

A new method for protein characterization and classification using geometrical features for 3D face analysis: an example of tubulin structures

Luca di Grazia¹, Maral Aminpour^{2,2}, Enrico Vezzetti¹, Vahid Rezania³, Federica Marcolin¹, and Jack Tuszynski²

¹Politecnico di Torino

²University of Alberta

³MacEwan University City Centre Campus

April 28, 2020

Abstract

This paper reports on the results of research aimed to translate biometric 3D face recognition concepts and algorithms into the field of protein biophysics in order to precisely and rapidly classify morphological features of protein surfaces. Both human faces and protein surfaces are free-forms and some descriptors used in differential geometry can be used to describe them applying the principles of feature extraction developed for computer vision and pattern recognition. The first part of this study focused on building the protein dataset using a simulation tool and performing feature extraction using novel geometrical descriptors. The second part tested the method on two examples, first involved a classification of tubulin isotypes and the second compared tubulin with the FtSZ protein, which is its bacterial analogue. An additional test involved several unrelated proteins. Different classification methodologies have been used: a classic approach with a Support Vector Machine (SVM) classifier and an unsupervised learning with a k-means approach. The best result was obtained with SVM and the radial basis function (RBF) kernel. The results are significant and competitive with the state-of-the-art protein classification methods. This opens a new area for protein structure analysis.

Hosted file

paper_proteins_MA.pdf available at <https://authorea.com/users/311604/articles/442305-a-new-method-for-protein-characterization-and-classification-using-geometrical-features-for-3d-face-analysis-an-example-of-tubulin-structures>