture that uses the inferred alignments to learn to generate novel descriptions of image regions. We demonstrate that our alignment model produces state of the art results in retrieval experiments on Flickr8K, Flickr30K and MSCOCO datasets. We then show that the generated descriptions significantly outperform retrieval baselines on both full images and on a new dataset of region-level annotations.



Figure 1. Motivation/Concept Figure: Our model treats language as a rich label space and generates descriptions of image regions.

CVPR 2015 : 28th IEEE Conf. Computer Vision and Pattern Recognition

A gépi látás fejlődése



Human: "A group of men playing Frisbee in the park."

Computer model: "A group of young people playing a game of Frisbee."

A gépi látás fejlődése



Human: "A green monster kite soaring in a sunny sky."

Computer model: "A man flying through the air while riding a snowboard."

A kurzus teljesítése

Szóbeli vizsga:

- ~10-perces turnusokban
- max. 5 fő