Computer Vision

Introduction to CV

Lecture on Segmentation – Mean Shift, EM, K-Means

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

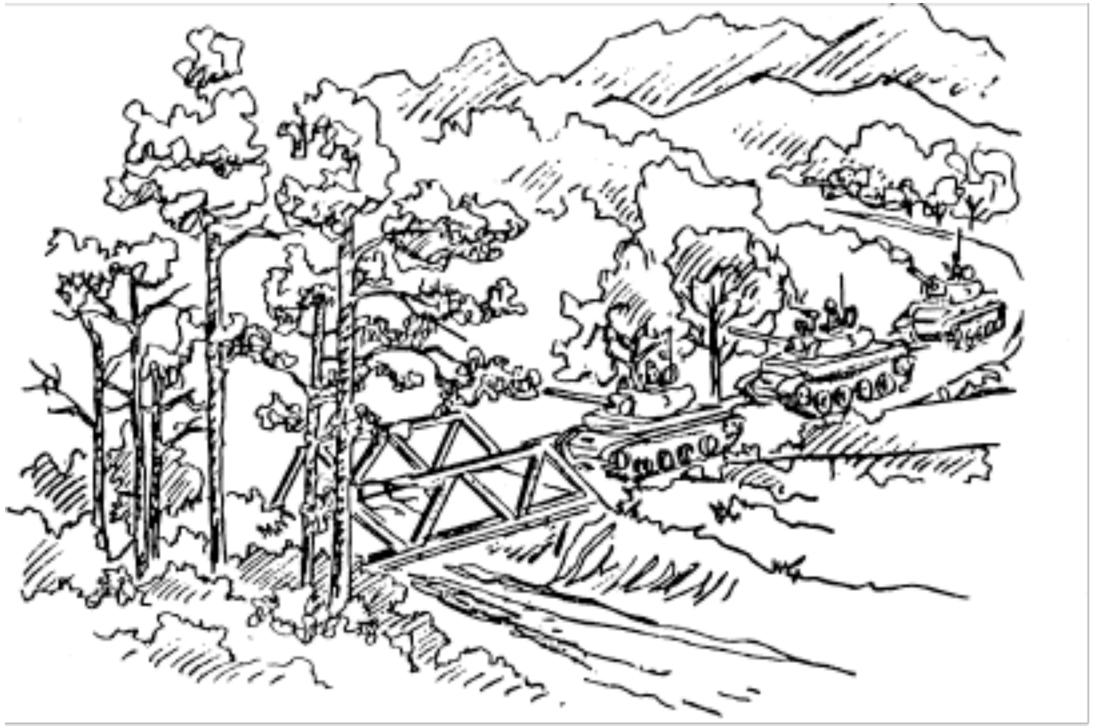
Artificial Intelligence Group July 7, 1966
Vision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

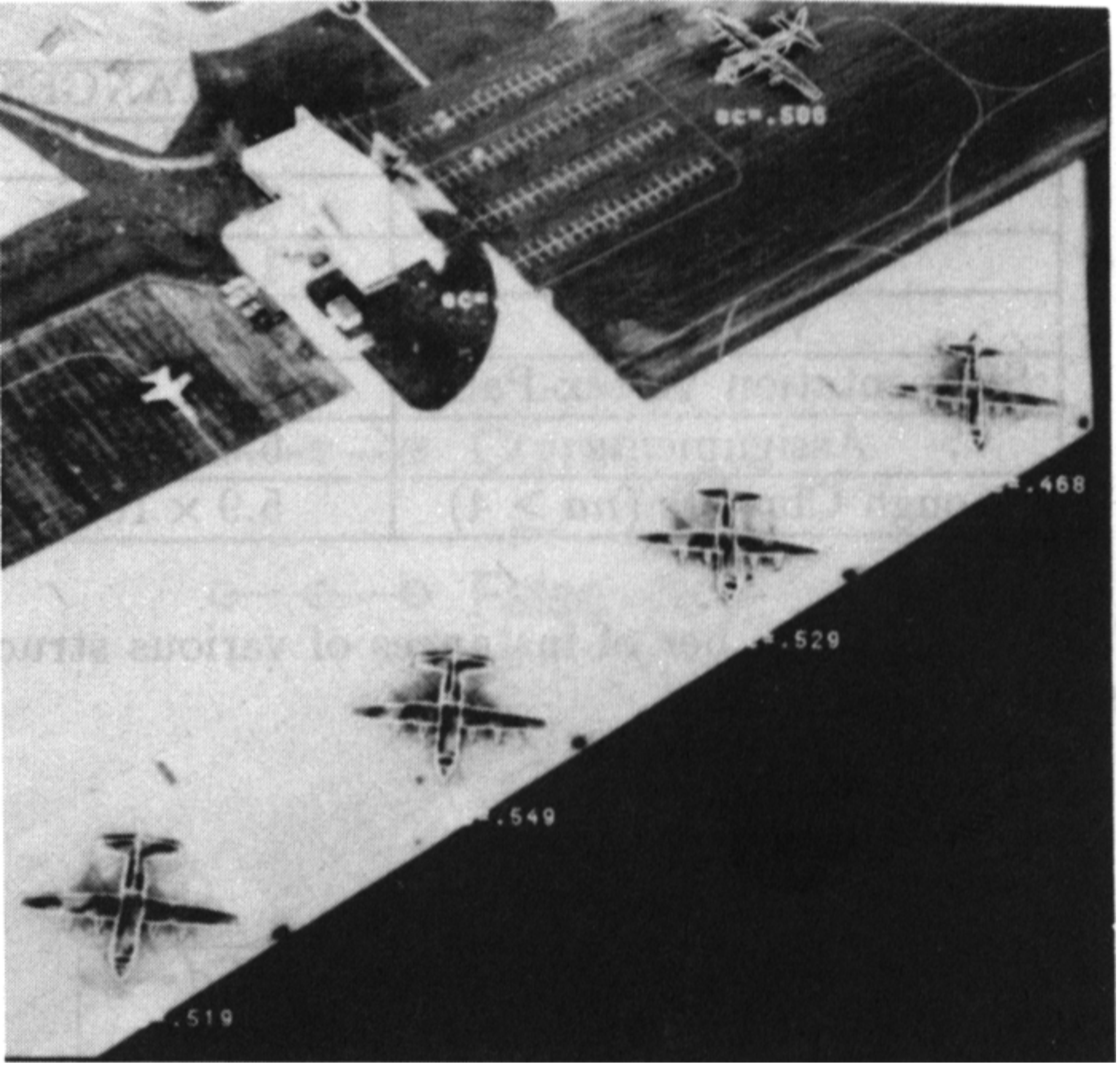
The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the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

The DARPA Image Understanding Program began in 1975



Tank Convoy Is Approaching Bridge Over Elbe River On Route #410. The Elbe River Is the Border Between East and West Germany. Route #410 Leads From Neuhaus.

Druffel L.E. (1982) Summary of the DARPA Image Understanding Research Program. In: Kittler J., Fu K.S., Pau LF. (eds) Pattern Recognition Theory and Applications. NATO Advanced Study Institutes Series (Series C — Mathematical and Physical Sciences), vol 81. Springer, Dordrecht

