Visione negli AIBO

Slide tratte in parte dal materiale in rete di Manuela Veloso & Paul E. Rybski

http://www.andrew.cmu.edu/course/15-491

Computer Science Department

Carnegie Mellon

e di:

Rob Salkin &

Russell Goldstein

What is Vision?

- Process of extracting information from an image
- Identifying objects contained in the image
- Throwing away useless information

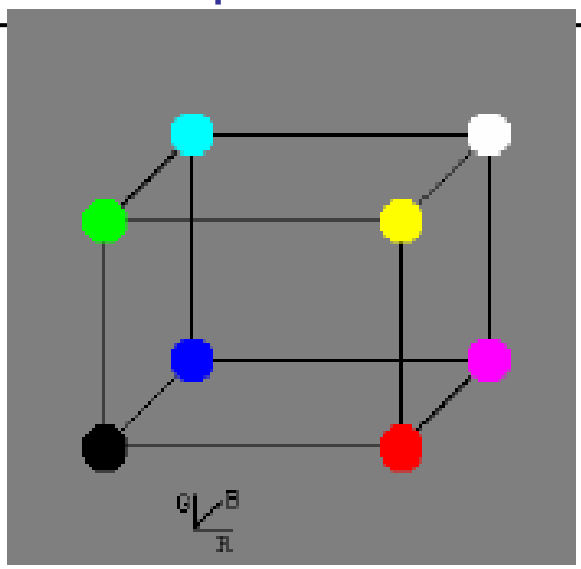
Color Spaces

- AIBO camera provides images at different resolutions: 208x160 pixels is the highest.
- Image are in YUV color space
- Each pixel has a 3-dimensional value
- Dimensions are called channels

Color Space Representation

- RGB: R = Red, G = Green, B = Blue
- YUV: Y = Brightness, UV = Color
- HSV: H Hue, S = Saturation,
  V = Brightness

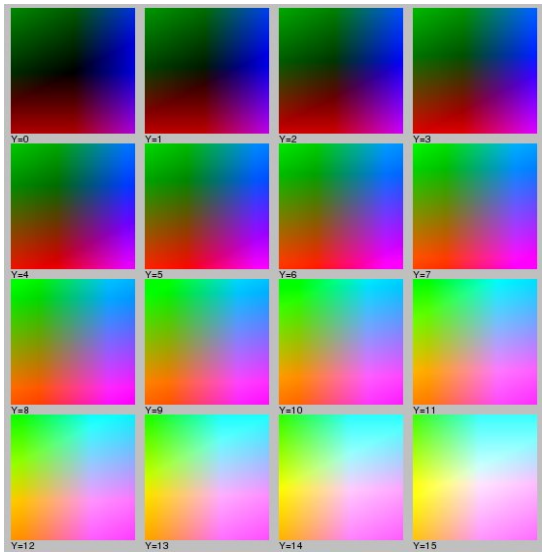
Color Spaces - RGB



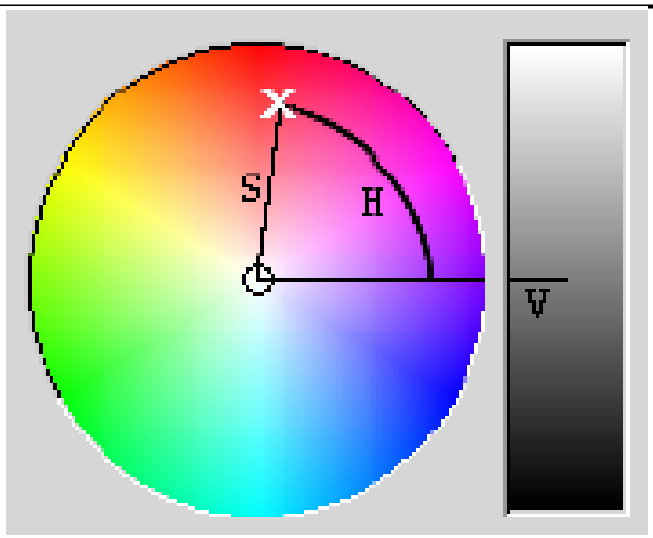
Color Spaces -YUV

- AIBO provides images in this environment
- Y – Luminance (brightness)
- U/Cb – Blueness (Blue vs. Green)
- V/Cr – Redness (Red vs. Green)

