An Information System for Material Microstructures

Kathrin Roberts*, Frank Mücklich**, Ralf Schenkel***, Gerhard Weikum***

* University of Karlsruhe	** Saarland University	*** Max-Planck Institute of
Dept. of Computer Science	Dept. of Material Science	Computer Science
D-76131 Karlsruhe	D-66123 Saarbruecken	D-66123 Saarbruecken
Germany	Germany	Germany
roberts@ira.uka.de	muecke@matsci.uni-sb.de	{schenkel,weikum}@mpi-sb.mpg.de

Abstract

This paper presents an information system that supports a materialographic laboratory in classifying material samples based on microstructure images. The system uses database and Web technologies to manage its information and make it accessible to Internet users. Its core is a classifier, based on support vector machines, that provides an automatic diagnosis of the material class of a given sample. The classifier uses texture features from an underlying image analysis, the so-called Haralick parameters, and stereologic features such as fractal dimension, Euler parameter, etc. In addition to the classifier, the system provides a sensitivity analysis that allows the user to understand which features are most influential for certain classification decisions. The system is fully operational and can be used on the Web.

1. Introduction

Materialography is the science of analyzing and synthesizing the microstructure of functional materials such as ceramics, steels, or multi-layer hybrid materials. A common task for a material laboratory is to analyze a sample material based on a microscopic image of the sample's microstructure. A concrete example is cast iron with lamellar graphite, which is widely used in the automobile industry, e.g., in motor blocks or brake discs, and mechanical engineering in general. When designing and eventually casting such components, material properties like abrasiveness, absorbability, etc. play a critical role. These do in turn depend on the microstructure of the material, as observed in microscopic or diffractometric images. Figure 1 shows a microscopic image of some piece of cast iron. Material engineers have developed a classification system for many of these functional materials, based on their microstructures. In the case of cast iron there is a standard, EN ISO 945:1994, which divides all kinds of cast iron into five categories, A through E, according to the structures formed by the included lamellar graphite. Figure 2 shows these five classes.

Classifying a given material sample into one of the applicable categories is a challenging task even for an experienced material engineer. Thus, it is highly desirable to apply automatic classification techniques to this problem, for two reasons: first to provide a more objective and quantitative basis for classification, and second to aid less experienced materialographs with this task. In this paper we outline an information system that manages microstructure images in a database of materials, and allows Internet users to either interactively train themselves with pre-classified images or upload their own images and have them automatically classified.

In Section 2 we sketch the architecture of the overall system. In Section 3 we discuss the classifier, based on support vector machines, and the underlying feature space, using texture and stereologic parameters. In Section 4 we present an interface and algorithms for sensitivity analysis, which is useful to understand the most influential features that contribute to the classification result for a given material sample. Section 5 contains experimental results, using the five classes of cast iron as a case study.
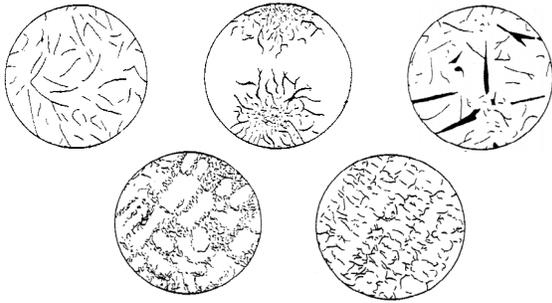

Fig. 1: Sample image of cast iron

Fig.2: Five classes A through E of lamellargGraphite

2. System Architecture

Our system is based on a three-tier architecture with an HTML plus Javascript Web front-end, a Java servlet middle tier running Apache and Tomcat, and a database back-end using Oracle 9i. The servlets invoke external C code for compute-intensive tasks, namely, the extraction of features from the microstructure images, and for external components.

We use the ImageMagick toolkit [4] and the ImageC product [1] for preparatory steps on the images, and the SVMlight classifier [5]. Typically, the original images are in JPEG gray-scale format, but our system requires binary images (i.e., only black or white pixels) in PBM format. ImageMagick provides the conversion routines, and ImageC performs a gray-scale shading correction before the conversion.

The database consists of 10 relations that contain images as BLOBs, the extracted features as numerical values (see Sections 3.2 and 3.3), training data and parameters, results of invoking the classifier, etc.

The system is operational and accessible at the URLs www-dbs.cs.uni-sb.de/~roberts/bildanalyse.html for examples and www-dbs.cs.uni-sb.de/~roberts/servlets/klassifikation for trials with your own images.[1] More details on the implementation can be found in [9].

3. Material Classification

3.1. Support Vector Machines

Our classification is based on support vector machines (SVMs) [2,10], which operate on numerical feature vectors with m=20 features characterizing statistical properties of the pixels in a microstructure image. We use 14 so-called Haralick parameters [3] that capture texture properties, and 6 stereologic parameters [8] that capture stochastic geometry properties of the images

---

[1] These direct links may be temporarily unstable; the anchors in http://www.materialography.de/ should be stable.

