An Algorithm for Identifying Texture Components in the Framework of Statistical Crystal Plasticity Models

K V Ostapovich a, P V Trusov and A Yu Yanz

Department of Mathematical Modelling of Systems and Processes, Perm National Research Polytechnic University, 29 Komsomolsky prospekt, Perm 614990, Russian Federation

E-mail: a ostkirvad@gmail.com

Abstract. An approach based on involving clustering techniques to identify texture components for a given polycrystalline aggregate is formulated. A heuristic iterative clustering algorithm operating only with a sample of lattice orientations and the concept of closeness between them is proposed for this task. The mentioned closeness concept is introduced via a special pseudometric distance, which is induced by the natural Riemann metrics and takes the lattice symmetry into consideration. The developed procedure allows representing an inhomogeneous orientation distribution by localized clusters in the space of orientations and thus can be used to reduce the dimension of orientation distribution data in statistical crystal plasticity models. Two examples of such an application of the clustering procedure are provided.

1. Introduction

Polycrystalline metals and alloys consist of anisotropic elements with lattice structure, i.e., crystallites, the orientations of which have a significant influence on macroscopic properties of such materials. During severe inelastic deformation, the crystallites can rotate leading to inhomogeneity in the orientation distribution, i.e. the formation of a crystallographic texture. The analysis of the texture plays an important role in material science problems related to the study and description of the interdependence between the evolution of the microstructure and changes in macroscopic properties of polycrystals.

Quantitative description of texture formation is generally based on the methods of multilevel modeling and explicitly or implicitly refers to the orientation distribution density in a polycrystalline representative volume element or its statistical model, i.e., the polycrystalline aggregate. This density (or the corresponding measure) determines volume fractions of a crystal lattice in various orientational configurations. For its numerical representation in most applications, certain reduction techniques are being involved. Some approaches [1–3] include the Fourier expansion with respect to the symmetrized generalized spherical harmonics; the proper orthogonal decomposition and the snapshot method, which is its simplified analogue; the texture components method; and omnifarious techniques based on the explicit discretization of the orientational space.

The mentioned methods and their modifications are in widespread use in texture analysis, in particular, for the purpose of calculating orientation distribution densities by the inversion of diffraction pole figures. In [4–11], these approaches are applied to formulating reduced-order models of texture formation during deformation. Such models are of significant interest in polycrystalline material design. Relevant inverse boundary value problems in this field require solving repeatedly the equations for microstructure evolution during various thermo-mechanical processes. Due to high
resource-intensity of these problems, the possibility of compressing orientation distribution data is quite important. One of the ways to reduce such information in statistical polycrystalline models, which is developed in this work, is based on identifying pronounced texture components by using clustering techniques [12–13]. Here, an emphasis is placed on the formulation of a heuristic algorithm appropriate for this task.

Clusters are generally treated as disjoint subsets of similar (close) elements. The procedure described below allows distinguishing orientation clusters for a given polycrystalline aggregate by considering only its sample of orientations. In fact, these clusters can be considered as regions in the space of orientations, in which the distribution density is higher.

2. A procedure for clustering orientations

One should note that almost all clustering schemes in one way or another refers to the concept of closeness between data elements. In the current case, these points are the orientations of a crystal lattice, the geometric symmetry of which needs to be taken into account. Such a problem is solved in this work by introducing a special pseudometric distance for comparing orientations. Unlike a metric one, a pseudometric distance does not assume the non-degeneracy on unequal arguments. Moreover, it should be equal to zero if and only if the orientations being compared are geometrically indistinguishable from each other.

Consider a single-phased polycrystalline aggregate of $M$ crystallites with the same volume fractions. Their orientations can be defined by the elements, $o^{(m)}$, $m \in [1, M]$, of the special orthogonal group, $O^+$. Let $S^+ \subset O^+$ be the rotational symmetry group of the given lattice type, so any $o \in O^+$ is geometrically indistinguishable from so for all $s \in S^+$, i.e., symmetrically equivalent to them. The pseudometric distance, $d[\cdot, \cdot]$, between arbitrary $o_1, o_2 \in O^+$ is introduced as

$$d[o_1, o_2] = \min_{s \in S^+} d_s[o_1, o_2],$$

where $d_s[\cdot, \cdot]$ is some metric distance. As for this distance, it is convenient to use the one $[,]$ induced by the natural Riemann metrics. In such a case, by adopting the quaternion parametrization, one can write:

$$d[o_1, o_2] = 2 \min_{s \in S^+} \arccos \left| \Re o_1 \Re o_2 + \Im o_1 \cdot \Im o_2 \right|$$

$$- \Im \left| \Re o_1 \Im o_2 - \Re o_2 \Im o_1 + \Im o_1 \times \Im o_2 \right|.$$  \hspace{1cm} (2)

Here, $\Re o$ and $\Im o$ denote, respectively, the real and the vector (imaginary) parts of the quaternion corresponding to $o \in O^+$. This distance determines the minimal angle of rotations from one orientation-argument to any symmetrically equivalent one of another.

