CAP 6635 – Artificial Intelligence

Lecture 10: How do machines learn? (Part 2)

Oge Marques, PhD

Professor

College of Engineering and Computer Science

College of Business

FLORIDA ATLANTIC UNIVERSITY

@ProfessorOge

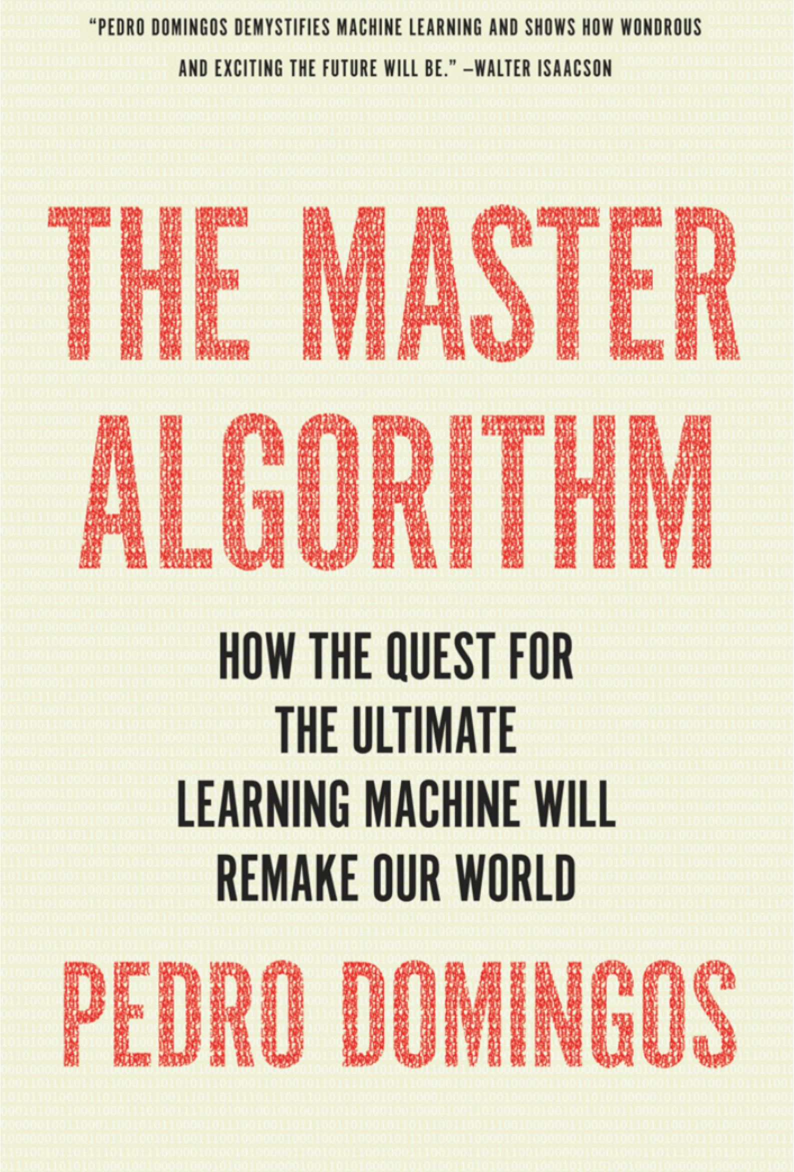
ProfessorOgeMarques

Previously on CAP 6635…

The Master Algorithm

(by Pedro Domingos)

"All knowledge—past, present, and future—can be derived from data by a single, universal learning algorithm."



The symbolists

No free lunch theorem - David Wolpert and William Macready (1997)

- "Any two optimization algorithms are equivalent when their performance is averaged across all possible problems."
- No single ML algorithm can be universally better than any other one on all domains.
- "The practical consequence of the "no free lunch" theorem is that there's no such thing as learning without knowledge. Data alone is not enough. Starting from scratch will only get you to scratch. Machine learning is a kind of knowledge pump: we can use it to extract a lot of knowledge from data, but first we have to prime the pump." -- Pedro Domingos

Overfitting

- “Overfitting is the central problem in machine learning.”
- “Learning algorithms are particularly prone to overfitting, though, because they have an almost unlimited capacity to find patterns in data.”
- “Overfitting happens when you have too many hypotheses and not enough data to tell them apart.”
- “Bottom line: learning is a race between the amount of data you have and the number of hypotheses you consider.”

New material starts here…

Test-set accuracy

- "You don't believe anything until you've verified it on *data that the learner didn't see*."
- Accuracy on previously unseen data is a pretty stringent test; so much so, in fact, that a lot of science fails it.
  - That does not make it useless, because <u>science is not just about prediction; it's also about explanation and understanding</u>.
  - But ultimately, if your models don't make accurate predictions on new data, you can't be sure you've truly understood or explained the underlying phenomena.
- And for "machine learning, testing on unseen data is indispensable because it's the only way to tell whether the learner has overfit or not.

The tank recognizer anecdote

- Even test-set accuracy is not foolproof.
- According to legend, in an early military application a simple learner detected tanks with 100% accuracy in both the training set and the test set, each consisting of one hundred images.
    - Amazing—or suspicious?
    - Turns out all the tank images were lighter than the non-tank ones, and that's all the learner was picking up."