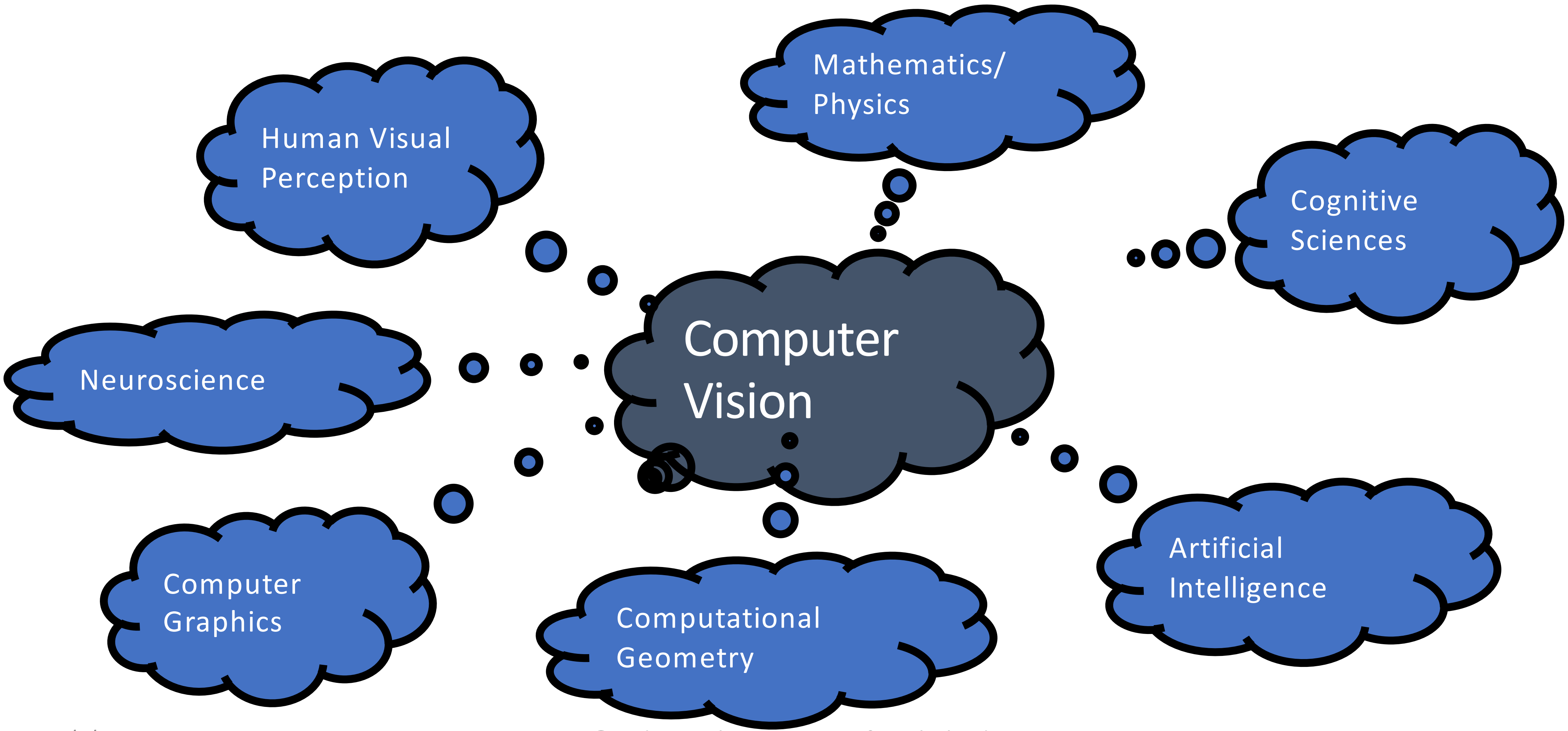


Brooks, R. A., Creiner, R., & Binford, T. O. (1979, August). The ACRONYM model-based vision system. In *Proceedings of the 6th international joint conference on Artificial intelligence-Volume 1* (pp. 105-113). Morgan Kaufmann Publishers Inc..

Complexity of Human Vision

- 1/3 of the brain devoted to perception
  - Eye’s just the front end.
- Most of the processing not available for introspection
- Most reliable sensory mode
- Human vision
  - Collection of slapped together modules doing various kinds of tasks
    - e.g. face recognition

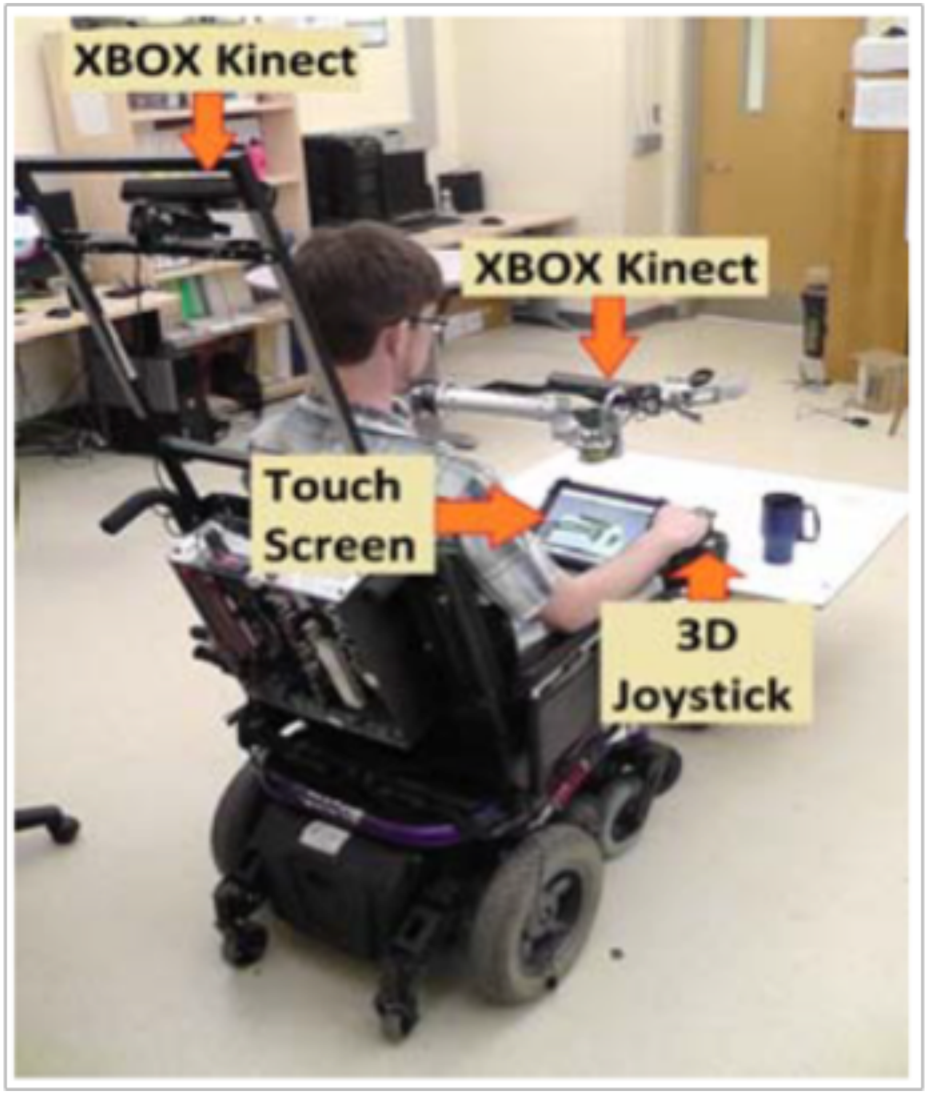
CV draws from many disciplines



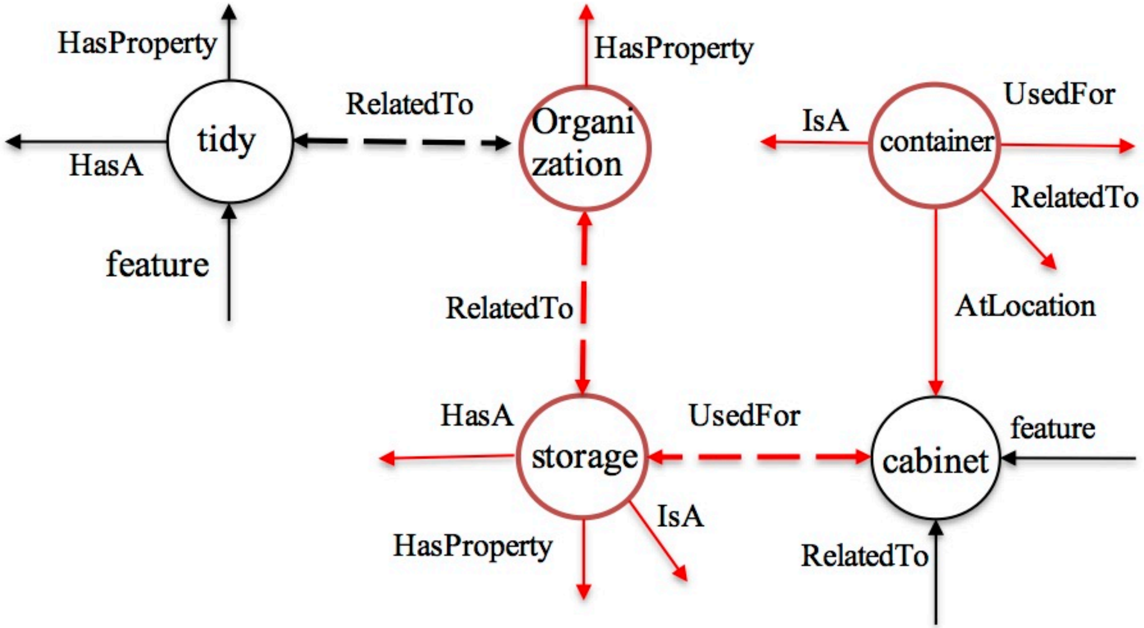
AI that anticipate our intentions

- Achieving Autonomy by Learning from Sensor-Assisted Control in a Wheelchair-Based Human-Robot Collaborative System.
- National Science Foundation (NSF) project