Due to high complexity of the space being considered, a clustering approach based exclusively on the comparison of the sample orientations seems to be quite performant. In such a framework, it is reasonable to introduce the misorientation matrix, $\{D_{mn}\}$ (of $M \times M$ size):

$$D_{mn} = d[o^{(m)}, o^{(n)}].$$  \hspace{1cm} (3)

This is a symmetric matrix with zero main diagonal, which contains almost all information required for applying abstract clustering methods of a rather wide class.

2.1. Extracting a layer

Generally, the exact amount of clusters for a given texture is not known a priori and the possibility to distinguish them is quite suppositional. In order to obtain an adequate partition, one should associate the locations of clusters with the regions of higher orientation distribution density. Appropriate techniques are based on grouping close orientations (or maybe their neighborhoods) into distinct subsets. However, such methods are usually sensitive to the “noise” created by rarefied orientations. To overcome this flaw, it is proposed to successively exclude from the sample under consideration the
layers of the orientations that have the highest density in their neighborhoods. Each of the layers obtained in such a way is supposed to be further independently clustered.

Let $M_{l\varepsilon}$ be the number of the orientations that left in the sample after excluding the $l$-th layer and belongs to the pseudo-metrical $\varepsilon$-neighborhood of $o^{(m)}$. In the framework of the approach being described, one constructs a layer by choosing a lower bound, $M_{l\varepsilon}$, and grouping the orientations with $M_{l\varepsilon} \geq M_{l\varepsilon}$.

### 2.2. Clustering by reachability

Orientations of a layer are considered as the vertexes of a connectivity graph. The connectivity criterion is defined as the closeness between orientations or their neighborhoods. In the simplest case, $o^{(m)}$ and $o^{(n)}$ belongs to an edge if and only if $D_{mn} \leq \varepsilon$. The maximal connected subgraphs of the introduced graph are treated as the initial clusters.

The selection of appropriate values for $M_{l\varepsilon}$ is restricted by the variations of distribution density in the $\varepsilon$-neighborhoods and hence depends on the texture under consideration. To define them, a histogram analysis for the distributions of $M_{l\varepsilon}$ over sample orientations at different layers can be engaged. Here, it seems reasonable to choose $\varepsilon$ in such a manner that the amount of orientations in almost each $\varepsilon$-neighborhood is statistically sufficient for introducing their local density (with the exception of relatively few neighborhoods containing rarefied orientations). In the case of the uniform distribution, the fraction, $\mu_0$, of orientations within an $\varepsilon$-neighborhood under the condition that there are no pairs of symmetrically equivalent elements (which can be always achieved by choosing not too large values of $\varepsilon$) is given by

$$\mu_0 = \frac{S}{\pi} \left( \varepsilon - \sin \varepsilon \right),$$

where $S$ is the cardinality of $S^\dagger$.

### 2.3. Medoid-based reclustering

Clustering by reachability does not provide an efficient control of the forms, sizes and amount for the clusters being obtained. Meanwhile, in order to simplify the reduced representation of a texture, the possibility to distinguish well enough localized clusters with pronounced centers is of a significant interest. Due to this, it seems reasonable to introduce an additional procedure for identifying the concentration points of orientations in clusters and then, if necessary, performing a repartition around them. This task is resolved in this work by applying the so-called medoid-based reclustering.

The medoid of a set is understood as its element such that the mean pseudometric distance from this element to other elements in the set is minimal. Medoid-based reclustering on the $l$-th layer with the fixed number, $K_l$, of clusters being formed can be considered as a heuristic procedure to solve the following discrete optimization problem. It is required to find the indexes, $m_{lk}$, $k \in \{1, K_l\}$, of orientation-medoids and the discrete function, $c_{l}[\cdot]$, which maps orientation indexes onto corresponding cluster indexes, such that the objective,

$$J_l^{K_l} = \sum_{k=1}^{K_l} \sum_{m \in \{1, \ldots, |S|\}} D_{mnk},$$

reaches its minimum value.