Husky or wolf?

https://becominghuman.ai/its-magic-i-owe-you-no-explanation-explainableai-43e798273a08

Husky or wolf?

(a) Husky classified as wolf (b) Explanation

Figure 11: Raw data and explanation of a bad model's prediction in the "Husky vs Wolf" task.

"Why Should I Trust You?"
Explaining the Predictions of Any Classifier

Marco Tulio Ribeiro
University of Washington
Seattle, WA 98105, USA
marcotcr@cs.uw.edu

Sameer Singh
University of Washington
Seattle, WA 98105, USA
sameer@cs.uw.edu

Carlos Guestrin
University of Washington
Seattle, WA 98105, USA
guestrin@cs.uw.edu

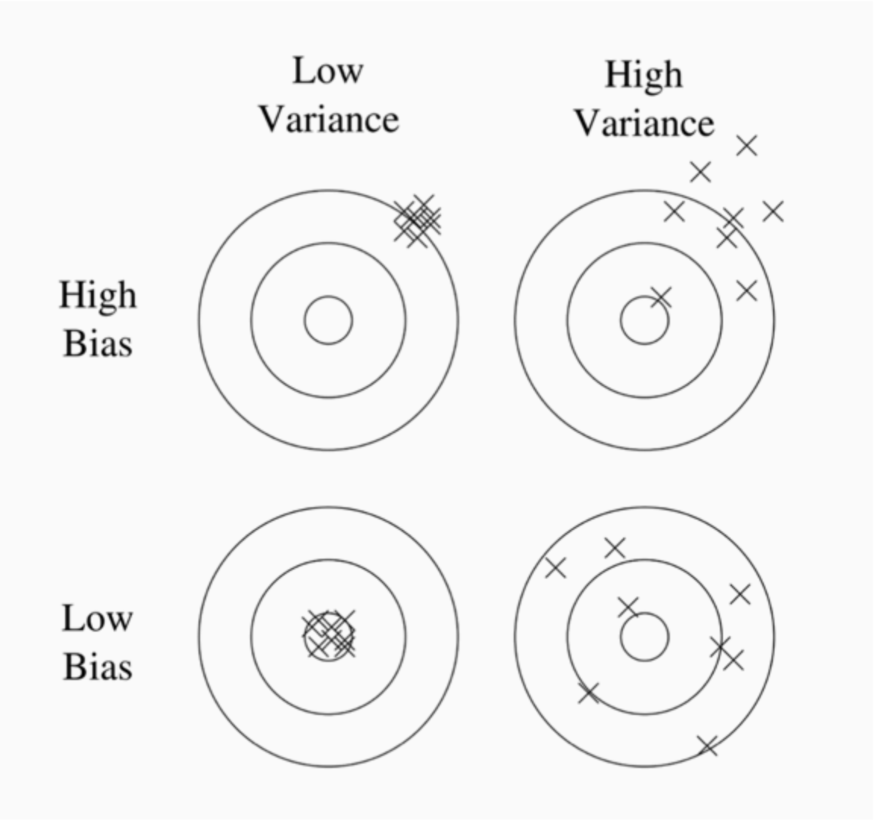
https://arxiv.org/pdf/1602.04938.pdf

Occam’s razor

Occam’s razor, also spelled **Ockham’s razor**, also called **law of economy** or **law of parsimony**, principle stated by the Scholastic philosopher William of Ockham (1285–1347/49) that *pluralitas non est ponenda sine necessitate*, “plurality should not be posited without necessity.” The principle gives precedence to simplicity: of two competing theories, the simpler explanation of an entity is to be preferred. The principle is also expressed as “Entities are not to be multiplied beyond necessity.”

Bias vs. Variance

- You can estimate the bias and variance of a learner by comparing its predictions after learning on random variations of the training set.
  - If it keeps making the same mistakes, the problem is bias, and you need a more flexible learner (or just a different one).
  - If there's no pattern to the mistakes, the problem is variance, and you want to either try a less flexible learner or get more data.
- Most learners have a knob you can turn to make them more or less flexible, such as the threshold for significance tests or the penalty on the size of the model.
  - Tweaking that knob is your first resort.



Induction vs. Deduction

- Read the book ☺

Decision trees

- Read the book ☺

The symbolists’ core belief

- “The symbolists’ core belief is that all intelligence can be reduced to manipulating symbols”
  -- Pedro Domingos. “The Master Algorithm.”

Symbolists: last words

- Symbolist machine learning is an offshoot of the knowledge engineering school of AI.
- Because of its origins and guiding principles, symbolist machine learning is still closer to the rest of AI than the other schools.
- If computer science were a continent, symbolist learning would share a long border with knowledge engineering.

Symbolists: last words

- Symbolism is the shortest path to the Master Algorithm.
  - It doesn’t require us to figure out how evolution or the brain works, and it avoids the mathematical complexities of Bayesianism.
  - Sets of rules and decision trees are easy to understand, so we know what the learner is up to.
  - This makes it easier to figure out what it’s doing right and wrong, fix the latter, and have confidence in the results.