- Rajiv Dubey, Redwan Alqasemi, Kyle Reed (Robotics) and CSE faculty Sudeep Sarkar (Computer Vision)



AI that assist in daily activities



Generating Open World Descriptions of Video using Commonsense Knowledge in a Pattern Theory Framework
Sathyanarayanan Aakur, Fillipe DM de Souza, Sudeep Sarkar
Quarterly of Applied Mathematics, 2019

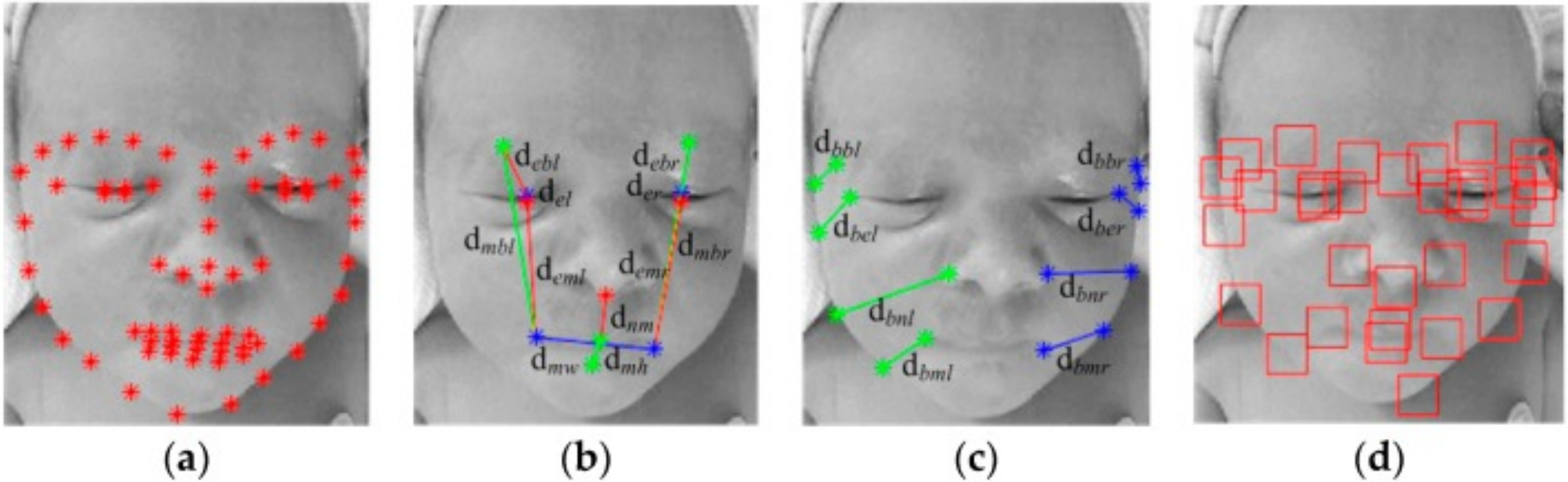
Deep Learning AI detect and classify facility, buildings, and land use from satellite imagery.

Urban planning, food security, transportation, disaster response

USF placed 3rd in IARPA Functional Map of the World challenge and 3rd in the DoD sponsored xView challenge

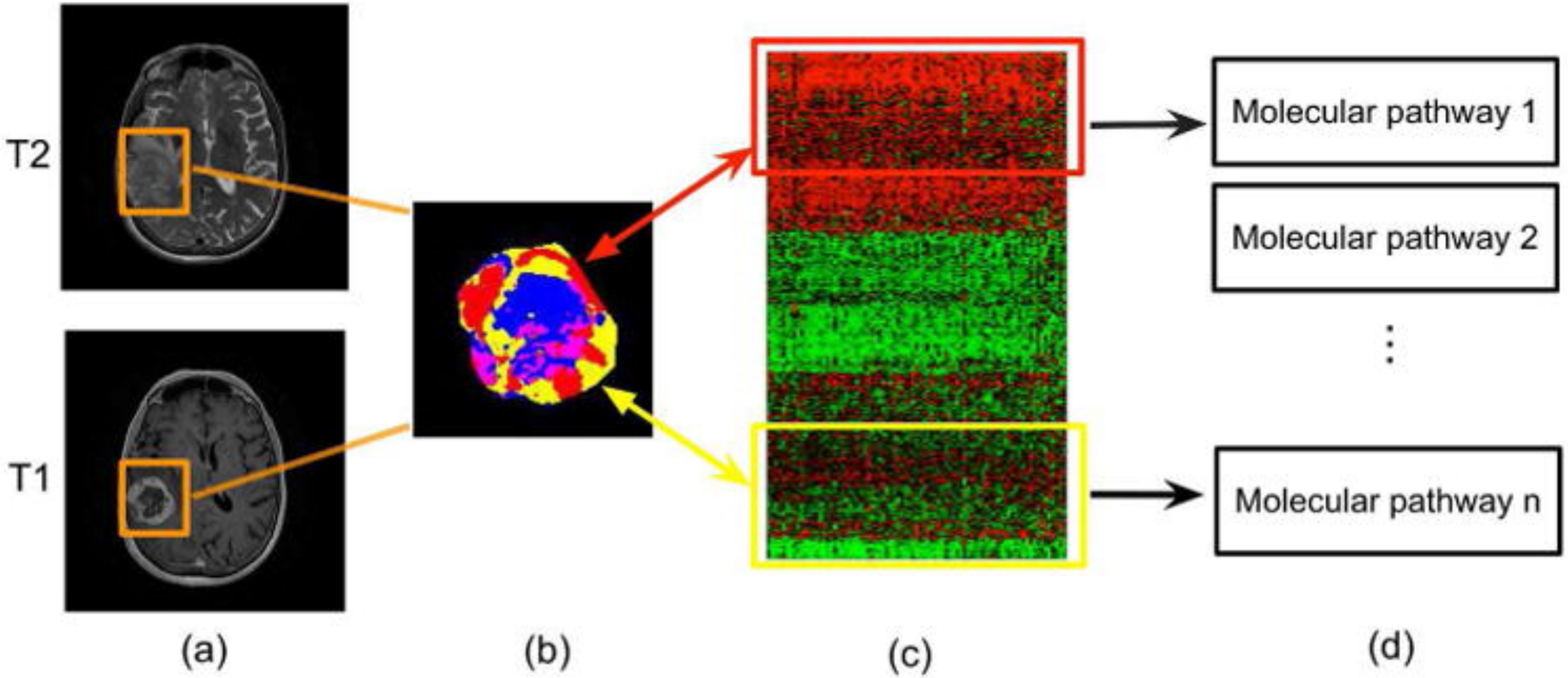
Hydra: an Ensemble of Convolutional Neural Networks for Geospatial Land Classification
R Minetto, MP Segundo, S Sarkar, arXiv preprint arXiv:1802.03518

