



Multimedia Similarity Searching in Practice

Pivot selection techniques

Michal Batko

Faculty of Informatics
Masaryk University



Course Outline



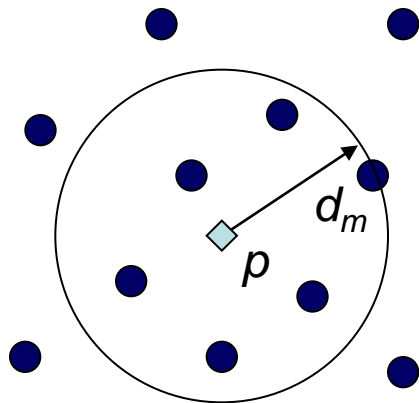
- Introduction, setup of the development environment
- Data collections and similarity functions
- Extraction of multimedia data descriptors
- Search and data manipulation operations
- Indexing algorithms, gathering statistics
- Command line interface, batches
- Data storage
- **Pivot selection techniques**
- Advanced index algorithms
- User and application interfaces

Pivot Selection

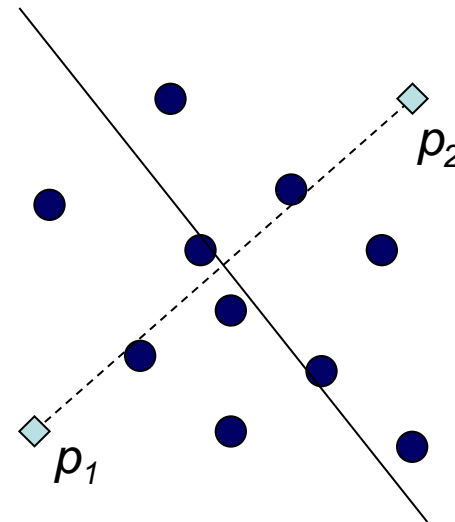


- Pivots
 - Reference objects
 - Used for example in partitioning

Ball partitioning



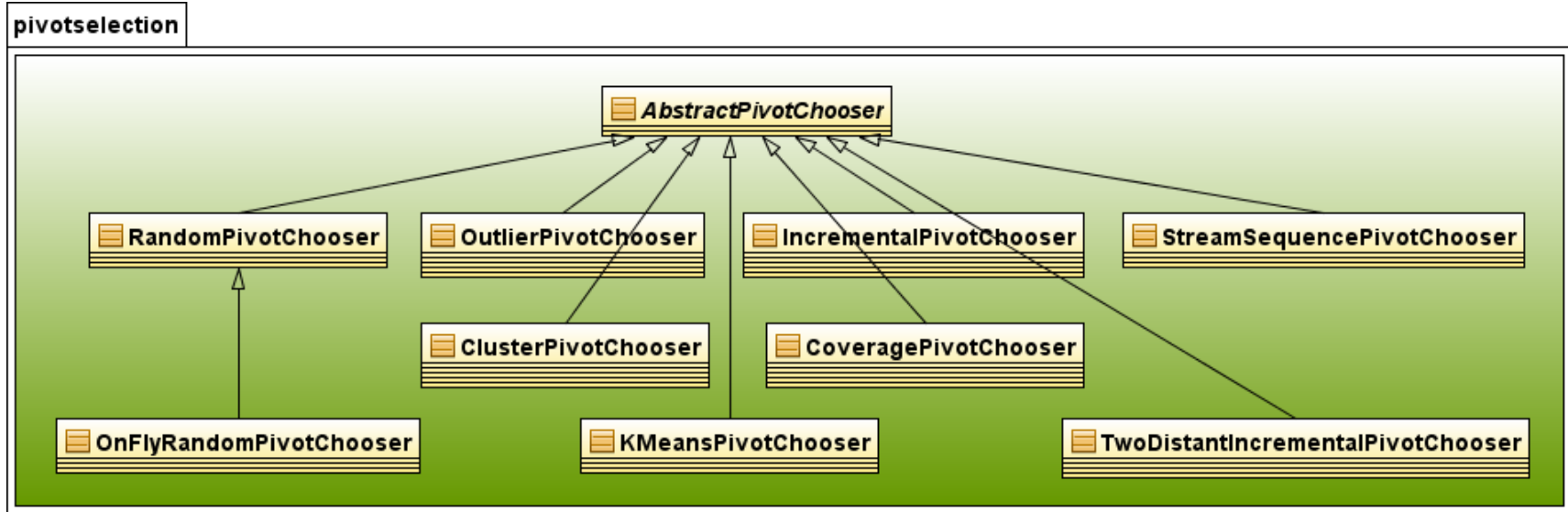
Hyperplane partitioning



Pivot Selection in MESSIF



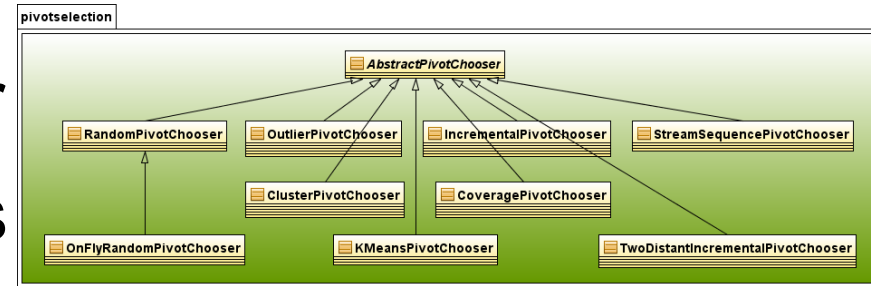
- Package `messif.pivotselection`



Pivot Selection Package



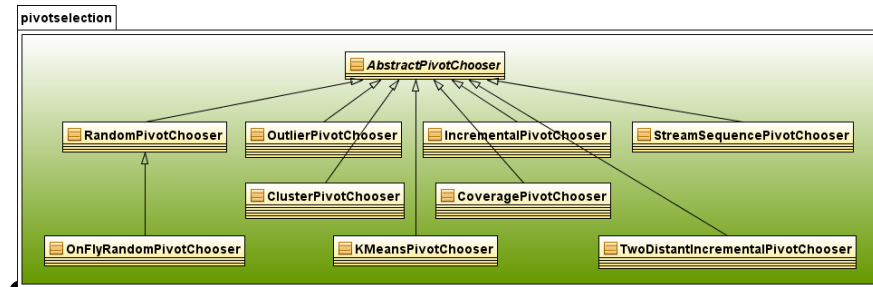
- AbstractPivotChooser
 - Ancestor of all classes
 - Selects pivots from sample set
 - registerSampleProvider method
 - selectPivot method
 - Retrieval by getPivot methods