Sidebar: David Marr (1945-1980)

- One of the most influential neuroscientists of vision.
- Thought of vision as an information-processing task.
- In his book *Vision* (1982), he distinguished three different levels of description involved in understanding complex information processing systems:
  - Computational level
  - Algorithmic level
  - Implementation level
    - An important point is that the levels can be considered independently.

Source: Palmer, S. *Vision Science*.

Marr's computational framework

Computational theory
What is the nature of the problem to solved, 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?

Representation and algorithm
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 from input to output?

Hardware implementation
How can the representation and algorithm be realised physically?

Source: Frisby and Stone. *Seeing, 2nd Edition*.

Information Processing Theory

- 3 levels of information processing
  - Computational level
    - The most abstract description level.
    - Sets out the goal of a process and an outline of how it can be achieved in principle.
    - This includes defining the input, the output and establishing the constraints that will be used in computing one from the other.
    - This level of theorizing specifies what computation needs to be performed and on what information it should be based, <u>without specifying how it is accomplished</u>.

Source: Palmer, S. *Vision Science*.

Information Processing Theory

- 3 levels of information processing
  - Algorithmic level
    - Algorithmic descriptions are more specific than computational descriptions in that they specify how a computation is executed in terms of information processing operations.
    - Gives details of how the input and output are represented and a set of rules for the transformation between the two.
    - Thus, the algorithmic level corresponds most closely to the concept of a program as it is understood in computer science.
    - Important point: more than one algorithm can satisfy a given computational description.

Source: Palmer, S. *Vision Science*.

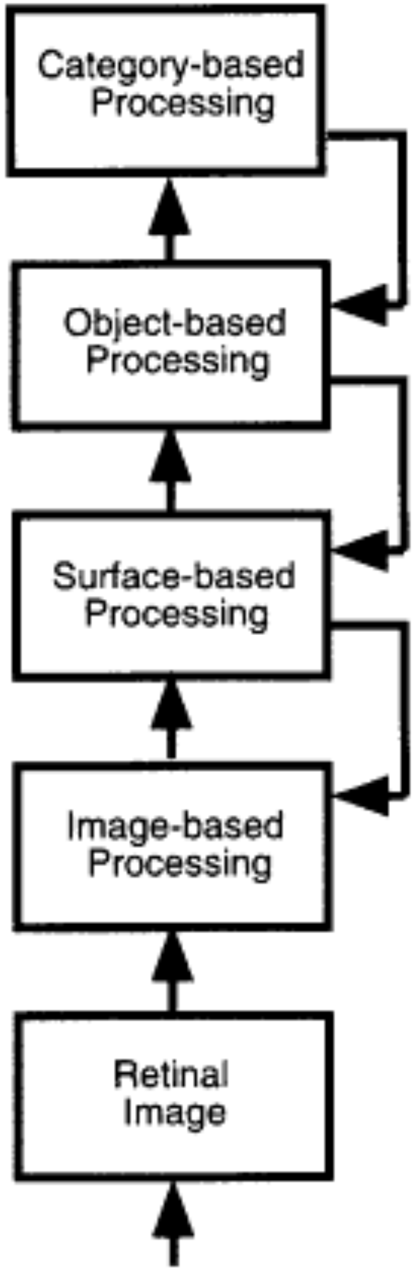
Information Processing Theory

- 3 levels of information processing
  - Implementation level
    - It specifies the physical method for carrying out an algorithm, for example in computer hardware or using neurons.
    - Just as the same program can be run on many computers that differ in their physical construction, so the same algorithm can be implemented using many physically different devices.

Source: Palmer, S. *Vision Science*.

Marr’s impact

- Palmer’s model
  - Decomposition of visual perception at the algorithmic level into <u>four major stages</u> beyond the retinal image itself.
  - Each stage is defined by a different kind of output representation and the processes that are required to compute it from the input representation.
  - These four stages provide a fairly general and robust framework for understanding vision as a computational process.



Source: Palmer, S. *Vision Science*.

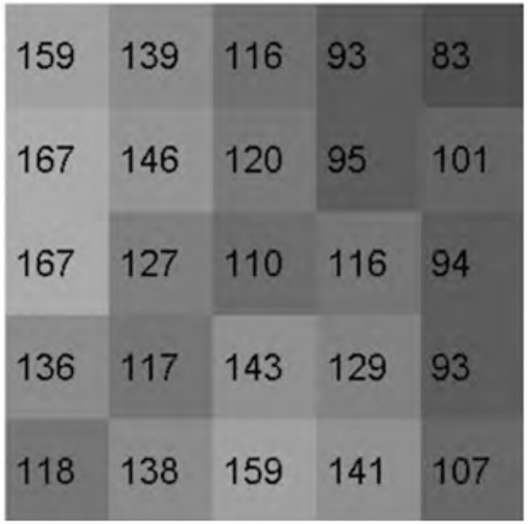
Four Stages of Visual Perception

- The retinal image
  - The proximal stimulus for vision is the pair of 2D images projected from the environment to the viewpoint of the observer's eyes.
  - The complete set of *firing rates* in all receptors of both eyes therefore constitutes the first representation of optical information within the visual system.
    - This retinal representation is complicated by the distribution of receptors and is almost always simplified and regularized by approximating it as a homogeneous, 2D array of receptors.