AI that monitor pain in infants



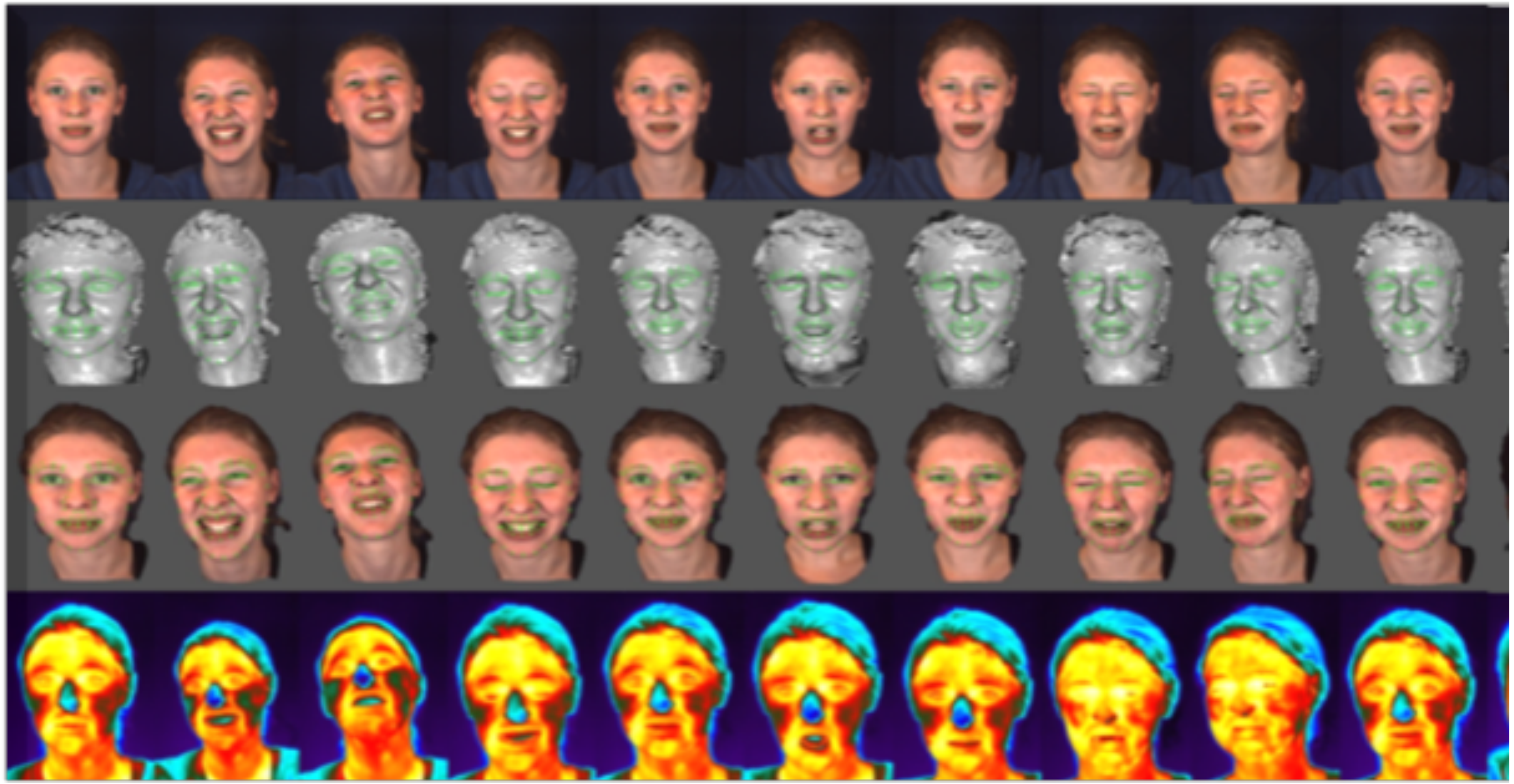
Zhi, R., Zamzmi, G., Goldgof, D., Ashmeade, T., & Sun, Y. (2018). Automatic Infants' Pain Assessment by Dynamic Facial Representation: Effects of Profile View, Gestational Age, Gender, and Race. Journal of clinical medicine, 7(7), 173. doi:10.3390/jcm7070173

AI that help make accurate cancer diagnosis



M. Zhou, J. Scott, B. Chaudhury, L.O. Hall, D. Goldgof, K.W. Yeom, M. Iv, Y. Ou, J., Kalpathy-Cramer, S. Napel, R. Gillies, O. Gevaert, R. Gatenby, Radiomics in Brain Tumor: Image Assessment, Quantitative Feature Descriptors and Machine-learning Approaches, American Journal of Neuroradiology, American Journal of Neuroradiology Oct 2017,

AI that recognize our emotions
Affective Computing

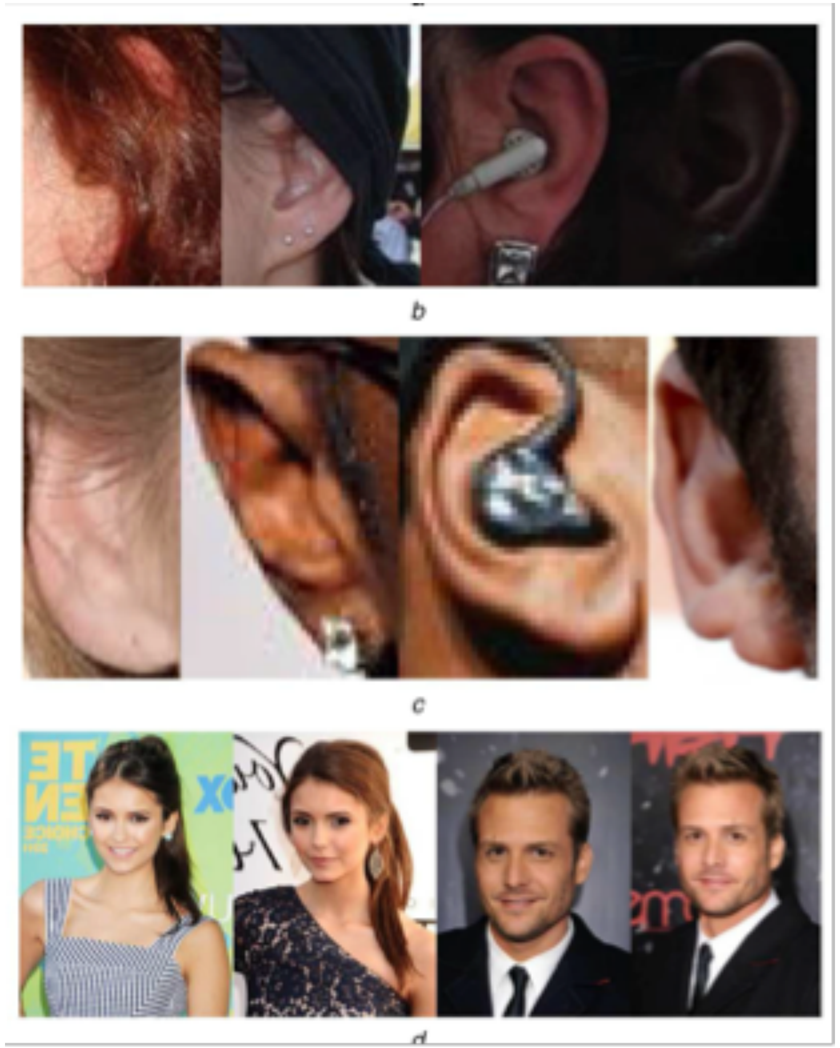


D. Fabiano and S. Canavan, "Spontaneous and Non-Spontaneous 3D facial expression recognition Using a Statistical Model with Global and Local Constraints," *International Conference on Image Processing, 2018*

AI that recognize us from a distance

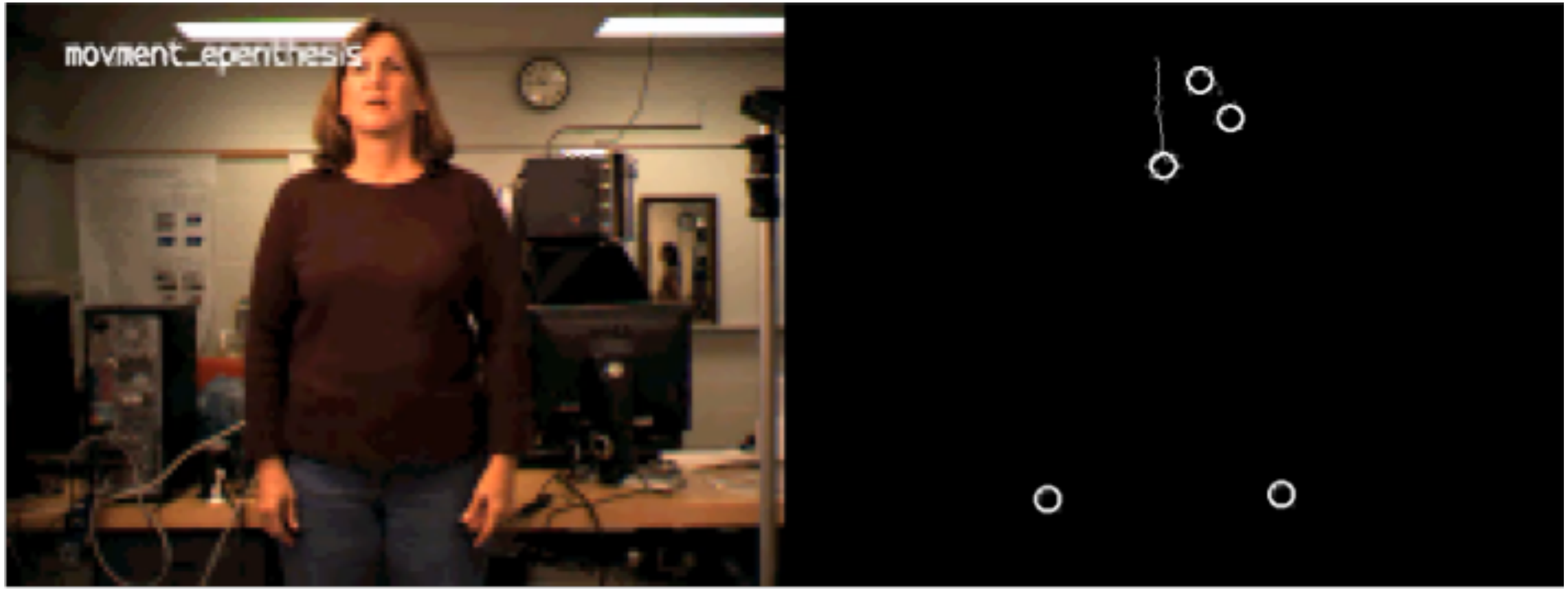
- The HumanID gait challenge problem: Data sets, performance, and analysis, S Sarkar, PJ Phillips, Z Liu, IR Vega, P Grother, KW Bowyer, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 27 (2), 162-177