Color Spaces – YUV



Color Spaces - HSV

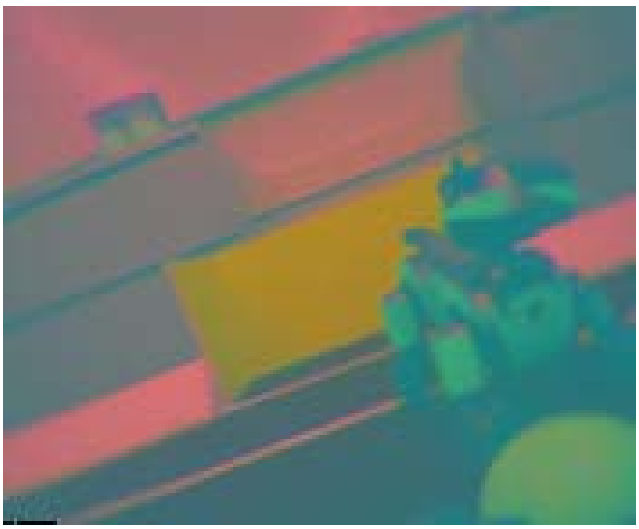


Color Spaces - Discussion

- RGB
  - Handled by most capture cards
  - Used by computer monitors
  - Not easily separable channels
- YUV
  - Handled by most capture cards
  - Used by TVs and JPEG images
  - Easily workable color space
- HSV
  - Rarely used in capture cards
  - Numerically unstable for grayscale pixels
  - Computationally expensive to calculate

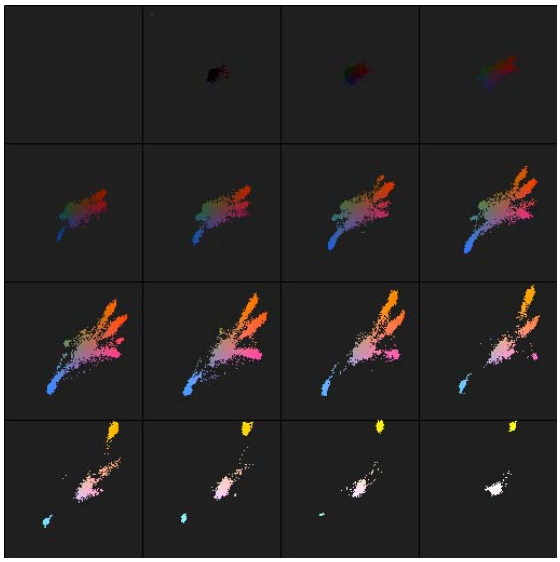
Image RGB

Image Raw



R=Y
G=U
B=V

YUV Histogram



Note: the U and V axes are swapped from the histogram in the previous slides (blue is in lower left corner in this slide but blue is in upper right corner in previous slide)

Vision Overview

- CMRoboBits vision is divided into two parts
- Low level
  - Handles bottom-up processing of image
  - Provides *summaries* of image features
- High level
  - Performs top-down processing of image
  - Uses *object models* to filter low-level vision data
    - Identifies objects
    - Returns properties for those objects

Low-Level Vision Overview

- Low level vision is responsible for summarizing *relevant-to-task* image features
  - Color is the main feature that is relevant to identifying the objects needed for the task
  - Important to reduce the total image information
- Color segmentation algorithm
  - Segment image into *symbolic colors*
  - Run *length encode* image
  - Find *connected components*
  - Join nearby components into *regions*

