# Learning Strategies to Select Point Cloud Descriptors for 3D Object Classification: A Proposal


FernUniversität in Hagen

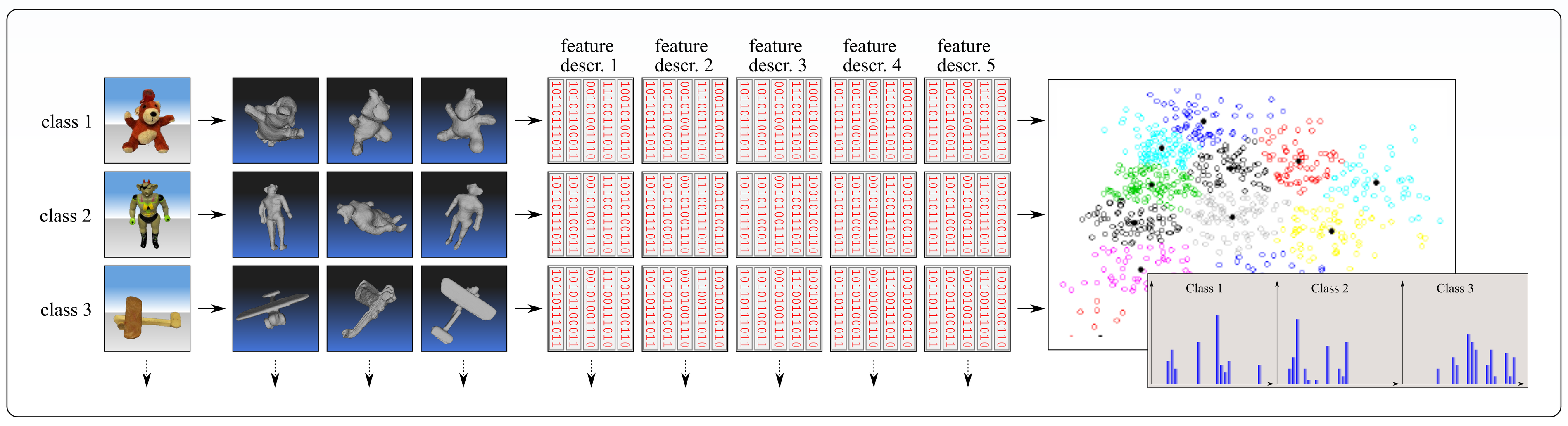
**Jens Garstka and Gabriele Peters**

Chair of Human-Computer-Interaction
University of Hagen - Germany