Source: Palmer, S. *Vision Science*.

Four Stages of Visual Perception

- The retinal image: an approximation
  - Think for a minute… how hard it might be to perceive 3D just on the basis of a 2D array of numbers….

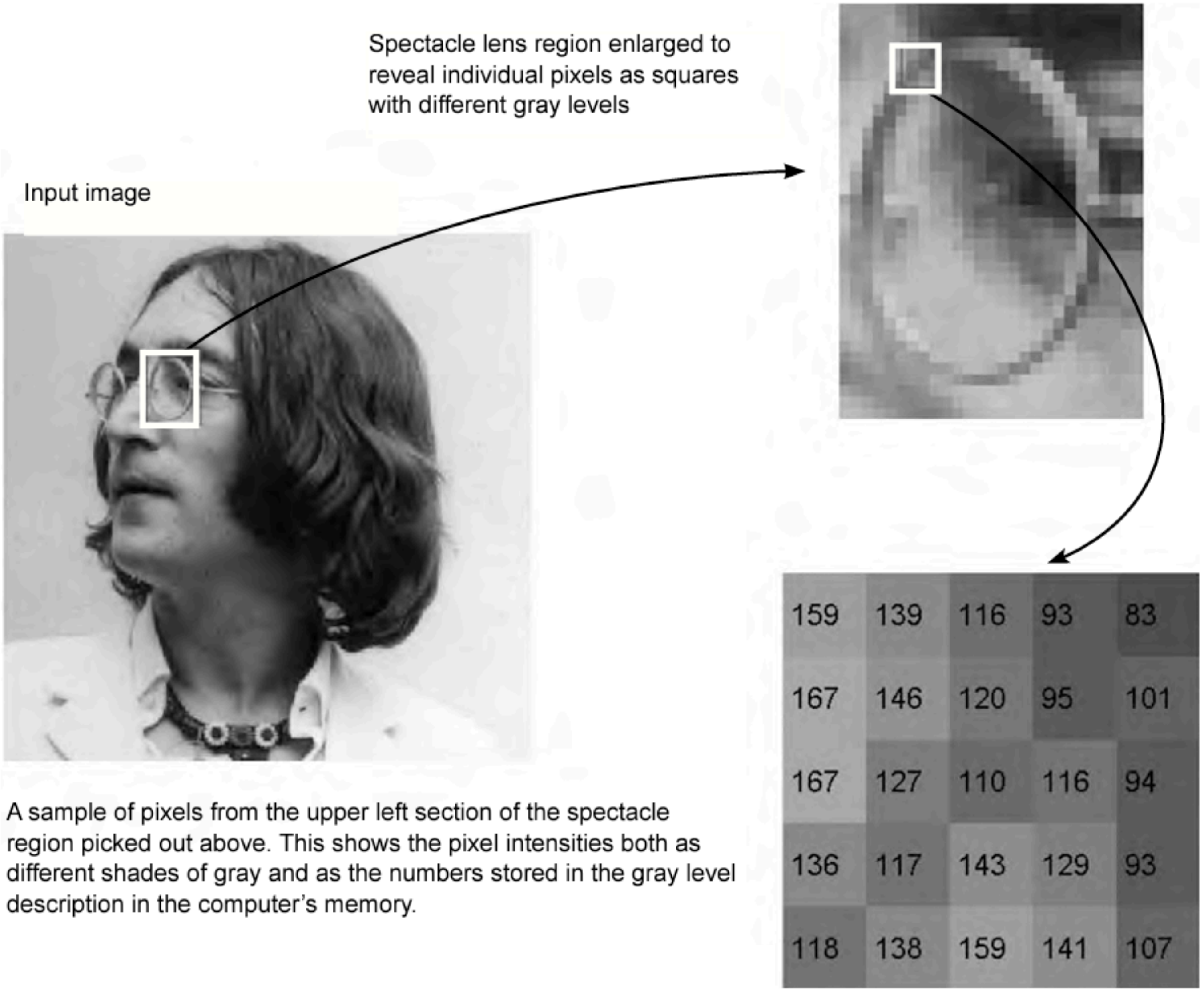


  - Q: What is this?
  - A: A portion of John Lennon's spectacles!

Source: Frisby and Stone. *Seeing, 2nd Edition*.

Four Stages of Visual Perception

- Don't believe me?



159	139	116	93	83
167	146	120	95	101
167	127	110	116	94
136	117	143	129	93
118	138	159	141	107

A sample of pixels from the upper left section of the spectacle region picked out above. This shows the pixel intensities both as different shades of gray and as the numbers stored in the gray level description in the computer's memory.

Source: Frisby and Stone. *Seeing, 2nd Edition*.