AI that recognize when we are near



Employing fusion of learned and handcrafted features for unconstrained ear recognition, EE Hansley, MP Segundo, S Sarkar IET Biometrics 7 (3), 215-223

AI that help Communicating with the Deaf



Handling movement epenthesis and hand segmentation ambiguities in continuous sign language recognition using nested dynamic programming
R Yang, S Sarkar, B Loeding,
IEEE transactions on pattern analysis and machine intelligence 32 (3), 462-477


NAE GRAND CHALLENGES FOR ENGINEERING
NATIONAL ACADEMY OF ENGINEERING
Challenges
News
Community
Grand Challenges Report
Advance Personalized Learning
Make Solar Energy Economical
Enhance Virtual Reality
Reverse-Engineer the Brain
Engineer Better Medicines
Advance Health Informatics
Restore and Improve Urban Infrastructure
Secure Cyberspace
Provide Access to Clean Water
Provide Energy from Fusion
Prevent Nuclear Terror
Manage the Nitrogen Cycle
Develop Carbon Sequestration Methods
Engineer the Tools of Scientific Discovery
Reverse-enginee brain
The intersection of engineering and neuroscience promises great advances in health care, manufacturing, and communication.

Problems (NAE Grand Challenges)

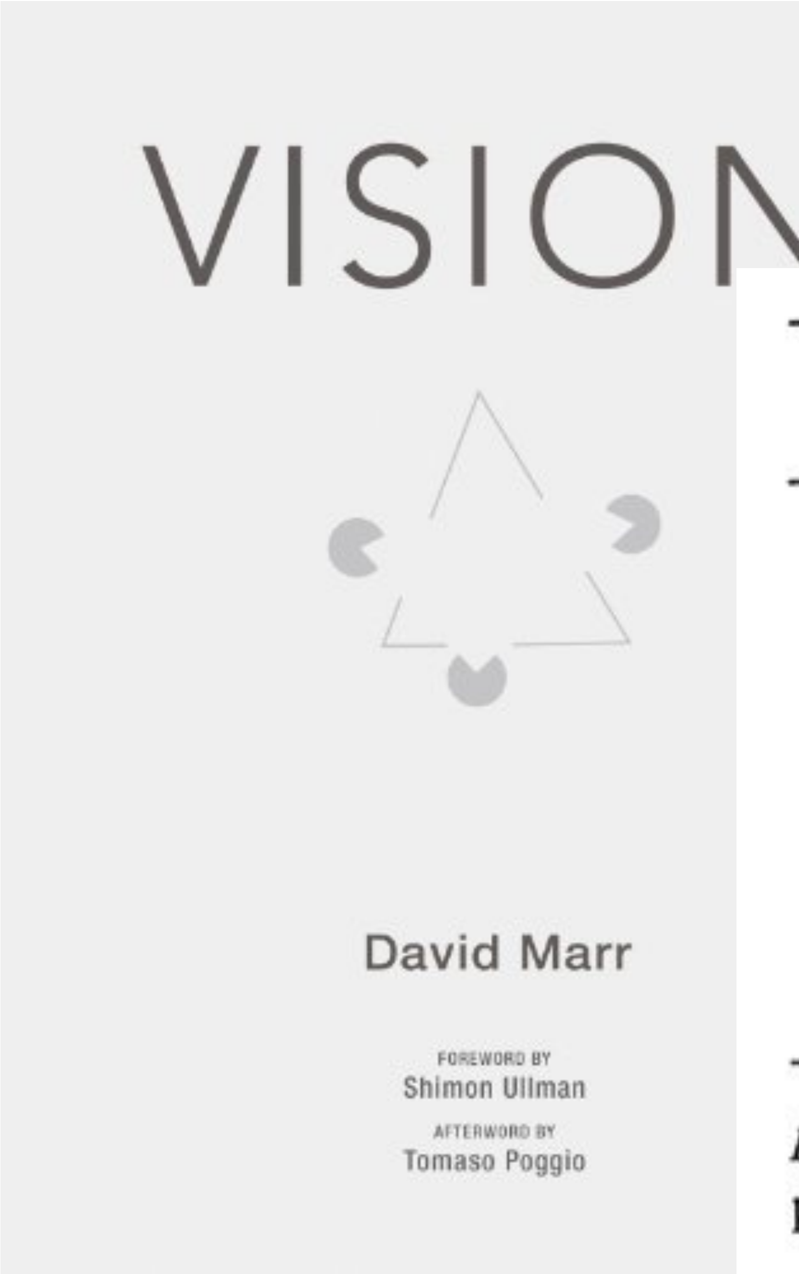
Advanced personalized learning, Enhance virtual reality, Reverse engineering the brain, Engineering better medicine, Advance health informatics, Improve urban infrastructure, secure cyberspace, engineer the tools of scientific discovery,…

Core Vision Problem Solved

Segmentation, 3D Motion estimation, 3D Shape estimation, Symbolic description of content, Recover material property

Mathematical Methods

EM, Particle Filtering, Graph Cuts, Graphical Models, Markov Chain Monte Carlo, Deep Learning, SVD, Linear solvers, etc.



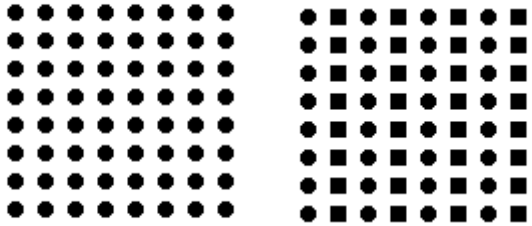
Computational theory	Representation and algorithm	Hardware implementation
What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?	How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?	How can the representation and algorithm be realized physically?

Figure 1–4. The three levels at which any machine carrying out an information-processing task must be understood.

