Shape Index Descriptors Applied to Texture-Based Galaxy Analysis

Kim Steenstrup Pedersen[†], Kristoffer Stensbo-Smidt[†], Andrew Zirm[‡], Christian Igel[†]
† Department of Computer Science, and ‡ Dark Cosmology Centre, Niels Bohr Institute
University of Copenhagen, Denmark
http://image.diku.dk/MLLab/SkyML.php

Abstract

A texture descriptor based on the shape index and the accompanying curvedness measure is proposed, and it is evaluated for the automated analysis of astronomical image data. A representative sample of images of low-redshift galaxies from the Sloan Digital Sky Survey (SDSS) serves as a testbed. The goal of applying texture descriptors to these data is to extract novel information about galaxies; information which is often lost in more traditional analysis. In this study, we build a regression model for predicting a spectroscopic quantity, the specific star-formation rate (sSFR). As texture features we consider multi-scale gradient orientation histograms as well as multi-scale shape index histograms, which lead to a new descriptor. Our results show that we can successfully predict spectroscopic quantities from the texture in optical multi-band images. We successfully recover the observed bi-modal distribution of galaxies into quiescent and star-forming. The state-of-the-art for predicting the sSFR is a color-based physical model. We significantly improve its accuracy by augmenting the model with texture information. This study is the first step towards enabling the quantification of physical galaxy properties from imaging data alone.

1. Introduction

This paper investigates a novel combination of texture descriptors and applies them for automated analysis of galaxy images. We follow the line of filter-based approaches [25, 31, 32] to texture analysis. Specifically, we focus on derivative filters. We construct differential invariants from these filters and agglomerate this information in histogram representations [23]. Descriptors such as SIFT, HoG, and DAISY [26, 12, 30] capture the local structure in images using first order differential structure in the form of gradient orientation histograms. We propose to extend these descriptors by a representation of the second order differential structure. To this end, we suggest using the shape index and the accompanying curvedness mea-

sure [21] as the basis for our descriptor, since they provide a summary of the second order structure. The novelty of our approach lies in using localized shape index histograms combined with gradient orientation histrograms both measured at multiple scales. For texture analysis, adding this higher order information will in some applications be necessary in order to improve the discriminative performance of texture representations—and quantifying physical properties of galaxies from imaging data is such an application.

Galactic structure (i.e. how the mass is generally distributed within galaxies) and morphology (i.e. how that mass is arranged on smaller scales) are important diagnostics of the formation and evolutionary mechanisms and timescales for galaxies. It is well known that this structure is correlated with other physical properties of the galaxies such as star-formation rate and dust content (e.g. [7]). However, the means to formalize these relationships are yet to be realized. Extremely large galaxy surveys from the ground, such as the SDSS, have compiled vast, homogeneous imaging of millions of galaxies. Furthermore, ever since the launch of the Hubble Space Telescope (HST) and the advent of adaptive-optics (AO) on large aperture ground-based telescopes enabling high physical-resolution images of galaxies, the study of galaxy structure and morphology has entered a data-rich era.

Galaxies are made of stars, gas and dust. Each of these components emits light over different wavelength ranges and with different intensities. To use the observed light, for example, to determine the mass of stars or the rate at which new stars are being formed, we need to be able to disentangle the various luminous contributions. To do so, astronomers build models of the emission for each source. Gas will primarily emit in emission lines, which appear at a set of discrete wavelengths associated with the emitting element. These emission lines can only be observed spectroscopically and give the most direct measurement of the rate at which new stars are being formed (SFR). Stars, on the other hand, emit continuum radiation over a large range of wavelengths. We can use models of populations of stars as a function of time to extract the mass and age of the stars

in a galaxy. These models can be used for spectroscopy as well as (broad band) imaging in multiple filters (colors).

The mass and SFR of a galaxy can therefore be (coarsely) measured by comparing a set of models with the shape of the spectral energy distribution traced by multiple filters. The specific star formation rate (sSFR) is simply the current SFR divided by the mass of stars. Usually, even if the SFR is determined from emission lines spectroscopically, the mass is determined from the colors of the galaxy in multi-filter imaging. The dominant approach for estimating sSFR from imaging data alone is based on analysis of the color of the galaxy.