Four Stages of Visual Perception

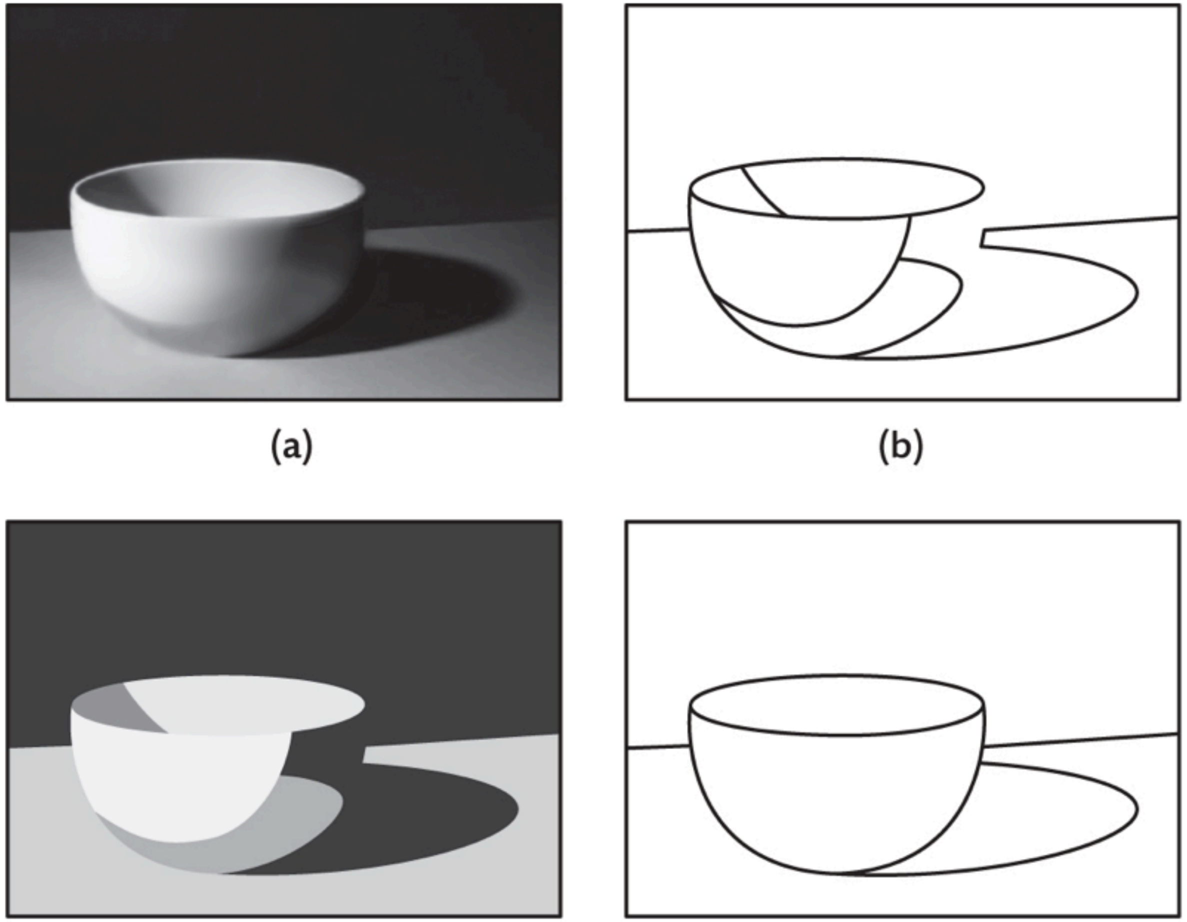
- The image-based stage
  - Includes image-processing operations such as detecting local edges and lines, linking local edges and lines together more globally, matching up corresponding images in the left and right eyes, defining two-dimensional regions in the image, and detecting other image-based features, such as line terminations and "blobs."
    - These 2D features of images characterize their structure and organization before being interpreted as properties of 3D scenes.

Source: Palmer, S. *Vision Science*.

Four Stages of Visual Perception

- The image-based stage (cont'd)
  - Luminance edges

- How does the visual system derive a representation like (d) from a retinal image like (a)?



(a) (b) (c) (d)

Source: Yantis, Steven. *Sensation and Perception, 2nd Edition*.

Four Stages of Visual Perception

- Primal sketches (Marr)
  - Raw: includes just the results of elementary detection processes that locate edges, bars, blobs, and line terminations.
  - Full: also includes global grouping and organization among the local image features present in the raw primal sketch.

Source: Palmer, S. *Vision Science*.

Four Stages of Visual Perception

- The surface-based stage
  - Concerned with recovering the intrinsic properties of visible surfaces in the external world that might have produced the features that were discovered in the image-based stage.
    - The fundamental difference is that the surface-based stage represents information about the external world in terms of the spatial layout of visible surfaces in three dimensions, whereas the image-based stage refers to image features in the 2D pattern of light falling on the retina.
  - Constructing a surface-based representation is the first step in recovering the third spatial dimension from 2D images.
    - It does not contain information about *all* the surfaces that are present in the environment, only about those that are visible from the current viewpoint.

Source: Palmer, S. *Vision Science*.