The minimization strategy is realized by alternating between determining medoids of current clusters and reassigning each orientation to the cluster with the closest medoid. One should note that, in general, such an approach provides only a sub-optimal (i.e., unimprovable by any further iterations) solution. Other optimization techniques [14] can be also involved to solve this problem.
2.4. Splitting clusters

To get a partition such that each cluster is well enough localized in the sense that, e.g., the mean pseudometric deviation of its orientations from the medoid does not exceed a critical value, \( \delta_{\text{avg}} \), the described procedure is iterated. Herewith at each step, the least localized cluster is splitted by the pair of its most distant orientations which are treated as new medoids.

The scheme of the clustering procedure described above is shown in Figure 1. This operation sequence is repeated until there are no orientations left unassigned to any cluster. The orientation clusters obtained by the proposed algorithm can be treated as texture components. Their local characteristics give a reduced representation of the orientation distribution in a polycrystalline aggregate.

More specifically, the orientation distribution measure, \( F[\cdot] \), of the given polycrystalline aggregate is formally defined for any Borel subset, \( O \in O^+ \), as

\[
F[O] = \frac{1}{M} \sum_{m=1}^{M} o_{m}^{(n)} \Delta_{\delta=\varepsilon}[O].
\]

Here, \( \Delta_{\delta}[\cdot] \) denotes the Dirac measure concentrated in \( o \in O^+ \), i.e., \( \Delta_{\delta}[O] = 1 \) if and only if \( o \in O \) and \( \Delta_{\delta}[O] = 0 \) otherwise. Further, let \( L \) be the amount of layers, \( K_{l}, l \in [1, L] \) be the amounts of clusters at the layers, \( M_{l_k}, l \in [1, L], k \in [1, K_{l}] \) be the cardinalities of the clusters and \( m_{l_k}, l \in [1, L], k \in [1, K_{l}] \) be the indexes of the corresponding medoids. Then one can propose the simplest variant for a reduced orientation distribution measure, \( \hat{F}_{l_{1}, l_{2}, \ldots} \), in the form:

\[
\hat{F}_{l_{1}, l_{2}, \ldots} = \frac{1}{M} \sum_{l=1}^{L} \sum_{k=1}^{K_{l}} M_{l_k} \Delta_{\delta=\varepsilon_{l_k}}[O],
\]

where \( L \) is the amount of the layers taken into consideration.

![Figure 1. The scheme of the clustering procedure. Dashed arrows point at conditioned stages.](image)

3. Some examples

In this section, examples of applying the clustering procedure to two numerically obtained textures are presented. The two-scale elasto-visco-plasticity model [15] is used in simulations. Kinematic loads of a copper (the face-centered cubic lattice) polycrystalline aggregate are realized in a simple shear and an one-axial tension tests with a constant velocity gradient. The aggregate is consisted of \( M = 1000 \) crystallites with initially uniformly randomly distributed orientations. The clustering parameters are set as \( \varepsilon = \delta_{\text{avg}} = 0.158 \) (\( \mu_{0} = 0.005 \)).

To represent input and output data of the procedure in each test, direct and reduced (characterizing cluster cardinalities) pole figures are plotted for the \{100\} family of planes. One should note that these figures are used only for the purpose of visualization since the formulated algorithm operates directly...
with orientations, not the projections of the corresponding directions. To specify characteristic axes of loads, a stationary (laboratory) coordinate system with the biorthogonal basis vectors, \( \mathbf{1}, i \in \mathbb{1}, \mathbb{3} \), is defined.

### 3.1. Simple shear
In this case, the shear along \( \mathbf{1}_2 \) in the plane perpendicular to \( \mathbf{1}_3 \) is simulated. The sample of obtained orientations is depicted in figure 2a. Partitioning into \( L = 7 \) layers is used for clustering. However, the orientations of the last layer (13.6% of the total amount) are highly rarefied so the procedure distinguishes a large number of degenerate “clusters” consisting of only one element. This layer is excluded from consideration. The medoids of the first 6 clustered layers are shown in figure 2b.