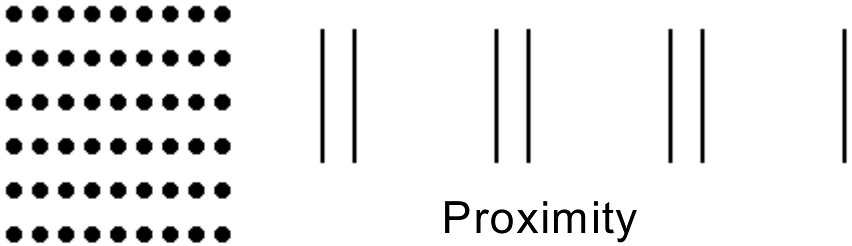
Segmentation

Figure-Ground problem, Mean Shift, EM, Graph cuts

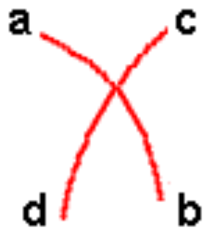
Gestalt Principles



Similarity

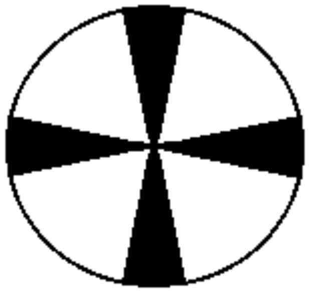


Proximity

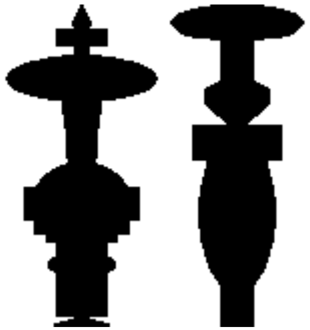


Continuity

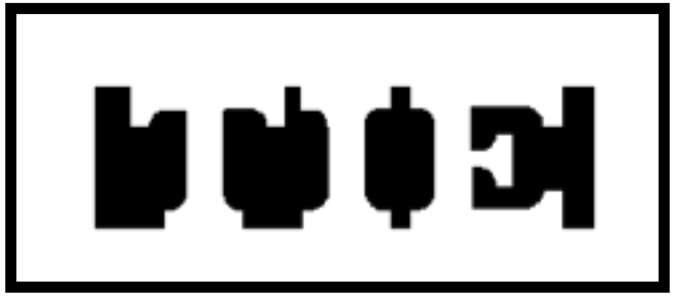
Closure



Small as figure



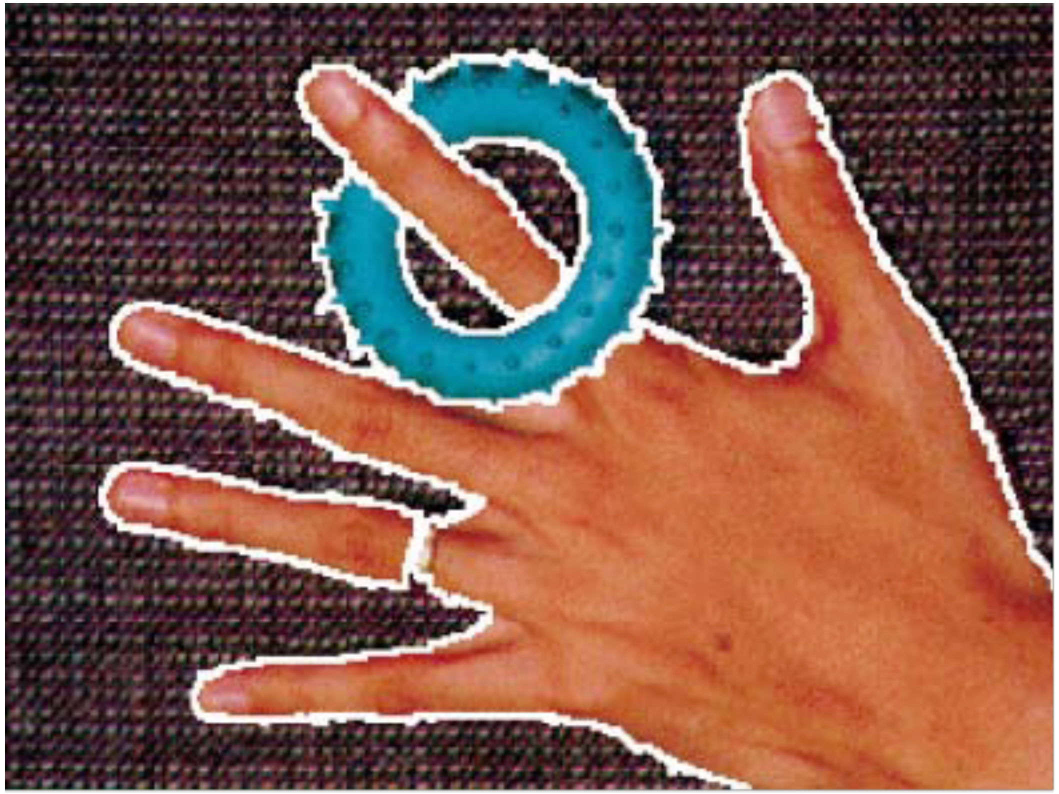
Symmetry



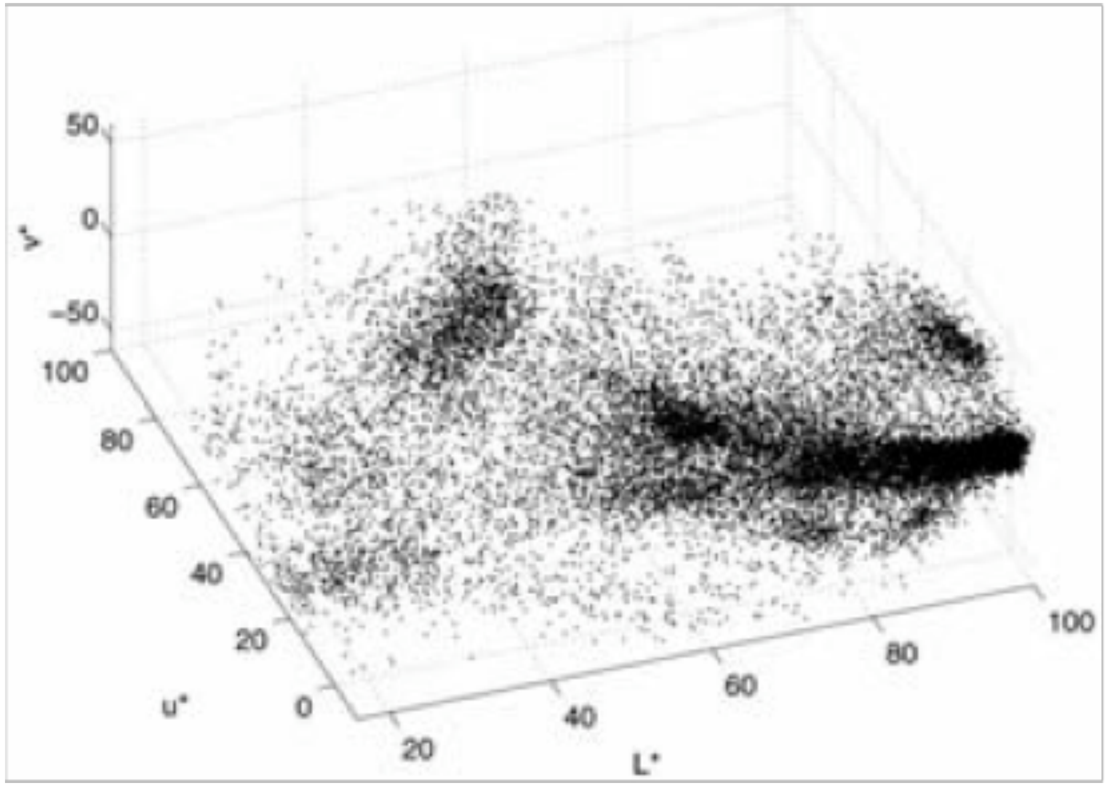
