Characterization of different crystal shapes using multivariate analysis

Dr.-Ing. Kerstin Wohlgemuth, TU Dortmund University, Laboratory of Plant and Process Design

1. Tasks and Objectives

In the chemical and pharmaceutical industry, crystallization is a process frequently used for obtaining and purifying a substance in solid form with desired properties. The control of the crystal shape/habitus is often very important, as it affects both the solid-liquid separation following crystallization, and bioavailability. Moreover, active pharmaceutical ingredients are especially affected by a change in bioavailability, i.e. if different polymorphic forms are involved. As a rule, the final product at the end of the process chain (consisting of crystallization, solid-liquid separation, washing and drying) is evaluated with regard to its properties (e.g., polymorphic form, purity, habitus, or particle size distribution). This can lead to products being produced in batch processes which do not meet the specification, and have to be rejected consequently. In the case of continuous crystallization processes, the reaction can be delayed accordingly. In both cases, increased resource and energy consumption can be expected.

The aim of this project, therefore, was the extension of an existing online image analysis tool [1, 2] in such a way that it is able to differentiate between different crystal shapes. These can be different polymorphic forms or crystals that change their habit due to different impurity concentrations because of different face-specific growth rates.

Multivariate analysis is central to the image analysis tool. Here, the image descriptors initially determined for all crystals of a training set are combined in such a way that a classification function – the so-called classifier – is able to differentiate crystals with different properties. So far, multivariate analysis has been successfully used to distinguish single crystals from crystal agglomerates, and to divide the classical particle size distribution into two subpopulations (single crystals, agglomerates). At present, a separate classifier has to be created for each material system. In order to create these classifiers quickly and reliably, we have developed a method in cooperation with Technical University of Munich which allows a systematic selection of the image descriptors with the help of a selection criterion called proportional similarity [3].

In order to distinguish between different crystal shapes, however, a different type of classifier is required. The requirements are many times higher as the different sizes will play a decisive role. The design of a shape-independent classifier, which is applicable for different material systems, is the aim of this project. The results of the project are published in [4] already, and are summarized here in a highly abbreviated form.

2. Solution Approach

As it turned out in [3], the use of an artificial neural network (ANN) for classification is the most sensible way, since non-linear relationships of image descriptors can be represented with little effort. The general idea of this project is as follows: In a first step, single crystals are discriminated from agglomerates, followed by habit discrimination of the single crystals previously determined to obtain the desired information on crystal shapes present. For better understanding of this so-called ANN cascade, see Figure 1. As it turned out, particular attention should be paid to the creation of the training set for the ANN: For the classifier to be applicable to the majority of crystal shapes possible, the classifier was trained on three different material systems that cover a wide range of possible crystal shapes; namely dipyramidal L-alanine, hexagonal platelet adipic acid, and needle-like β L-glutamic acid,

all crystallized from aqueous solution. To create the two classifiers necessary, proportional similarity is used for the pre-selection of image descriptors.

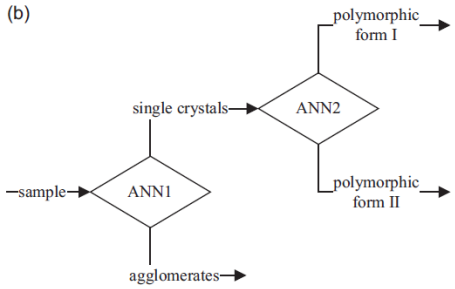

Figure 1: Simplified representation of an ANN cascade for shape discrimination

The results are presented in the following for the model system L-glutamic acid/water, exemplary. During the crystallization process and depending on the choice of process conditions, the two different polymorphs of L-glutamic acid (α (rhombic platelets) and β (needles)) crystallize in different habits (Figure 2) and agglomerate also.

Figure 2: Polymorphic α form (left) and β form (right) of L-glutamic acid

3. Results

Figure 3 shows the comparison of the final shape-specific (solid symbols, non-transferable benchmark classifiers according to [1-3]) and shape-independent (empty symbols, ANN1) classifier results for β L-glutamic acid exemplarily. Please refer to [4] for the exact compilation of the classifiers used here. With the performance measurements PI_j and δAg_j showing slight deviations from the accuracy required (PI_j>0.9 and δAg_j<0.1) only (Figure 3 left), it is assumed that both the shape-specific and the shape-independent classifier give similar results. This assumption is confirmed in Figure 3 right, where the subpopulations of single crystals and agglomerates for both classifiers superimpose completely.

