Literature review

Max-tree

A tree created by finding the connected components of the image when thresholding the image at every intensity level. The tree creates a hierarchical representation of the image in which every node corresponds to a peak component. The root of the tree represents the entire image and the leaves represents the local maxima of the image. This tree can also be used to efficiently find interesting attributes of the nodes [4].

Background estimation

Before astronomical objects can be detected, the background has to be subtracted from the images, such that after subtraction each pixel's intensity is proportional to the amount of incoming light [4].

LVQ/GMLVQ

A supervised classification algorithm which is represented by the prototypes $W = \{w_1, w_2, .., w_n\}$ The algorithm is trained as follows: for each input vector L, find the closest prototype W_{min} according to a the euclidean distance, and then update (move) W_{min} to be closer if the label of L and W_{min} match, and else farther apart.

Generalized Relevance Matrix (LVQ) is an extension of LVQ. In GMLVQ, the distance is computed as: [3]

$d(w,\ a) = (a - w)T(a - w)$ with $T = N^T N$ and and $\sum_i T_{ii} = 1$

Using the matrix T allows to (de-)emphasize certain features and gives the ability to tailor the distance function according to the given problem.

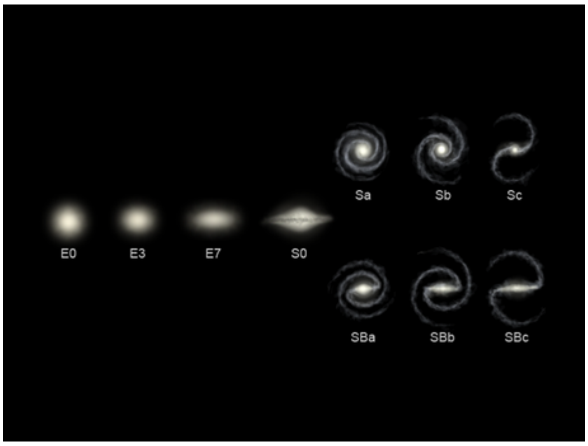
Astronomical data sets (catalogues)

SDSS data-release 7 (DR7)
GAMA (Galaxy and Mass Assembly)

Galaxy classes [Kelvin et al, 2013]

Kelvin et. al devised a method where galaxies are classified into the hubble classes from E to Sd-Irr. [5]

1. Ellipticals
2. Little blue spheroids
3. Early-type spirals
4. Early-type barred spirals
5. Intermediate-type spirals
6. Intermediate-type barred spirals
7. Late-type spirals & irregulars
8. Stars



In addition, it is also possible to detect merging galaxies and overlapping galaxies. [4]

Image moments

an image moment is a certain particular weighted average of an image pixels' intensities. [1] The raw moment of order $p + q$is defined as $M_{ij} = \sum_x \sum_y x^i y^j I(p, q)$where $I(p, q)$is the pixel intensity at the point (p, q).

Moment invariants

From the raw moments it is possible to construct moment invariants [1] which are invariant under translation, scale, and rotation. Because of the invariance, they are very useful for classifying shapes. Flusser later improved on these moment invariants, showing that they are incomplete and not independent and derived a complete set of invariants [2].

The idea for this project is to use the mtobjects library, an implementation of the max-tree algorithm, to compute an entirely new feature vector. Since each object is represented by a node in the max-tree, we can compute an attribute for each node in the subtree of an object. In each node we will compute the Flusser moment invariants, and then form a feature vector composed of the changes of these invariants in these subtrees. In other words, the feature vector will indicate roughly how an object's shape changes with its intensity level.

References

1. Hu, M.K.: Visual pattern recognition by moment invariants. IRE Transactions on Information Theory 8(2), 179–187 (1962)
2. Flusser, J.: On the independence of rotation moment invariants. Pattern Recognition (PR) 33(9), 1405–1410 (2000)
3. Nolte, A., Wang, L., Bilicki, M., Holwerda, B., Biehl, M.: Galaxy classification: A machine learning analysis of GAMA catalogue data, Elsevier (2019)
4. Teeninga, P., Moschini, U., Trager, S.C., Wilkinson, M.H.F: Statistical attribute filtering to detect faint extended astronomical sources. Math. Morphol. Theory Appl. 2016; 1:100–115 (2016)
5. Kelvin, L.S. et. al: Galaxy And Mass Assembly (GAMA): ugrizY JHK Sérsic luminosity functions and the cosmic spectral energy distribution by Hubble type. Mon. Not. R. Astron. Soc. 000, 1–23 (2013)