Surroundedness

Taken from Daniel Chandler's webpage

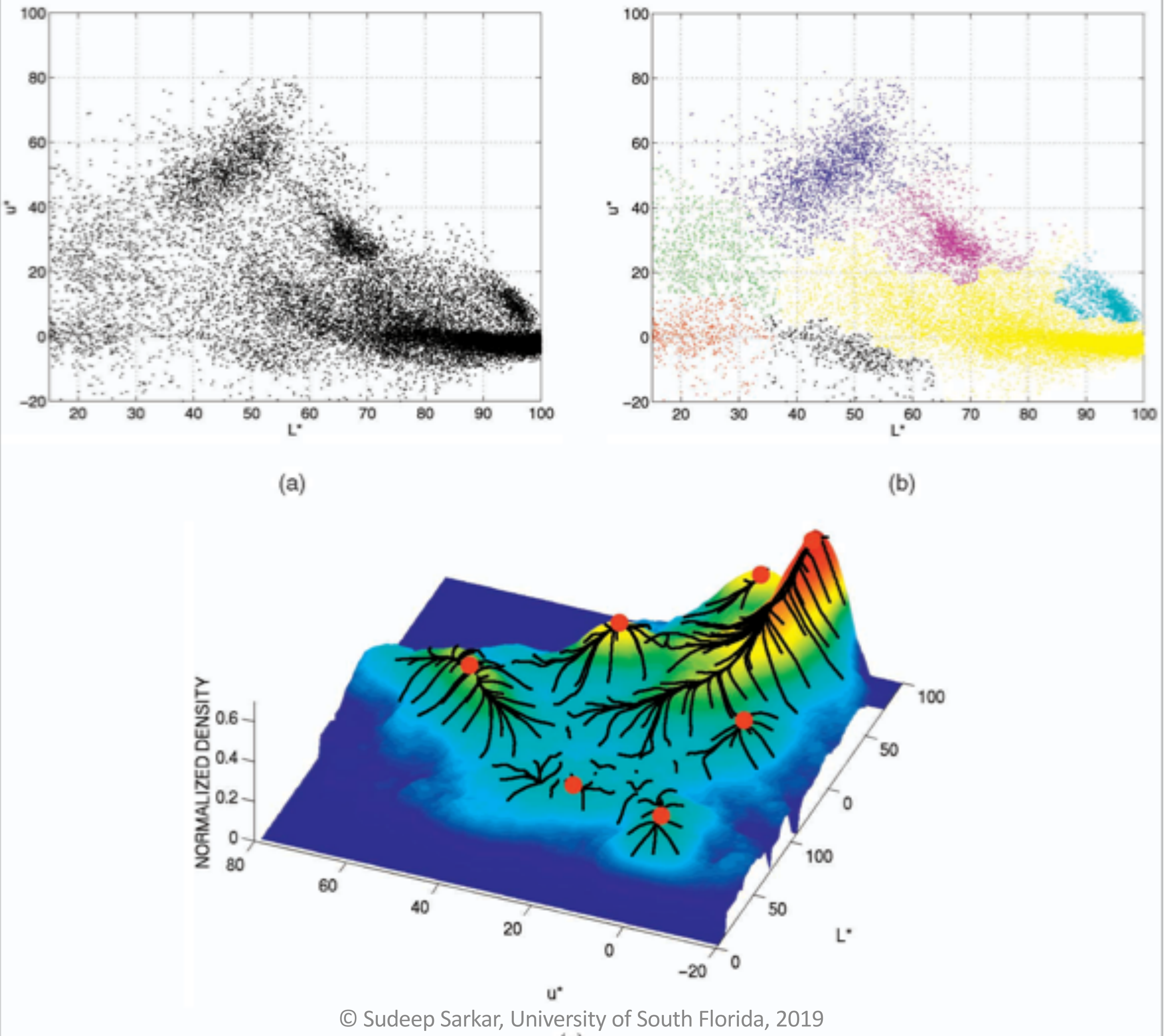
Segmentation by exploiting similarity of color



Distribution of pixel values



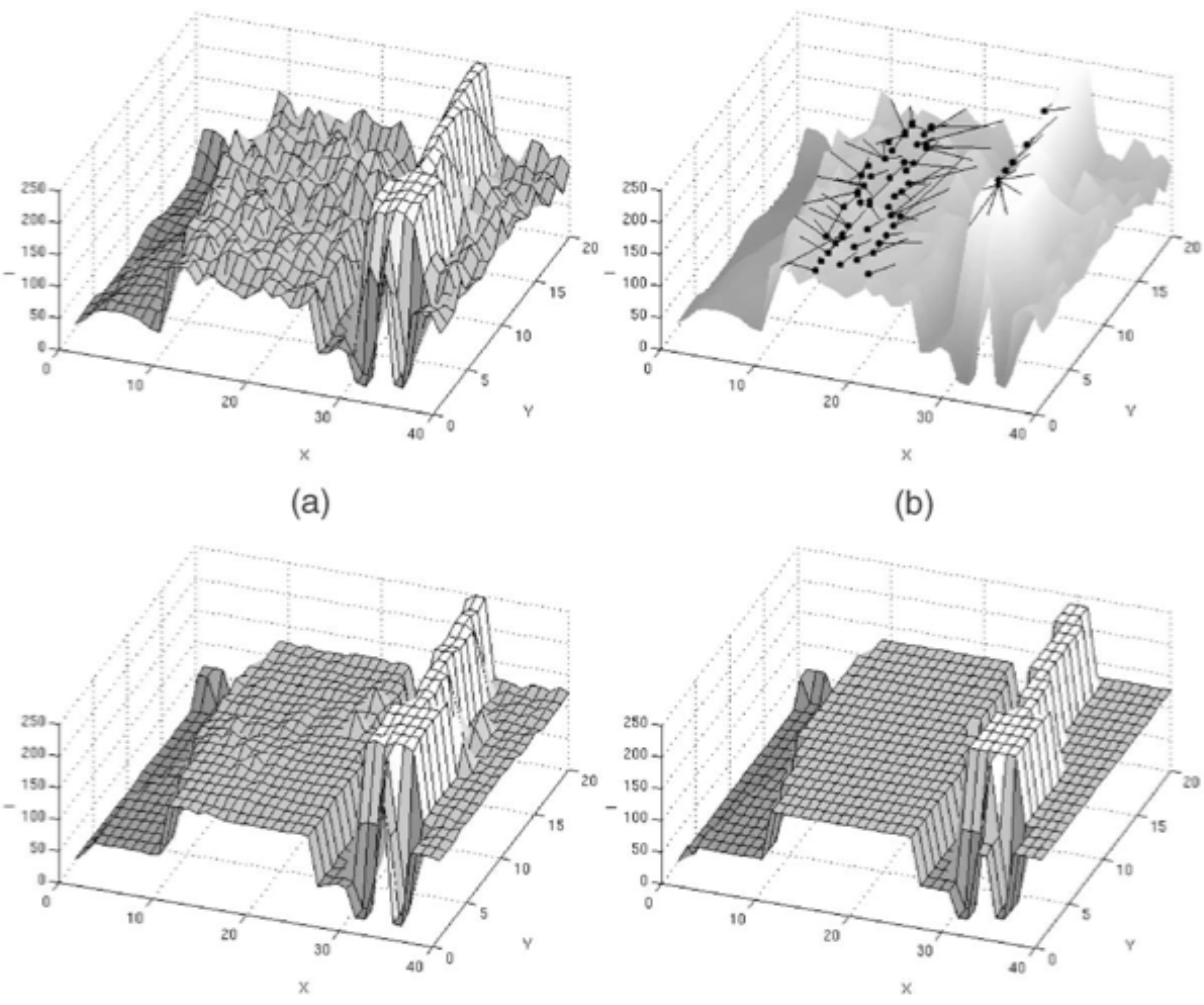
Comaniciu, D. and Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(5):603–619.



100
80
60
40
20
0
−20
u*
20 30 40 50 60 70 80 90 100
L*
(a)
(b)
NORMALIZED DENSITY
0.6
0.4
0.2
0
80
60
40
20
0
−20
u*
100
50
0
L*

(c)

(a) (b)



(c) (d)