## 1 Initialization

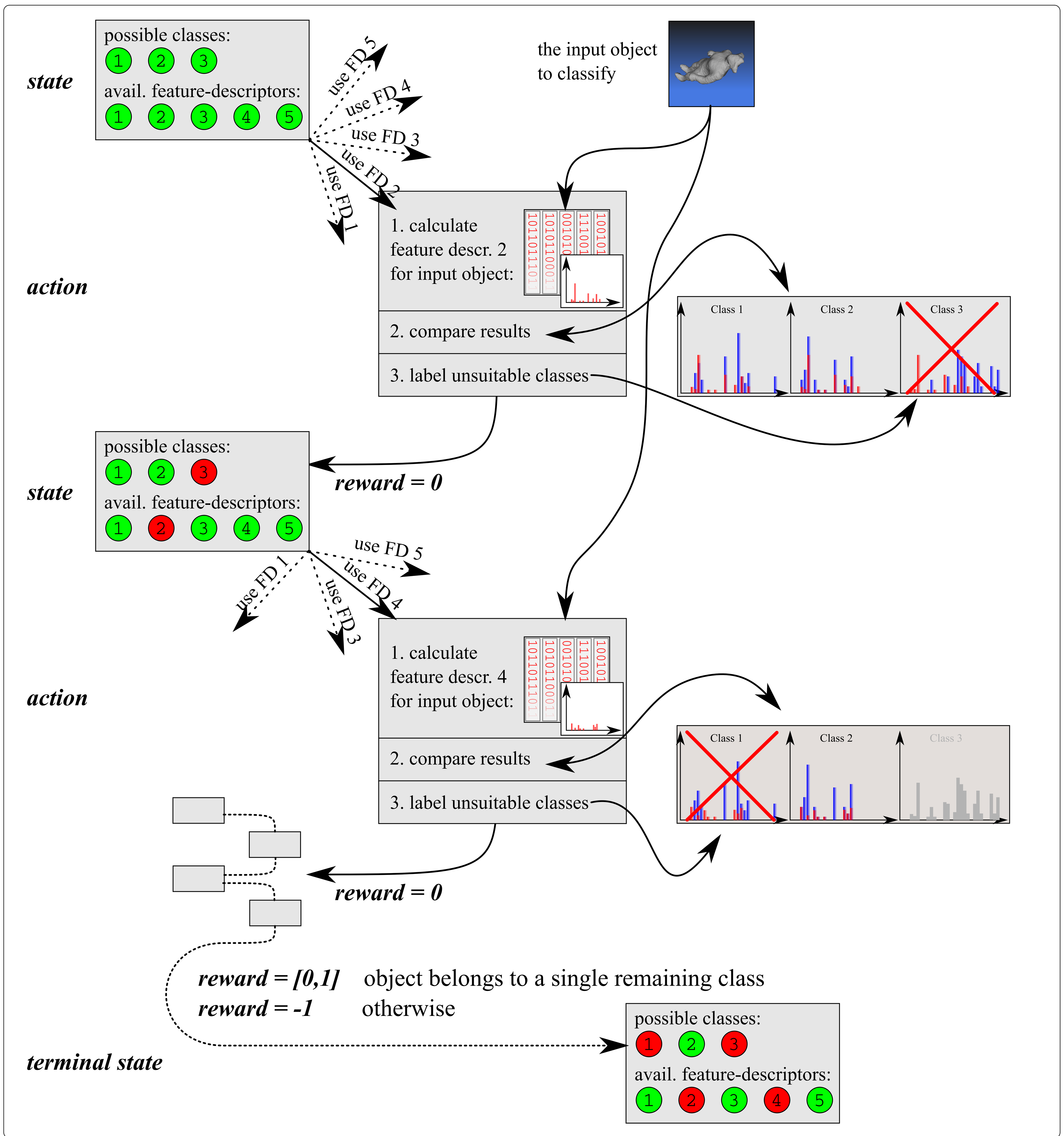
Building a dynamic set of feature vectors and object classes, e.g., cars, planes, cups, teddy bears, etc.:

- using point clouds of some objects per class
- using different types of feature descriptors for point clouds




