Feature selection algorithms for fast video search, classification, and retrieval

Technology #m04-053

To accelerate searching for concepts of interest in video data, the technology provides a set of algorithms for selection of features useful for video search and exploration. These algorithms automatically and efficiently construct a hierarchy tree of feature subsets using a novel combination of existing dimensionality reduction and classification methods to determine how subsets reduce the error of concept prediction. To reduce tree construction time, heuristic parameters may be set that control which selected parts of the hierarchy are set up.

Fast construction of feature subset hierarchy enables efficient search and classification of video data that is robust against noise and sparseness

Algorithms for searching through large datasets for concepts of interest often employ techniques to identify, extract, and rank features that can be used to accelerate the search process by reducing the dimensionality of the data in question. Although effective feature selection methods have been devised for a variety of datasets such as text and genomic data, these methods are less useful when applied to video data because of their long run times. This technology combines advantages of forward selection, backward elimination, and genetic algorithm heuristic methods to rapidly select small but well-chosen feature subsets. A novel combination of Fastmap and the Mahalanobis measure are respectively used for dimensionality reduction and classification. This technology reduces the time required to construct a hierarchy of useful feature subsets using evaluation metrics to prune unnecessary parts of the hierarchy.

The accuracy and efficiency of this technology has been used to successfully retrieve frames of interest from within a variety of different videos.

Lead Inventor:

John R. Kender, Ph.D.
Applications:

- Fast search and retrieval of data from videos
- Search and analysis of audio data
- Video navigation systems
- Drug discovery
- Data mining
- Micro-array analysis
- Face detection and recognition in surveillance systems

Advantages:

- Fast execution time makes online video search/retrieval possible
- Effectively handles sparse and/or noisy training data
- Exhibits linear scaling with respect to the number of features
- Can be used as an effective machine learning method for efficiently searching through non-video datasets

Patent Information:

Patent Issued (US 7,277,574)

Tech Ventures Reference: IR M04-053, M04-054, M04-055

Related Publications:


Inventors

John Ronald Kender