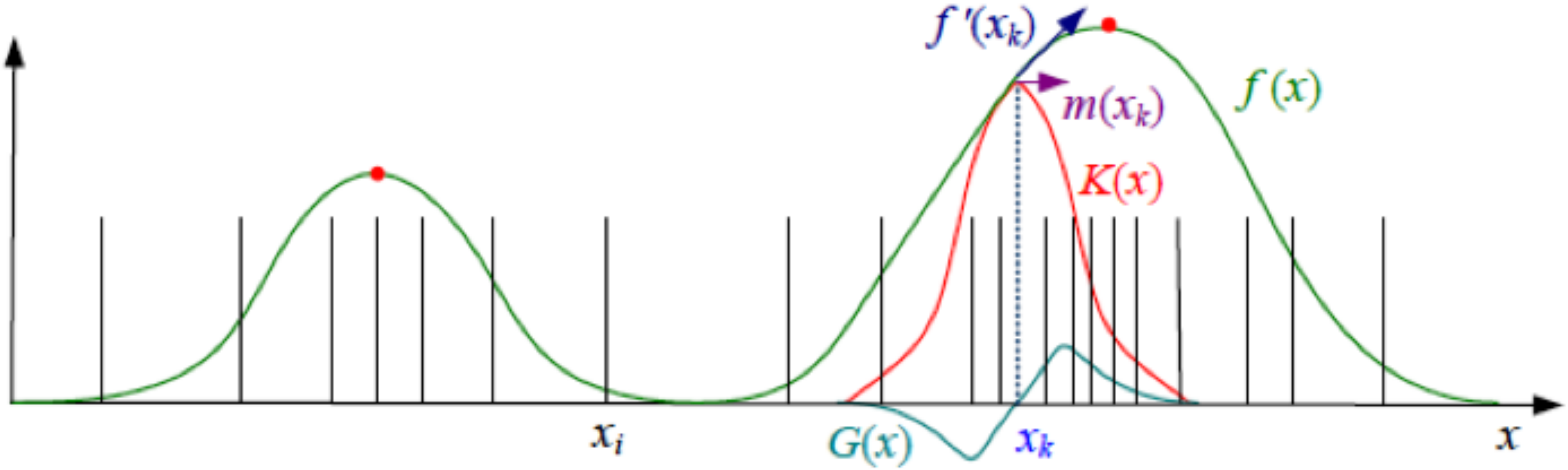
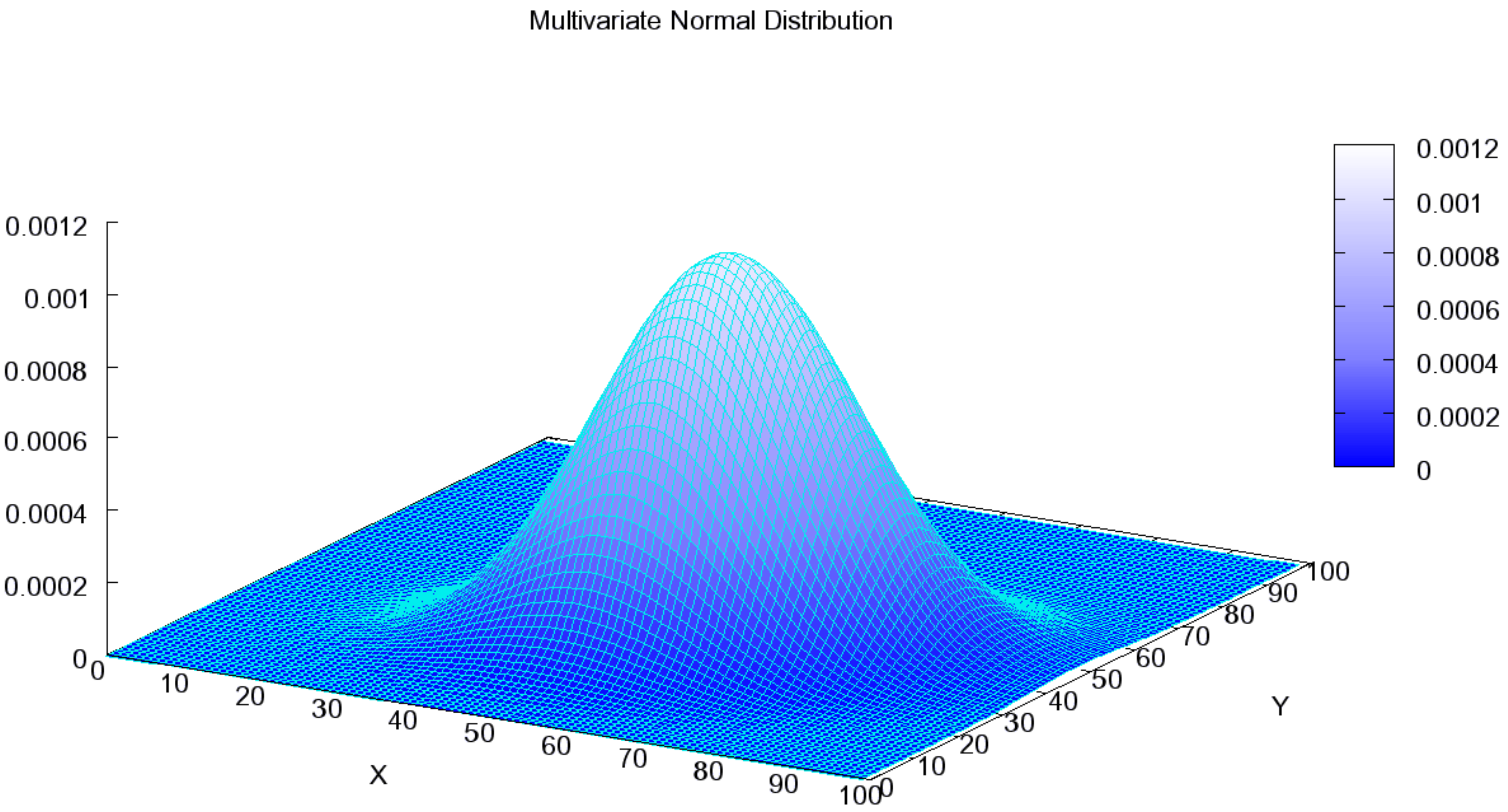
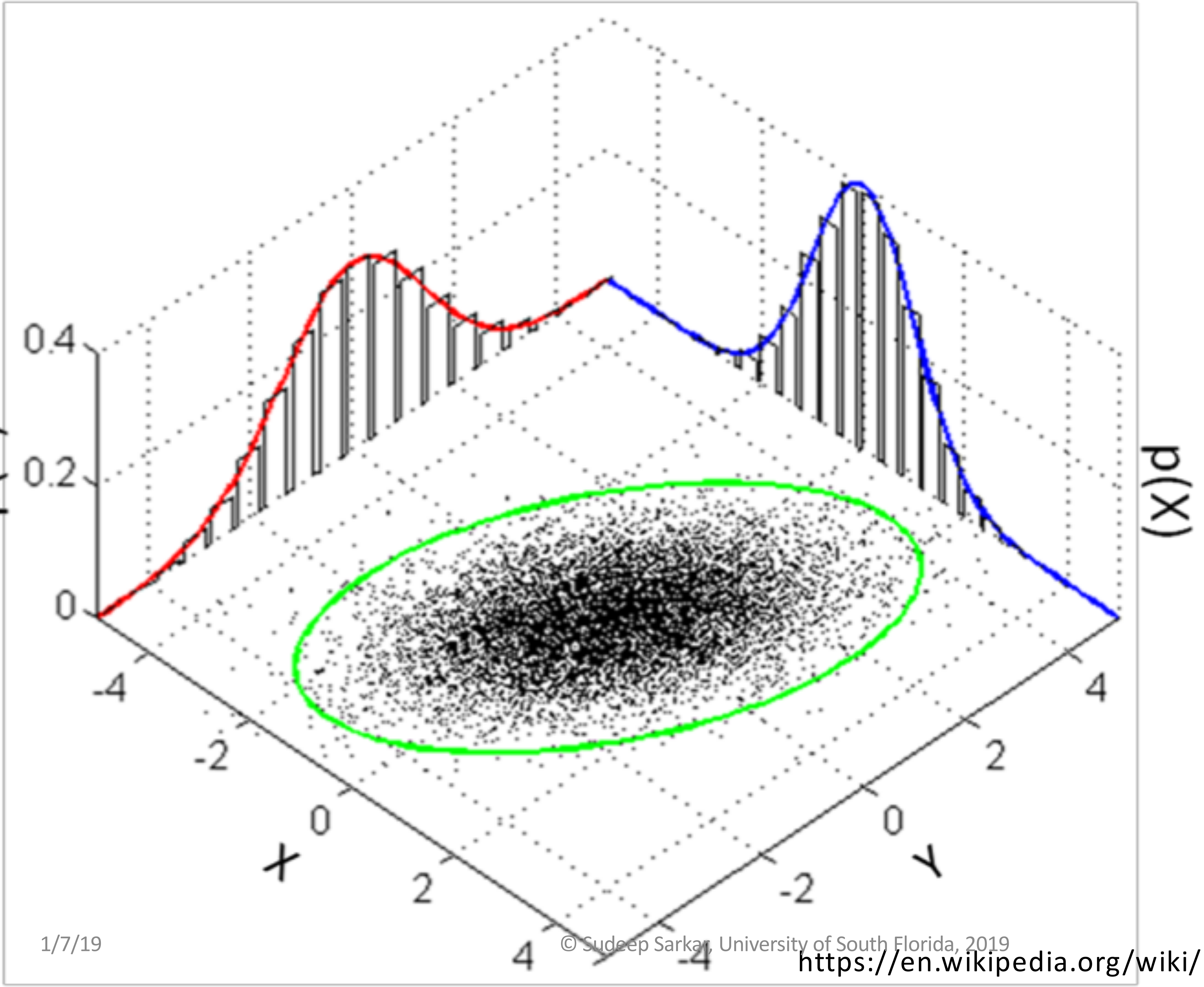


Figure 5.17 One-dimensional visualization of the kernel density estimate, its derivative, and a mean shift. The kernel density estimate $f(x)$ is obtained by convolving the sparse set of input samples x_i with the kernel function $K(x)$. The derivative of this function, $f'(x)$, can be obtained by convolving the inputs with the derivative kernel $G(x)$. Estimating the local displacement vectors around a current estimate x_k results in the mean-shift vector $m(x_k)$, which, in a multi-dimensional setting, point in the same direction as the function gradient $\nabla f(x_k)$. The red dots indicate local maxima in $f(x)$ to which the mean shifts converge.



https://en.wikipedia.org/wiki/Multivariate_normal_distribution



https://en.wikipedia.org/wiki/Multivariate_normal_distribution