(or, more generally, the 3D volumes of which the images are 2D slices).

We use one SVM for each class that we are interested. The SVM for class c learns a decision function f: $R^m \rightarrow \{+1,-1\}$ where $f(\vec{z}) = 1$ means that a previously unseen test vector $\vec{z}$ belongs to class c. We use the linear form of SVM where training amounts to finding a hyperplane in the m-dimensional feature vector space that separates a set of positive training examples from a set of negative examples. As negative examples for class c we use the positive examples from all classes other than c. The hyperplane can be described in the form $\vec{w} \cdot \vec{x} + b = 0$, with the components of vector $\vec{w}$ and the scalar b being the parameters to be learned by the SVM. Computing the maximum-margin hyperplane, which maximizes the distance of the closest positive or negative training point, is equivalent to solving a quadratic optimization problem. Our implementation uses the very efficient SVMlight package [5] where training is typically between $O(n^2)$ and $O(n^3)$ for n training vectors.

The decision of the SVM classifier for a new test vector $\vec{z}$ is very efficient. We merely need to test on which side of the hyperplane $\vec{z}$ lies, which requires computing the sign of the m-dimensional scalar product $\vec{w} \cdot \vec{z} + b$ and setting $f(\vec{z}) = 1$ for positive sign and $f(\vec{z}) = -1$ for negative sign. The Euclidean distance of the test point to the hyperplane can be viewed as a measure of the classifier's confidence.

3.2. Texture Features

The Haralick parameters [3] are a widely used feature set for characterizing image textures. They are based on the gray-scale transitions between adjacent pixels. In our setting, with black and white pixels only, we can simplify the computation. For each pixel we consider its 8 neighboring pixels and compute certain values, which are then aggregated over all pixels in the image. These values are the numbers of white-white, white-black, black-white, and black-black transitions in horizontal, vertical, and the two diagonal directions. The Haralick feature set contains averages, variances, correlations, and entropy measures over these statistics. The details of our computations are in [9].

3.3. Stereologic Features

Stereologic features are based on connected pixel sets, p-sets for short, where a p-set is a maximal set of pixels with two properties: 1) all pixels have the same color, and 2) for each pair of p, p' of pixels there is a path p=p_1, p_2, …, p_n=p' such that p_{i+1} is one of the

eight neighbors of p_i. We can now compute properties of each black p-set such as area, perimeter, curvature, etc, and then build statistics over all p-sets of the image. Two of these parameters that are particularly worthwhile mentioning are: 1) the Euler number, which is the number of p-sets in an image minus the number of holes where a hole is a p-set of the complementary color within a p-set, and 2) the fractal dimension, which essentially measures the ratio of p-set areas under different magnification scales. More details on how to compute the fractal dimension and the other stereologic parameters can be found in [7]. Note that these parameters actually serve to reconstruct the 3D structure of some material based on 2D data about slices, hence the name "stereologic"; in our context we use these computations in a much simpler way.

4. Sensitivity Analysis

Although we aim at automatic classification of microstructure images, the underlying feature set is, of course, still only a crude model of the physical properties of the material class, and our users are experts who want to interactively "play" with the system. To this end, when classifying an image we provide additional output about the SVM decision models that were learned from certain training sets and the actual classification decisions made for the given image:

- the estimated precision, recall, and accuracy of each classifier, determined by leave-one-out testing on the training set (i.e., training with n-1 out of n training vectors and testing the n-th vector, averaged over the n possible choices for the left-out test vector),
- or, alternatively, the estimated precision, recall, and accuracy, determined by the ξ-α estimator [6], which is a cruder estimator but avoids retraining the SVMs n times,
- the positive or negative distance of the test vector from the hyperplane, relative to the mean distance of positive and negative training vectors and the standard deviation of these distances.

The latter is graphically illustrated by the system; a sample screenshot of the output for classifying a given image is shown in the appendix.

In addition to these outputs, the system supports two kinds of sensitivity analyses that are helpful for the expert user:

- varying the training sets by starting from a small set and incrementally increasing it, showing the resulting precision, recall, and accuracy at each stage, and

- varying the feature sets by starting from the full set and incrementally dropping features, again showing the resulting classification quality measures at each stage.

The second kind of sensitivity analysis is particularly useful for identifying the most influential (or even *the* decisive) features for certain classification decisions. Driven by the consideration that there is an exponential number of feature subsets that we could explore, we developed a simple greedy algorithm that works as follows:

```
feature set F := all features;
while |F|>1 do {
    for each f∈F do {
        train classifier with F−{f} and
        evaluate its precision, using leave-one-out
testing;
        remember the feature f* for which the precision
        of F−{f*} exceeds those of all other F−{f}; };
    F := F −{f*} };
```

This greedy algorithm yields a relative order of importance among the features, with the most influential feature remaining last in F. It may seem that standard methods from multivariate statistical analysis and data mining, most notably, Principal Component Analysis (PCA) would give us the same kind of result in a much more elegant and better founded way. However, with PCA the "influential dimensions" are linear combinations of the original feature dimensions; so the result cannot be easily interpreted by the user. Even though we have expert users in mind, it is all but easy to conceptually "map" linear combinations back to individual features; so PCA would of no help in our application context.