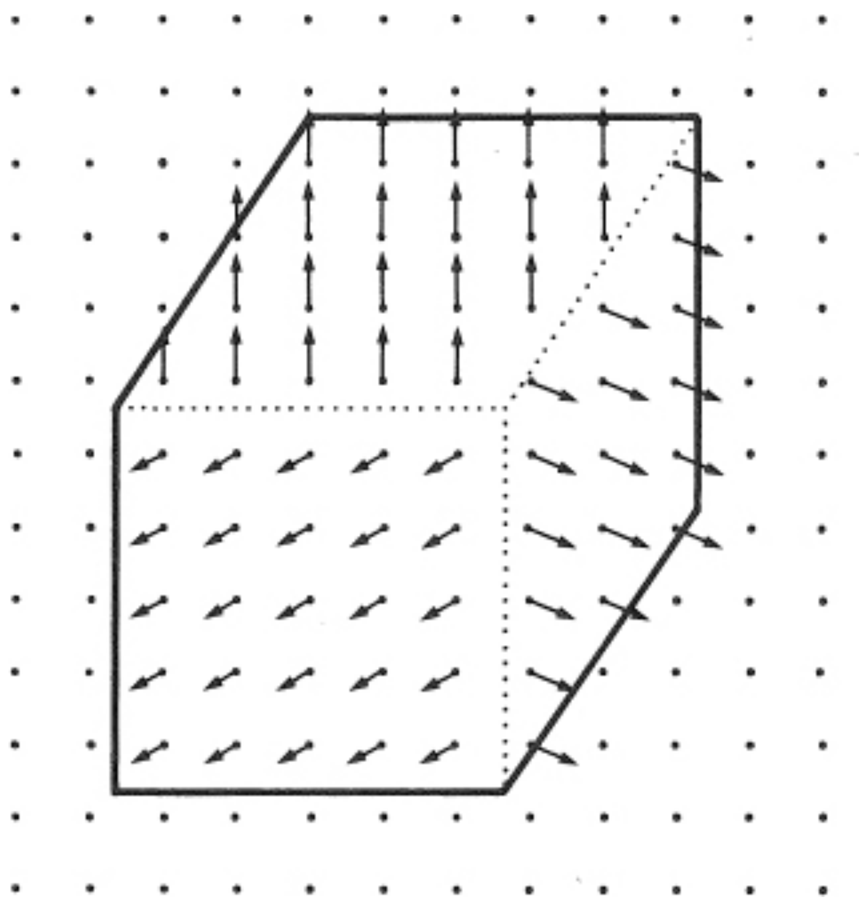
Four Stages of Visual Perception

- The surface-based stage (cont'd)
  - Helpful representation to remove ambiguities of edges
  - **Example**: a surface-based representation for a hat, composed of circles lying on the local surface patches and vectors sticking perpendicularly out of them at a sampling of locations, as though needles were sticking perpendicularly out of the small patches of surface.

Source: Frisby and Stone. *Seeing, 2nd Edition*.

Four Stages of Visual Perception

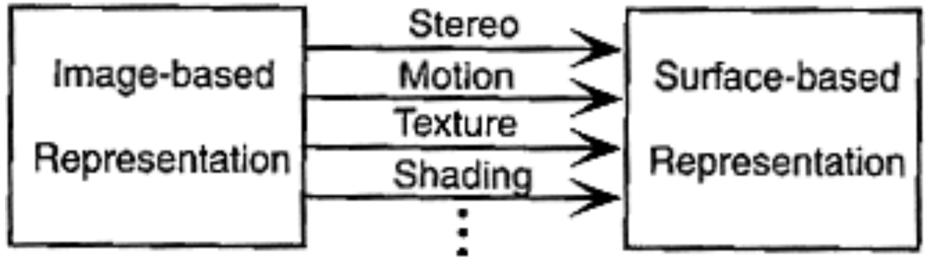
- The surface-based stage (cont'd)
  - Marr (1978): 21/2-D sketch
    - A clever name used to emphasize the fact that it lies somewhere between the true 2-D structure of image-based representations and the true 3-D structure of object-based representations.



Source: Palmer, S. *Vision Science*.

Four Stages of Visual Perception

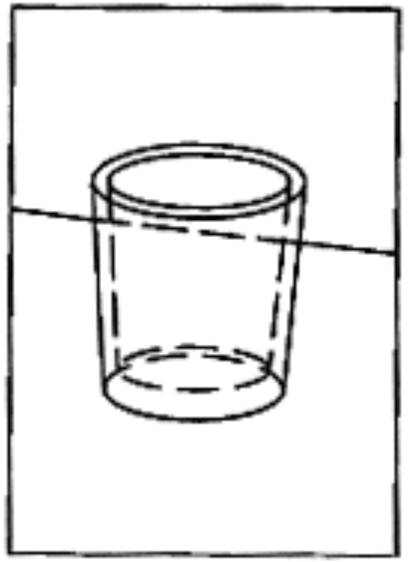
- The surface-based stage (cont'd)
  - Deriving a surface-based representation from an image-based representation:
    - stereopsis (the small difference between the lateral position of objects in the images of the left and right eyes)
    - motion parallax (differences in velocity of points at various distances due to motion of the observer or object)
    - shading and shadows
    - other pictorial properties (e.g., texture, size, shape, and occlusion).



