venturi

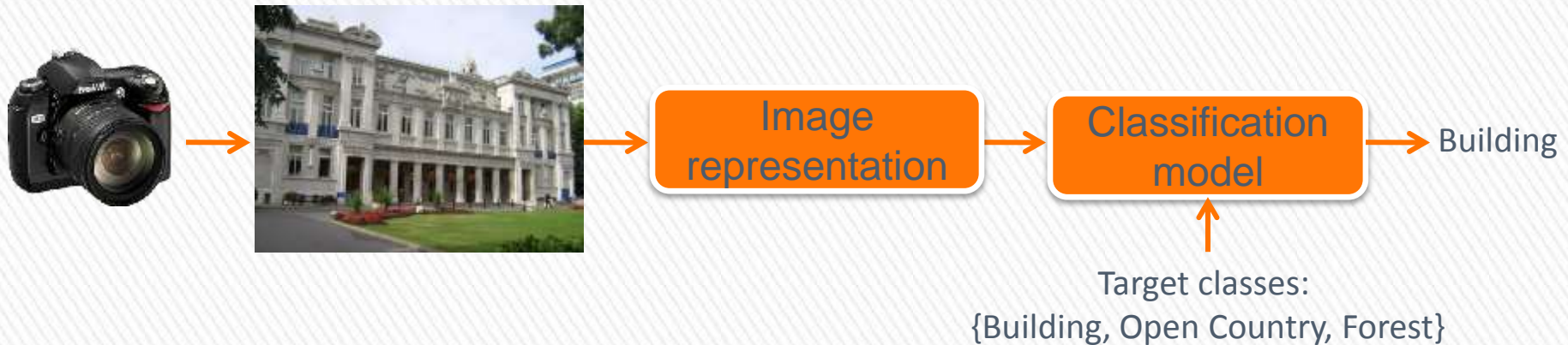# Video Scene Classification WP4

venturi

» Scene classification definition
» State of the art:
  > Representation:
    + Global centered;
    + Local centered;

  > Classification:
    + Discriminative;
    + Generative
» Functional pipeline
» Scene classes definition
» Database population
» Library experience/survey:
 > Pc environment
  > Library/SDK available

Agenda

» Scene classification (also known as *Visual categorization*) aims to automatically classify images into a set of different semantic classes, based on information extracted from the images.

Image representation → Classification model → Building

Target classes:
{Building, Open Country, Forest}

Definition

venturi

» Global Centered:

> Scene as a single entity. The identification of the constituent objects/concepts is bypassed.

+ Spatial Domain: *BoW (Bag of visual Words)*
+ Frequency Domain: *FFT Domain, Textons, DCT Domain*
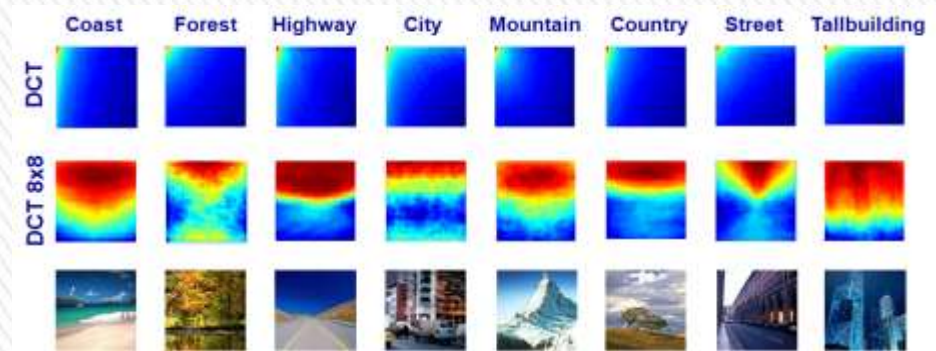
» Local centered:

> Scene defined as a collection of previously recognized objects/concepts within the scene;
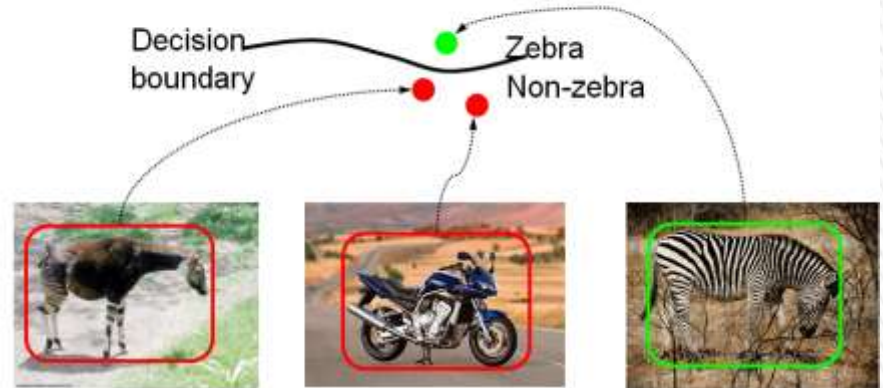> Require object recognition as first step *(Semantic Texton Forests)*

State of the art: Representation ❯ 4

venturi
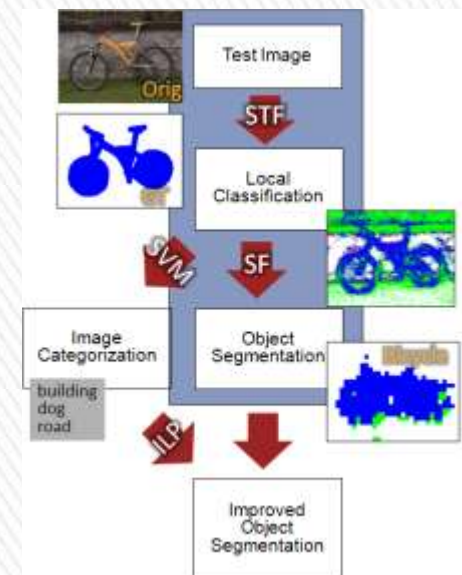
» Discriminative (e.g. SVM, Random Decision forest)