Our current knowledge of galaxies is built on imaging surveys and follow-up spectroscopy. Modern imaging surveys will acquire data in several band-pass filters and can be used to approximate galactic properties. However, better determinations of these quantities require deep spectroscopy covering a significant wavelength baseline. Furthermore, most surveys will only have a single band of high angular-resolution imaging (e.g. from space). In such resolved galaxy images, it is possible to use the structure as a proxy for internal dynamics that would require more time-consuming spectroscopic data to observe. Indeed, many of the future surveys will be imaging-only surveys that will not allow for spectroscopic follow-up observations of the vast majority of the observed galaxies. Therefore, being able to fully exploit the most well-resolved images as proxies for spectroscopic data is highly valuable.

Figure 1 illustrates examples of optical images of galaxy from the subset of the SDSS dataset that is used in this paper. The top row shows well-resolved galaxy images. Notice that the light profile of these galaxies contains intricate texture. This texture is caused by the distribution of stars and gas in the galaxy—an important cue for determining the sSFR. We propose to investigate the predictive power of texture when estimating sSFR from optical images. The bottom row of Fig. 1 illustrates problematic cases for our texture based analysis. These range from noise and nearby stars to faint distant galaxies which are poorly resolved in the images. At first glance, this may seem impossible. After all, making the leap from single-band or a few bands imaging data to spectroscopic quantities is a large jump. However, the properties of galaxies are correlated. We have known since the earliest galaxy surveys, that star-forming galaxies have more internal morphological structure due to dust obscuration and star-forming clumps than quiescent (elliptical) galaxies, which tend to be smoother.

There has been some prior work on automated analysis of optical images of galaxies [14, 9]. Much of this work, however, focuses on classification of galaxies based on morphology (e.g. [4]) and tends to ignore information found in the texture. Furthermore, these approaches have used somewhat standard image features as input. Here we present new
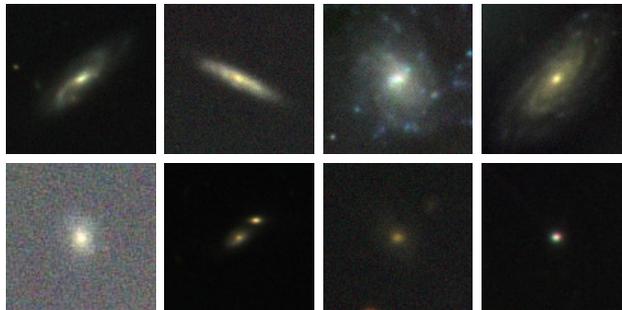

Figure 1. Examples of low-redshift galaxies in our subset of the SDSS dataset. We have mapped the gri-bands to the RGB color space ($gri \rightarrow$ BGR). The top row shows well-resolved galaxies and the bottom row shows problematic cases for our analysis. These color images are best viewed electronically.

image features which we believe can capture heretofore ignored information contained in resolved galaxy images.

The following section (§ 2) describes the galactic dataset we use in our experiments. The new texture features are introduced in § 3. Section § 4 explains how we perform regression in order to predict sSFR values from our features. The results are presented in § 5 before we discuss their implications and future extensions of our work in § 6.

2. Galaxy data

The primary data used for the current work are a sample of low-redshift galaxies drawn from the SDSS DR7, see Fig. 1. We use the g-, r-, and i-band images covering the wavelengths from 4000–5500, 5500–7000 and 7000–8500 Ångstroms, respectively. This sample is defined as all spectroscopic galaxies within the GAMA DR1 region [13] which also have entries in both the MPA-JHU and NYU-VAGC catalogs [11, 8]. The overlap with GAMA for these ~ 12000 galaxies is of particular interest because that survey will acquire spectroscopy of fainter targets and higher quality imaging (including at different wavelengths) thus allowing us, eventually, to extend our analysis to more galaxies and to longer wavelengths.

The images for our galaxy sample were obtained using the *skyview* software provided by NASA/GSFC. For each galaxy position, as defined in the SDSS DR 7, we downloaded a 100×100 pixel region (covering $39.6'' \times 39.6''$) around that position. These images are not background subtracted and do not include an object segmentation map. We used SExtractor [6] on each image to generate and subtract an estimate of the background and to produce a segmentation map including both the target and neighboring galaxies. We have not applied any additional smoothing to the galaxy pixels at this stage because that is a core part of our following analysis. We however compress the intensity range by applying a logarithmic function of the intensities.