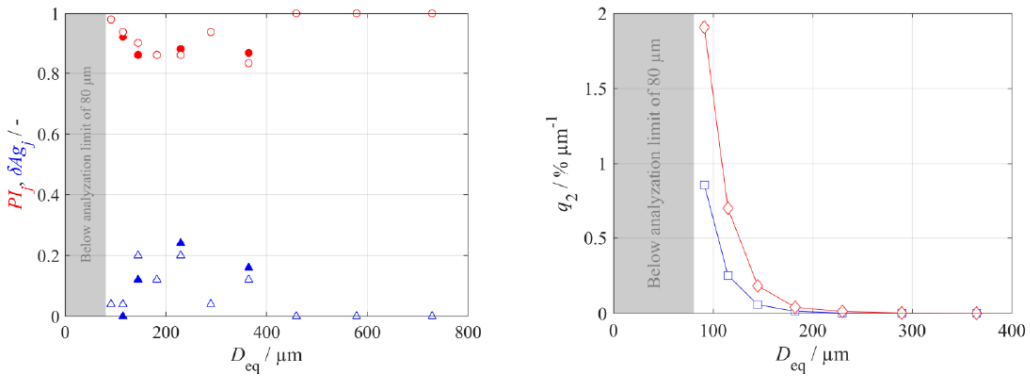

Figure 3: Comparison of the final shape-specific (solid symbols) and shape-independent (empty symbols) classifier results for β L-glutamic acid. Left: PI_j (red circles) and δAg_j (blue triangles) for the two classifiers. Right: Subpopulations of single crystals (blue squares) and agglomerates (red diamonds). The total number of crystals analyzed was 12,027. The two curves of the shape-independent classifier are visible only because they superimpose the curves of the shape-specific classifier completely.

The highly accurate shape-independent classifier (ANN1) was now used within the cascade (compare Figure 1) to discriminate a sample containing L-glutamic acid crystals of both polymorphic forms which are also available in agglomerated form. The subpopulations of single crystals and agglomerates of the sample created with ANN1 are given in Figure 4 left. It is shown that the sample is highly agglomerated. In the second step, the single crystal subpopulation is analyzed further with ANN2 to discriminate the single crystals into the α and β form (Figure 4 right).

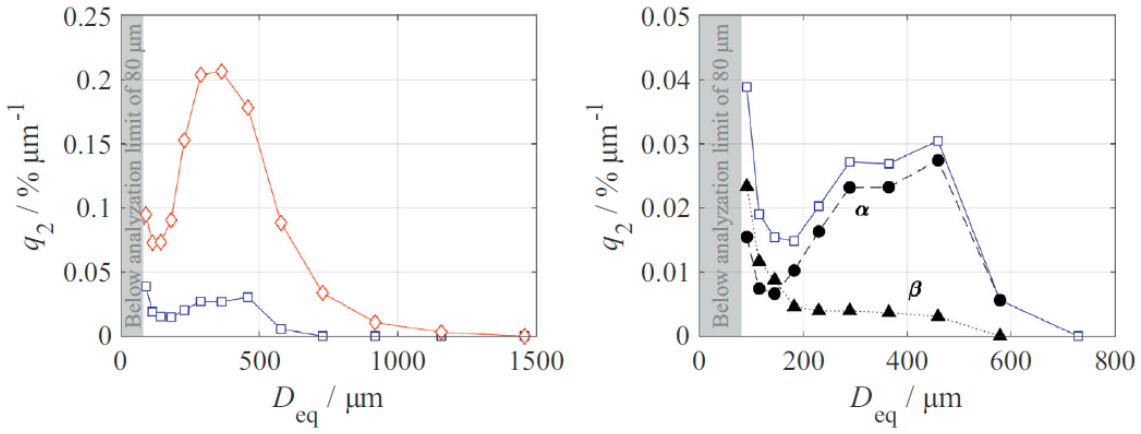

Figure 4: In-depth analysis of an exemplary sample containing both α and β L-glutamic acid. Left: PSD subpopulations of single crystals (blue squares) and agglomerates (red diamonds) created with ANN1. Right: Discrimination of single crystals into α (black circles) and β (black triangles) form created with ANN2. The overall (area-based) polymorphic amount is 79.28% of the α form and the total number of crystals analyzed was 5,641.

4. Conclusion

The use of an artificial neural network cascade not only for discrimination of single crystals/agglomerates but different crystal forms was shown successfully. The procedure presented showed high potential for gaining more in-depth insight into occurring agglomeration processes and makes it possible to determine polymorphic phase ratios within a sample also. This might be of interest for changing process parameters within crystallization processes and their effect on the product quality.

5. References

[1] L.-M. Terdenge, S. Heisel, G. Schembecker, K. Wohlgemuth
Agglomeration degree distribution as quality criterion to evaluate crystalline products
Chemical Engineering Science 133 (2015) 157-169

[2] L.-M. Terdenge, K. Wohlgemuth
Impact of agglomeration on crystalline product quality within the crystallization process chain
Crystal Research & Technology 51 (2016) 513-523

[3] S. Heisel; T. Kovačević; H. Briesen, G. Schembecker, K. Wohlgemuth
Variable Selection and Training Set Design for Particle Classification using a Linear and a Non-Linear Classifier
Chemical Engineering Science 173 (2017) 131-144

[4] S. Heisel, J. Ernst, A. Emshoff, G. Schembecker, K. Wohlgemuth
Shape-independent particle classification for discrimination of single crystals and agglomerates
Powder Technology 345 (2019) 425–437