 - getPivot, getPivot(position), getNextPivot(), getLastPivot()



Pivot Selection Package



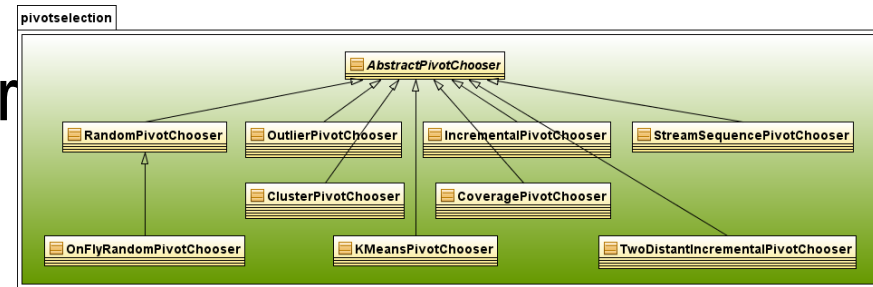
- RandomPivotChooser
 - Selects randomly
- OnFlyRandomPivotChooser
 - Registered as bucket filter – select pivots as objects are added to the bucket
- OutlierPivotChooser
 - first pivot is the object farthest from the other objects
 - second pivot is the farthest object from the first pivot
 - third pivots is the object farthest from the previous two
 - etc.



Pivot Selection Package



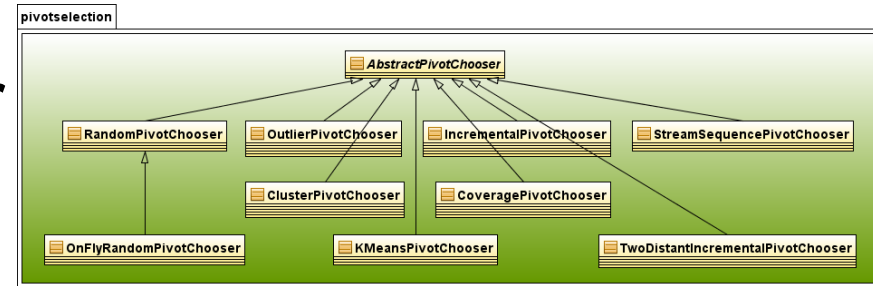
- IncrementalPivotChooser
 - Based on the technique Incremental selection
- StreamSequencePivotChooser
 - Returns pivots from a stream (no computation)
 - Wrapper class to map stream to pivot chooser
- ClusterPivotChooser
 - varying number of pivots based on cluster sizes which are limited by the parameter passed to the constructor



Pivot Selection Package



- KMeansPivotChooser
 - k-means algorithm for metric spaces



- CoveragePivotChooser
 - coverage of data by ball regions with centers in pivots and radius equal to a fixed value (threshold)



Questions?