

TWO-STEP DATA MINING METHOD FROM DIGITAL IMAGES

Mgr inż. Anna Wójcicka

SUMMARY

The main aim of this work is to develop a two-step method of extracting knowledge from digital images. The method integrates the analysis of digital images directed at the extraction of quantitative and qualitative characteristics and knowledge extraction methods. The proposed method let to conduct a new kind of exploratory research aimed at extracting knowledge from digital images. This method allows to create rule set knowledge bases for decision support systems.

Analysis of a coherent series of images allows the extraction of interesting qualitative and quantitative features. Put their further exploratory analysis may allow for the discovery of regularities occurring, generalizations, connections and relationships. The paper presents the research literature on both the available algorithms analysis and processing images as well as methods of exploration of knowledge.

Realized computer program implements the proposed method, while providing the environment of its experimental verification. Checks on the correctness of the system were carried out on real data: computer tomographic images of the surface of the tooth, dermatoscopic skin cancer images and microscopic images of Friction Stir Welding joints.

The proposed method of data mining of digital images was divided into two steps. The first step uses the selected methods of image analysis, focused on the extraction of quantitative and qualitative characteristics of objects presented on images. The basic input data for the selected methods are a series of images depicting the analyzed objects. At the step of extraction of characteristics:

- Determined their number, type, the names of the attributes and names or ranges of features.
- Establish a set of graphical transformations to be carried images to their standardization and to obtain the required characteristics.

The result of this stage are data table - information system. These data are subjected to pre-processing, covering the processing of missing, outliers, digitization data. Pre-processed data make decision table. In the table determined attribute decision-making and conditional attributes. The decision table is input for the second step of the proposed method. This step involves data mining, ended generating decision rules. This process is preceded by an analysis of the consistency of the two available methods: qualitative and quantitative. Data mining is

based on an approach based on rough set theory. The result of exploratory research are the decision rules generated by various methods. Cross validation allows to carry out the quality of the method. At each stage it is possible to support the use of domain experts who are using a dedicated system could verify the results obtained with regard to input images.

The last element of the method is the ability to use established knowledge base for the implementation of subject decision support system. This is done by inference forward module. The inference can be used for both practical and experimental verification of the received rule base, as well as to achieve ready to implement the user system.

The proposed method can maximize the automation capabilities of acquiring knowledge of the images, while allowing for the use of knowledge and competence domain experts.