Source: Palmer, S. *Vision Science*.

Four Stages of Visual Perception

- The object-based stage
  - The stage in which visual representation includes truly 3-D information.
    - Further hidden assumptions about the nature of the visual world are required, because now the inferences include information about unseen surfaces or parts of surfaces.
    - The inclusion of these unseen surfaces implies that they involve explicit representations of whole objects in the environment.
    - Recovering the 3-D structure of these environmental objects is the goal of object-based processing.



Source: Palmer, S. *Vision Science*.

Four Stages of Visual Perception

- The object-based stage (cont'd)
  - How do we recognize objects from different viewpoints?

Structural-Description Models

Propose that our ability to recognize 3D objects is based on 3D volumes (called *volumetric features*) that can be combined to create the overall shape of an object.

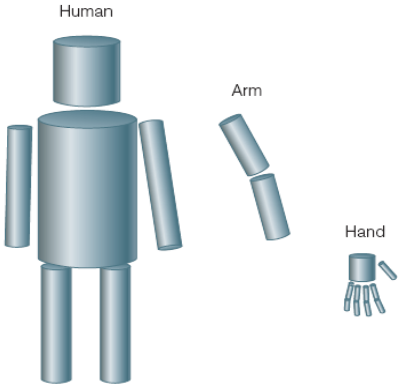


Image-Description Models

Propose that our ability to recognize objects from different viewpoints is based on stored 2D views of the object as it would appear from different viewpoints.

Which Model Is Correct? The actual mechanism for object recognition probably involves elements of both the structural-description and image-description models (Palmeri & Gauthier, 2004)

Source: Torralba and others.

Four Stages of Visual Perception

- The category-based stage
  - The final stage of perception, concerned with recovering the functional properties of objects: what they afford the organism, given its current beliefs, desires, goals, and motives.
  - It is widely believed that functional properties are accessed through a process of categorization, but the details aren't exactly clear.

Source: Palmer, S. *Vision Science*.

Four Stages of Visual Perception

- Example: chairs
  - The computer vision 3D reconstruction perspective
    - "Give me enough 2D images of a chair and I'll build a 3D model of it."

Source: Torralba and others.

Four Stages of Visual Perception

- Example: chairs
  - Affordances vs. mediated perception

The perception of function

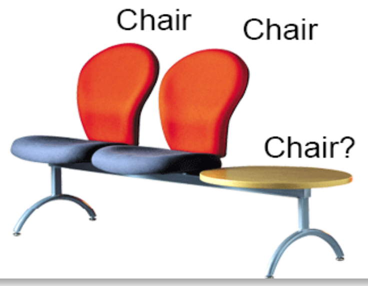
- Direct perception (affordances): Gibson

Flat surface
Horizontal
Knee-high
…
→ Sittable upon

- Mediated perception (Categorization)

Flat surface
Horizontal
Knee-high
…
→ Chair → Sittable upon

One caveat of this comparison: deciding that something is a chair might require access to more features than the ones needed to decide that we can sit on something… (it is a different level of categorization)



Source: Torralba and others.

Four Stages of Visual Perception

- Example: chairs
  - How to bundle all of these under a category?

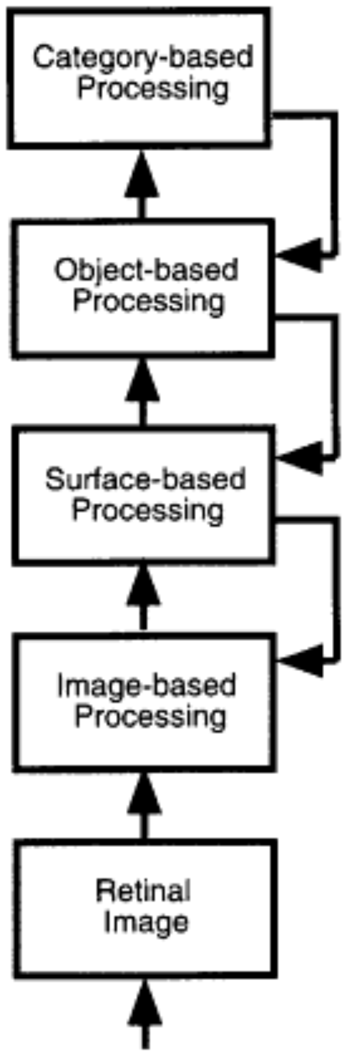
Four Stages of Visual Perception

- The category-based stage
  - It is possible — indeed, even likely — that people employ both types of processes (direct and indirect) in perceiving function.
    - Some objects such as chairs and cups have functional properties that are so intimately tied to their visible structure that one might not need to categorize them to know what they can be used for.
    - Other objects, such as computers and telephones, have functions that are so removed from their obvious visual characteristics that they almost certainly need to be categorized first.
  - The extent to which people use each of these strategies to perceive functionally relevant information about objects is currently unknown.

Source: Palmer, S. *Vision Science*.