Color Segmentation

- Goal: semantically label each pixel as belonging to a particular type of object
- Map the domain of raw camera pixels into the range of symbolic colors *C*

$$F : y,u,v \rightarrow c \in C$$

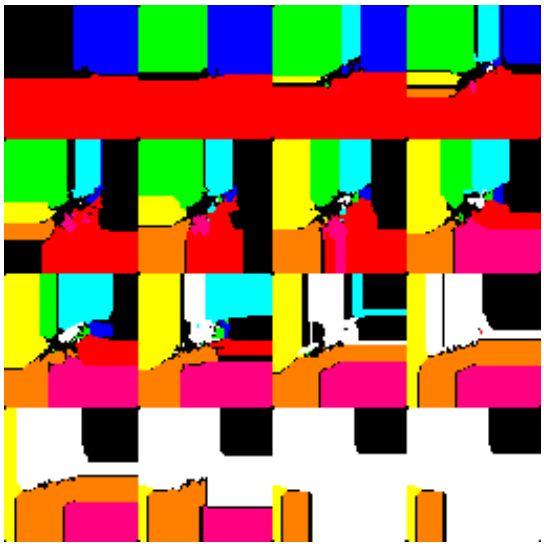
  - C includes ball, carpet, 2 goal colors, 1 additional marker color, 2 robot colors, walls/lines and unknown
- Reduces the amount of information per pixel roughly by 1.8M
  - Instead of a space of 256^3 values, we only have 9 values!

Before Segmentation

Ideal Segmentation

Result of Segmentation

Color Class Thresholds

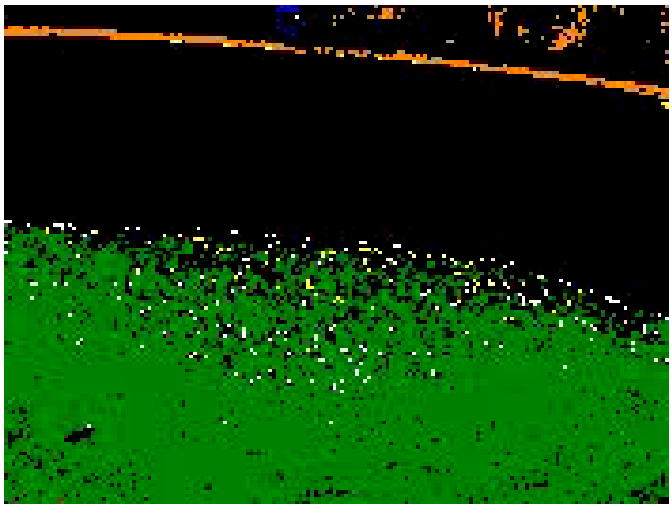


How to set a cluster

- The color volume can be subdivided ad hoc into 8 different clusters.
- The luminance channel (Y channel is subdivided into 32 bands or segments).
- For each segment the color boundaries for the cluster are identified by the following function:

  set(Y_segment , Cr_max, Cr_min, Cb_max, Cb_min)

Potential Problems with Color Segmentation



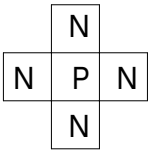
Color Segmentation Analysis

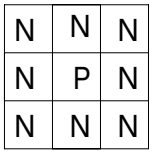
- Advantages
  - Quickly extract relevant information
  - Provide useful representation for higher-level processing
  - Differentiate between YUV pixels that have *similar* values
- Disadvantages
  - Cannot segment YUV pixels that have *identical* values into different classes
  - Generate smoothly contoured regions from noisy images

Turning Pixels into Regions

- A disjoint set of labeled pixels is still not enough to properly identify objects
- Pixels must be grouped into spatially-adjacent regions
  - Regions are grown by considering local neighborhoods around pixels

4-connected neighborhood





8-connected neighborhood

Summary

- Computer vision
- Color spaces
- Low-level vision
  - Color segmentation
  - Colored region extraction
- High-level vision
  - Object filters
  - Example: tracking the ball