Automated Classification of Bitmap Images Using Decision Trees

Pavel Surynek

Faculty of Mathematics and Physics
Charles University in Prague
Czech Republic

Graduate School of Maritime Sciences
Kobe University
Japan

Ivana Lukšová

Faculty of Mathematics and Physics
Charles University in Prague
Czech Republic

Bitmap Classification

- the task is to **automatically classify** bitmap images into predefined classes
  - finite set of bitmap images $\mathcal{J}$
  - finite set of classification classes $\mathcal{K}$
- for each $t \in \mathcal{K}$ a **characterization** d(t) of the class t in the natural language is given (*example:* „image depicting landscape")
- the correct classification of the set of images $\mathcal{J}$ is defined with respect to a **fixed user** using a function c:
  - $c: \mathcal{J} \rightarrow 2^{\mathcal{K}}$ such that $\forall I \in \mathcal{J} \;\; \forall t \in c(I)$ **d(t) characterizes *I well***
- we need to learn $c': \mathcal{J} \rightarrow 2^{\mathcal{K}}$ such that it gives the same answer as **c** on as many as possible images
  - c is not known **explicitly**
  - the condition cannot be checked for **all** the images
  - **training/testing** sets are used

New Method

- the concept of **decision tree** is used as underlying technology
- it is crucial to propose a set of good **characterizing attributes** and attribute **extraction techniques**
  - different classification classes have different important characteristics
  example: straight lines are characteristic for images of buildings
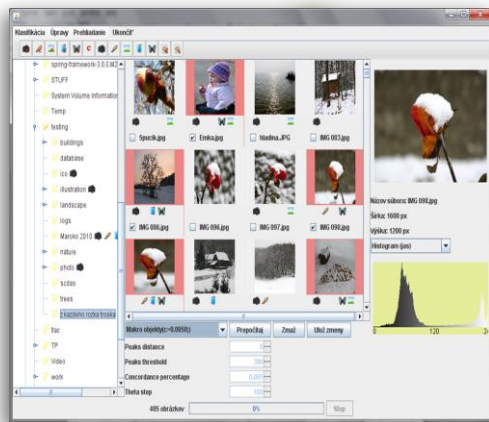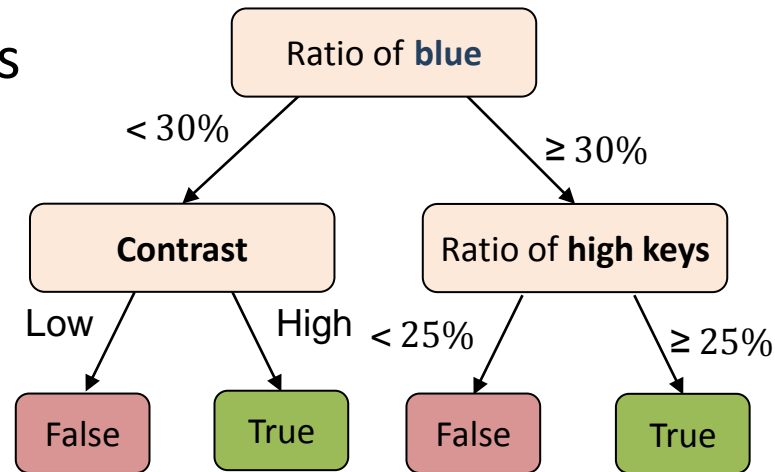
- successful classification of several classes of images
  - photography
  - drawings
  - macro objects
  - buildings

building

drawing

macro

Pavel Surynek and **Ivana Lukšová**