Four Stages of Visual Perception

- Final remarks
  - The four proposed stages of visual processing—image-based, surface-based, object-based, and category-based—represent the current best guess about the overall structure of visual perception.
  - They have been listed in the particular order in which they must logically be initiated, but that does not necessarily mean that each is completed before the next begins.
    - The arrows going backward in this figure indicate that later processes may feed back to influence earlier ones.

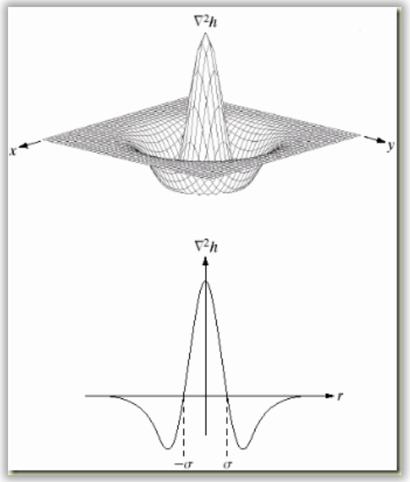


Source: Palmer, S. *Vision Science*.

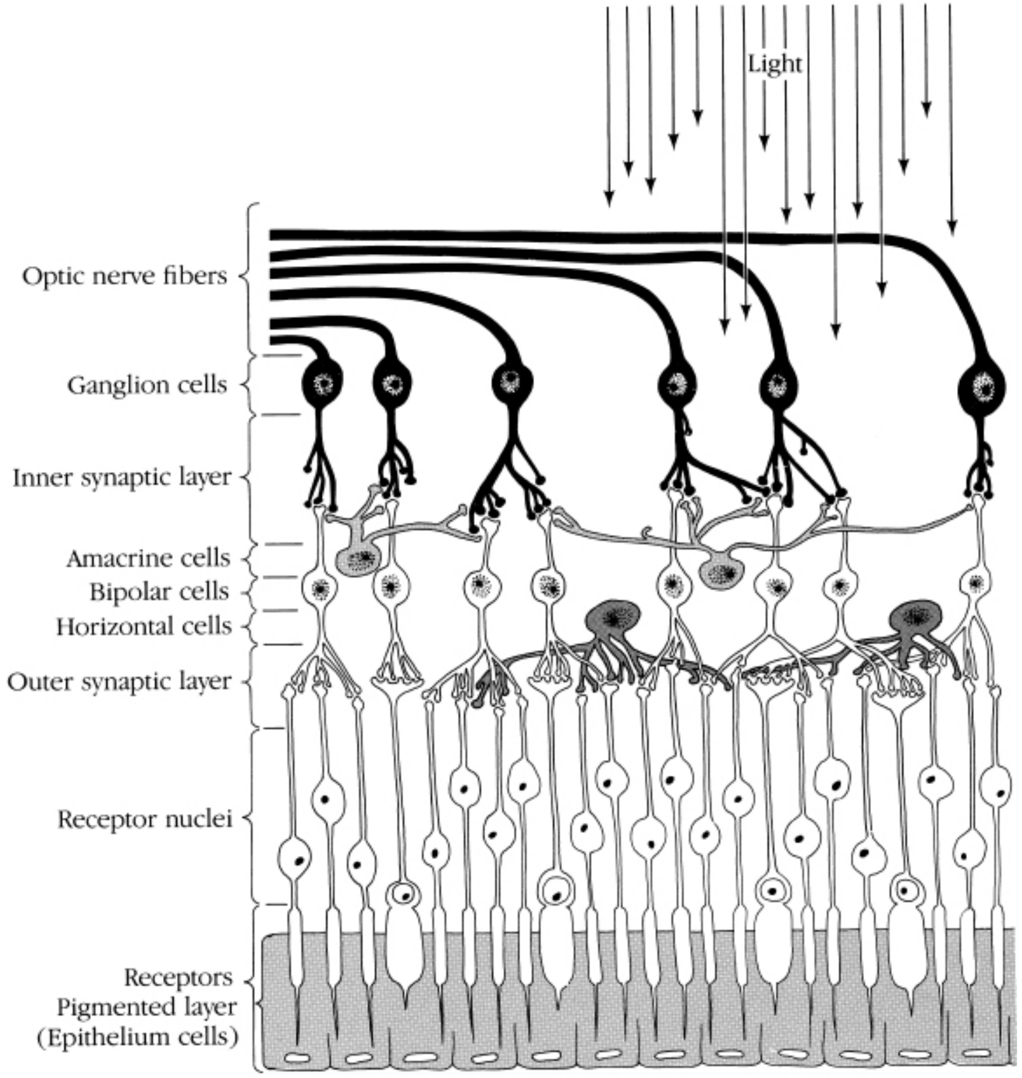
Marr's impact

- Edge detection

$$\nabla^2 G * I(x,y),$$
$$\text{where } \nabla^2 G(r) = -\frac{1}{\pi\sigma^4}\left(1-\frac{r^2}{2\sigma^2}\right)\exp\left(\frac{-r^2}{2\sigma^2}\right)$$



Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B: Biological Sciences*, *207*(1167), 187-217.



Source: Marr, D. *Vision*.