![Figure 2](image)

**Figure 2.** Pole figures for the simple shear texture: the direct pole figure for the aggregate orientations (a) and the reduced pole figure for the identified cluster medoids (b)

The adequacy of the resulted medoid-based representation is verified for aggregating elastic properties. The Voigt averages of the elasticity tensors are evaluated with the full orientation distribution measure (6) and with its reduced analogue (7) for \( L = 6 \). The difference between the corresponding responses defined by the anisotropic Hooke’s law is estimated to be less than 10%.

### 3.2. One-axial tension
In this test, the tension along \( \mathbf{1}_3 \) is simulated. The sample of obtained orientations is presented in figure 3a. Clustering with \( L = 6 \) layers is performed in this example. By reasoning as for the previous case, the last layer (8.1% of the total amount) is omitted. The medoids of the first 5 clustered layers are given in figure 3b.

![Figure 3](image)

**Figure 3.** Pole figures for the one-axial tension texture: the direct pole figure for the aggregate orientations (a) and the reduced pole figure for the identified cluster medoids (b)
The verification of the adequacy of this medoid-based representation is carried out in the same way as before. The use of the reduced measure (7) with \( \hat{L} = 5 \) instead of the full measure (6) leads to an error in the response, which is less than 5%.

4. Conclusion
To sum up, this work is devoted to the formulation of the heuristic iterative algorithm for identifying texture components by analysing the sample of lattice orientations with the help of clustering techniques. The proposed procedure consists of the following stages: extracting a layer of orientations; clustering the layer by reachability (in the sense of the transitive closure of the closeness criterion being adopted for orientations); medoid-based reclustering of the layer; and splitting weakly localized clusters. The mentioned methods operate with the concept of closeness between orientations. To formalize it, the pseudometric distance that is invariant to the lattice symmetry transformations of one argument is introduced on the basis of the natural Riemann metrics. The developed clustering approach can be used to reduce the dimension of orientation distribution data in statistical crystal plasticity models. Its application is considered in the examples of the numerically simulated simple shear and one-axial tension textures. The results demonstrate the satisfactory adequacy of the reduced orientation distribution measures corresponding to the medoid-based representations of the textures.

Acknowledgments
The work was supported by the Russian Federation Ministry of Science and Higher Education (the basic part of the PNRPU state assignment, No. 9.7434.2017/8.9), Russian Foundation for Basic Research (grant No. 17-01-00379-a) and the RF President Grant No. MK-2970.2019.1.

References
[1] Chateigner D 2013 Combined Analysis (New York: John Wiley & Sons) chapter 4 pp 95–190
[2] Holmes P, Lumley J L, Berkooz G and Rowley C W 2012 Turbulence, Coherent Structures, Dynamical Systems and Symmetry (Cambridge: Cambridge University Press) chapter 3 pp 68–105
[3] Kumar A and Dawson P R 1996 Comput. Methods Appl. Mech. Eng. 130(3–4) 227–46
[4] Adams B L, Henrie A, Henrie B, Lyon M, Kalidindi S R and Garmestani 2001 J. Mech. Phys. Solids 49(8) 1639–63
[5] Kalidindi S R, Houskamp J R, Lyons M and Adams B L 2004 Int. J. Plast. 20(8–9) 1561–75
[6] Kalidindi S R, Houskamp J R, Proust G and Duvvuru H 2005 Mater. Sci. Forum 495–497 23–30
[7] Brahme A, Winning M and Raabe D 2009 Comput. Mater. Sci. 46 800–4
[8] Sundararaghavan V and Zabaras N 2005 Acta Mater. 53(4) 1015–27
[9] Sundararaghavan V and Zabaras N 2009 Stat. Anal. Data Min. 1(5) 306–21
[10] Acharjee S and Zabaras N 2003 Acta Mater. 51(18) 5627–46
[11] Ganapathysubramanian S and Zabaras N 2004 Comput. Methods Appl. Mech. Eng. 193(45–47) 5017–34
[12] Ostapovich K V and Trusov P V Proc. 12th Int. Conf. on Mechanics, Resource and Diagnostics of Materials and Structures (Ekaterinburg, RF, 21–25 May 2018) pp 271–72.
[13] Mokrova S M, Petrov R P and Milich V N 2016 Vestn. Udmurt. Univ. Mat. Mekh. Komp. Nauk 26(3) 336–44
[14] Hastie T, Tibshirani R and Friedman J 2009 The Elements of Statistical Learning: Data Mining, Inference, and Prediction (New York: Springer-Verlag)
[15] Trusov P V, Shveykin A I and Yanz A Yu 2016 Phys. Mesomech. 19(2) 47–65