## 2 Reinforcement Learning

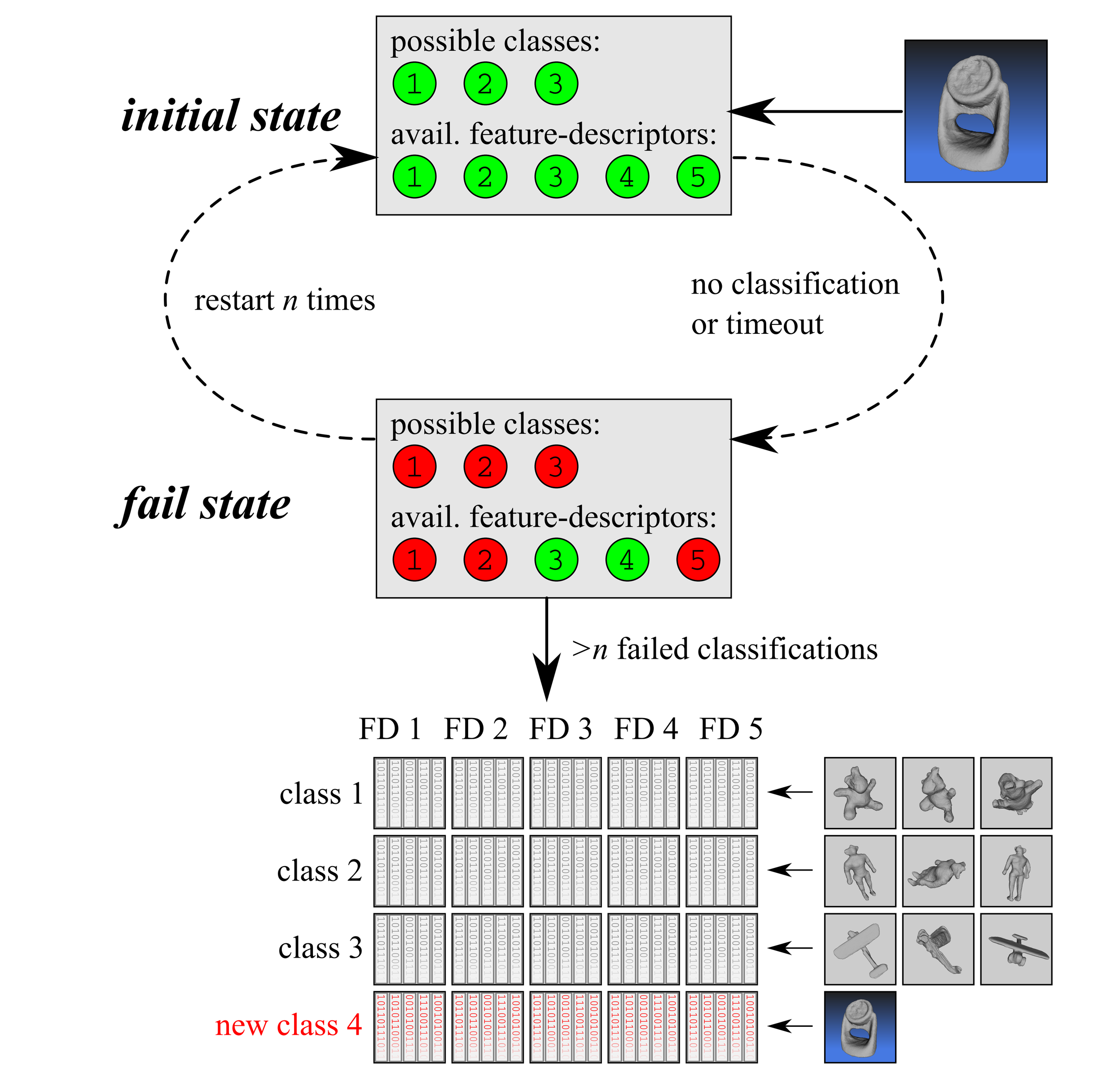
- **Task**
  The reinforcement learning (RL) uses periodic tasks. Each task ends, if
  - one class is remaining (successful classification),
  - no class is remaining (no classification), or
  - the learner runs into timeout.
- **Environment**
  The environment consist of
  - a dynamic set of object classes with preprocessed feature vectors, as presented in 1,
  - a set of available types of different feature descriptors that could be applied on the current input object, and
  - the input object.
- **Policy**
  While initially learning the first policy $\pi$, the RL selects randomly one of the available feature descriptors. During the subsequent course of classification the RL will use an adaptive $\epsilon$-greedy selection based on the learned policy $\pi$.
- **Action**
  The action based on the selected feature descriptor consists of the following steps:
  - The calculation of the feature vector(s) of the current object for the selected feature descriptor.
  - The comparison of the feature vector(s) against the set of classified feature vectors.
  - The labeling of all classes with an insufficient quality of matching feature vectors as unsuitable
- **Reward**
  The reward depends on the following rules:
  - If one class is remaining and the object belongs to this class, the reward value is calculated by linear interpolation of the time used for classification: $1$, if the time is zero, $0$ if the time equals the timeout.
  - In all other cases the reward value is $-1$.