Figure 3 shows a sample output of this analysis, for classifying class B graphite. Each line represents one feature set, with the green (light) boxes in the left part denoting present features and the red (dark) boxes denoting discarded features. The blue (dark) bars in the right part indicate the resulting precision and recall. Here, the remarkable result is that both precisions and recall stay very high until the last 4 remaining features, 3 of which are stereologic parameters. The feature that was kept until the end was the average area of the p-sets. When we showed this to materialography experts, they found this result to be very much in line with their understanding of class B graphite structures, but they said that they would not be able to easily predict this outcome a priori. Also, these effects and the most influential features varied quite a bit among the different classes of graphite, and may exhibit even higher variability when applied to other types of

materials (note that our feature set is universally applicable and in no way tied to cast iron).

Fig. 3: Output of the feature sensitivity analysis

5. Experimental Results

We ran experiments with a set of 350 microstructure images of cast iron with lamellar graphite, which were pre-classified by human experts into the five categories A, B, C, D, and E [8]. Some of the images contained multiple graphite classes and were considered to be acceptable for each of their corresponding classes. We used leave-one-out evaluation to measure precision, recall, and accuracy. With the contingency table of classifier f for class c shown in Figure 4 (e.g., n^{++} denoting the number of positively classified test vectors that do indeed belong to class c), precision is defined as $n^{++} / (n^{++} + n^{+-})$, recall as $n^{++} / (n^{++} + n^{-+})$, and accuracy as $(n^{++} + n^{--}) / (n^{++} + n^{+-} + n^{-+} + n^{--})$ (and error is $1 -$ accuracy).

	$\vec{z} \in c$	$\vec{z} \notin c$
$f(\vec{z}) = 1$	n^{++}	n^{+-}
$f(\vec{z}) = -1$	n^{-+}	n^{--}

Fig. 4: Contingency Table for Classifier Result

Figure 5 shows our experimental results for the five classes of lamellar graphite in cast iron. For four out of the five classes we obtained very good results; for two of them, A and E, the results were excellent. For class C, on the other hand, our approach obviously failed. Our explanation for this failure is twofold: 1) for this class we had an order of magnitude more negative training samples than positive ones (this also explains the good accuracy despite the zero precision and recall) which introduced a bias in the corresponding SVM, and 2) our set of features, texture and stereologic parameters were insufficient to characterize this particular type of lamellaes.

	A	B	C	D	E
Precision	0.976	0.807	0	0.849	0.966
Recall	0.993	0.700	0	0.725	0.691
Accuracy	0.972	0.880	0.898	0.858	0.843

Fig. 5: Classification quality for lamellar graphite

6. Concluding Remarks

The information system that we developed to support material laboratories for managing microstructure images has so far received very positive comments by its users. In particular, the presented form of feature sensitivity analysis has been appreciated.

Nevertheless, we realize that it still has many limitations. Most importantly, the set of 20 features that we considered, albeit forming a good core set, needs to be extended. This requires close collaboration with material scientists to understand further phenomenological properties that characterize the type of a material's microstructure. Another aspect that we need to address is the fact that a microstructure image may contain multiple material classes in different spatial regions (Figure 1 actually shows an example). This calls for appropriate partitioning of the image, and the challenge lies in finding adequate strategies.

References

[1] Aquinto AG, http://www.aquinto.de
[2] C.J.C. Burges: A Tutorial on Support Vector Machines for Pattern Recognition, Data Mining and Knowledge Discovery, Vol.2 No.2, 1998, pp 121-167.
[3] R.M. Haralick: Textural features for image classification, IEEE Transactions on Systems, Man, and Cybernetics, 1973, pp 610-621.
[4] ImageMagick, http://www.imagemagick.org
[5] T. Joachims: SVM light, http://svmlight.joachims.org
[6] T. Joachims: Learning to Classify Text using Support Vector Machines, Kluwer, 2002
[7] F. Mücklich, J. Ohser: Statistical Analysis of Micro-structures in Material Science, John Wiley & Sons, 2000.
[8] J. Ohser, K. Sandau, W. Stets, W. Gerber: Image Analytical Characterization of Graphite in Grey Cast Iron and Classification of Lamellar Arrangement, Practical Metallography, 2003.
[9] K. Roberts: Automatic Classification of Material Micro-structure Images, Master Thesis, Saarland University, 2003.
[10] V.N. Vapnik: The Nature of Statistical Learning Theory, Springer, 1995.