» Generative (e.g. LDA, HLDA, PLSA, Bayesian)

- Direct modeling of $\dfrac{p(zebra \mid image)}{p(no\ zebra \mid image)}$

Decision boundary
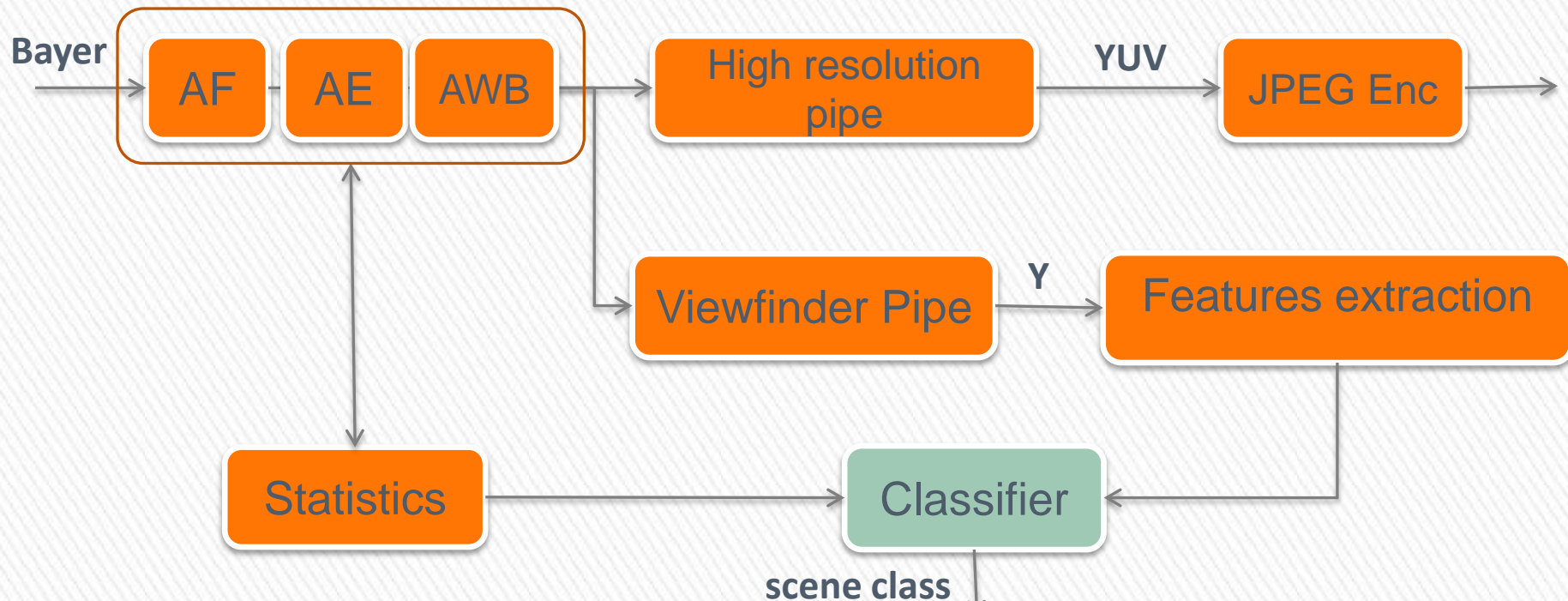
Zebra
Non-zebra

- Model $p(image \mid zebra)$ and $p(image \mid no\ zebra)$

	$p(image \mid zebra)$	$p(image \mid no\ zebra)$
	Low	Middle
	High	Middle→Low

State of the art: Classification

venturi

» The idea is to use two different types of features: camera statistics (focus, gain, white balance, exposure, histogram, etc.) and features extracted from the luminance plane (at least)

Bayer → [**AF** - **AE** - **AWB**] → High resolution pipe → **YUV** → JPEG Enc →

Viewfinder Pipe → **Y** → Features extraction

Statistics → Classifier ← Features extraction

scene class

Functional Pipeline ❯

» A list of scene classes has been defined according to the Use Cases (phone call March 28th 2012):

> Day/night: OK with camera parameters (when they will be available);
> Indoor/outdoor;
> Man-made/natural;
> Street/street intersection;
> Home rooms (bedroom, kitchen, living room, bathroom);
> Corridor/shelves (in a supermarket);
> Supermarket vs. (a precise list of classes has to be defined)

» Other classes:

> Crowded/non-crowded: a completely different approach is needed for this task;
> Text class could be also useful (some tests with our approach give very good results)

Scene classes definition

venturi

- » Download of publicly available datasets;
- » Automatic download from the Internet (e.g. flickr) according to specific tags;
- » Acquisition of images and statistics (when they will be available) through the STE U9500 platform;
- » Who, when, how will populate the Database?
- » Will camera and ISP statistics be available?

Database population

venturi

» Operating system:
  > Windows (No dependencies from specific API);

» Programming language:
  > C/C++;

» Library dependencies:
  > OpenCV (Camera, image handling);
  > LibSVM (source code available);

» Build system:
  > Qt Creator, qmake for Nokia N900
  > ad-hoc make file compatible with eclipse IDE should be possible

Library experience › 9