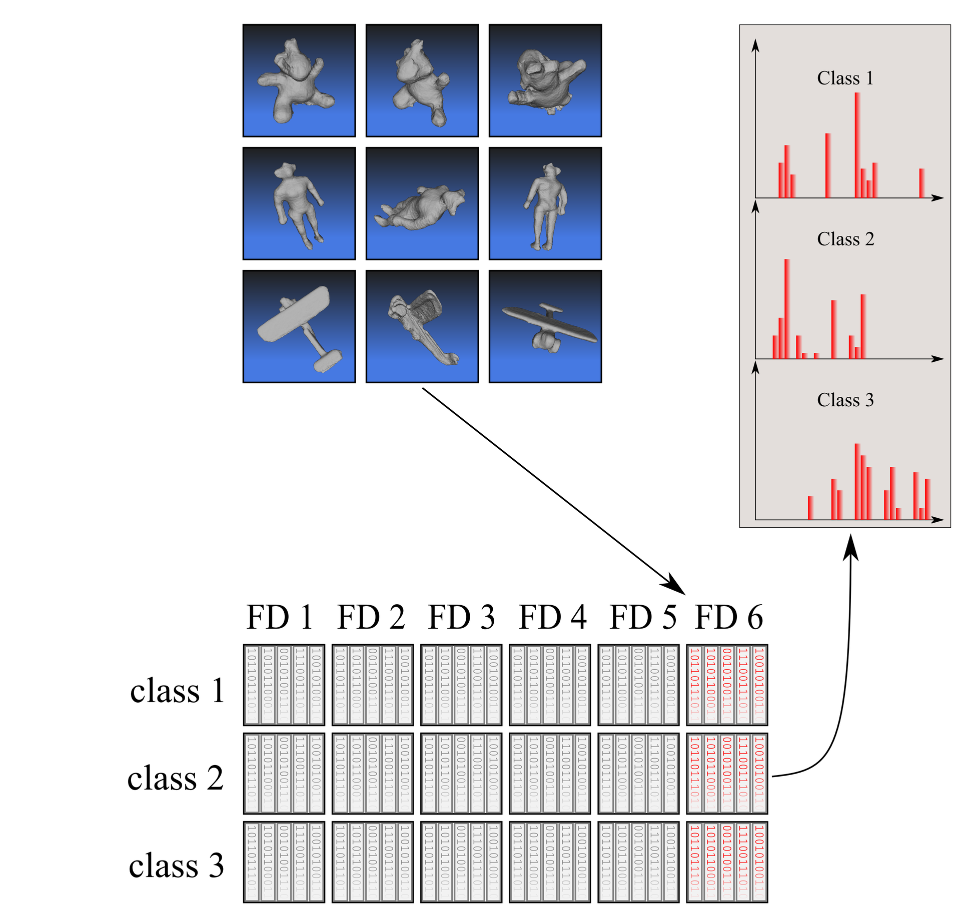

## 3 Classification and Online Learning

- **Classification**
  The RL follows policy $\pi$ with an $\epsilon$-greedy strategy to select actions, i.e., to select and apply feature descriptors. If a classification succeeds, the used feature vectors get added to the dynamic set of feature vectors.
- **New Unknown Object Classes** 3a
  In case all object classes are labeled as unsuitable (fail state), the process is repeated $n$ times while increasing the $\epsilon$-value. This leads to a high rate of randomly selected feature descriptors. If this additional iteration does not lead to a classification, a new object class is created *automatically*.
- **New Feature Descriptors** 3b
  New feature vectors are calculated for all classes learned so far. The $\epsilon$-greedy strategy *automatically* leads to the occasional use of the these new feature descriptors and an adaptation of the policy $\pi$.

## 3a Learning New Objects Online




## 3b Learning New Feature Descriptors Online





Contact: Jens.Garstka@FernUni-Hagen.de