The last step in the pre-processing of the images was to construct a refined and well-defined pixel segmentation mask indicating which pixels belonged to the galaxy of interest in each frame. We used a generalized Petrosian method to build these masks, similar to that presented in [3]. We first rank-order the pixels in the SExtractor segmentation map for the target from bright to faint. At each intensity level we calculate the average intensity brighter than that pixel. When the ratio of the pixel's intensity to that average reach a pre-determined value (the Petrosian η) we set that intensity as the lower limit for a pixel to be included in the following analysis. For some galaxies, even with low η, the resulting number of included pixels may be too small for proper analysis (see below for further details). We note that smoothing the data first and then creating the mask will push more pixels above η and create a more inclusive mask. However, these lower significance pixels will not add to the textural features at small scales because they will be highly correlated in a way determined by the smoothing kernel.

Each band image leads to slightly different masks, not only due to noise but also because some galaxy structure is only visible at certain wavelengths. We construct a combined mask by taking the union of the masks for each band. We use this combined mask for processing all of the bands.

The mask extraction (segmentation) occasionally leads to incorrect masks which includes non-galaxy pixels. In order to remove some of these outliers from the analysis, we apply a threshold on the ratio of galaxy pixels and pixels in the convex hull of the galaxy mask. We discard all images where this ratio is less than 0.7.

The galaxy images were extracted such that each galaxy is in the image center. We discard images from the analysis if the mask processing leads to a mask not overlapping with the image center. This may be caused by a faulty mask extraction that latches onto objects in the vicinity such as nearby stars.

In order to remove noise at the boundary of the produced masks and holes inside these, the masks where processed by applying a morphological closing followed by an opening operation with a disk structure element with radius 1 pixel. Following this the masks have been filtered with a linear Gaussian filter with $\sigma = 0.5$ and filter mask size equal to 3σ. This produces a cleaned galaxy mask with smooth boundaries.

Prior to applying the Gaussian filter, we estimate the Petrosian radius of the galaxy by

$$ R_p = \sqrt{\frac{N_{\mathrm{gal}}}{\pi}} \, , \qquad (1) $$

where N_{gal} denotes the number of galaxy pixels in the mask. Furthermore, we estimate a fiducial orientation of the galaxy from the binary mask, which we use to make the gradient orientation feature invariant to rotation. This esti-

mation is based on the masks prior to Gaussian filtering. We compute the spatial covariance of the galaxy pixels by

$$ \mathbf{C}_{\mathrm{gal}} = \frac{1}{N_{\mathrm{gal}} - 1} \sum_{x_{\mathrm{gal}}} (x_{\mathrm{gal}} - \mu)^T (x_{\mathrm{gal}} - \mu) \, , \qquad (2) $$

where $x_{\mathrm{gal}} \in \mathbb{R}^2$ is the position of galaxy pixels in the mask, the sum runs over all galaxy pixels in the mask, and

$$ \mu = \frac{1}{N_{\mathrm{gal}}} \sum_{x_{\mathrm{gal}}} x_{\mathrm{gal}} \qquad (3) $$

is the mean position of all galaxy pixels. We define the fiducial orientation of the galaxy as the eigenvector corresponding to the largest eigenvalue of the covariance matrix. This direction of most spatial variance in galaxy pixels usually corresponds to the major axis of ellipsoidal shaped galaxies. Since the eigenvector is computed up to a change of sign, we flip the sign of any eigenvector with a negative x-component in order to make the orientation consistent. In case of isotropic galaxies this way of picking a fiducial orientation will lead to a random choice, but as there is no natural orientation in this case, this is acceptable.

We note here that our image analysis does not strongly depend on the precise background level (as long as it does not vary greatly on galaxy scales), the choice of η, or on the absolute flux level in the galaxy pixels themselves. Our image features are dependent solely on the intensity texture within the galaxies—not the specific intensity level. That said, objects for which the number of pixels in the mask is smaller than ~ 100 will have insufficient data to reliably measure histogram based image features. We do, however, not remove such images from our study, which potentially leads to outliers in the analysis.

3. Texture descriptors

Discriminative information in textures may appear on several different scales—this is certainly the case for galaxy images—hence using a multi-scale representation appears to be a necessity when performing analysis of texture images. We use the linear scale-space representation [20, 29], where the scale-space of an image $I : \Omega \mapsto \mathbb{R}$, $\Omega \subset \mathbb{R}^2$ is defined as $L(x, y; \sigma) = (I * G)(x, y; \sigma)$, where $*$ denotes convolution with a Gaussian filter

$$ G(x, y; \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \, . \qquad (4) $$

The parameter $\sigma > 0$ is the scale of the representation. In this representation we can compute image derivatives of order n and m by

$$ L_{x^n y^m}(x, y; \sigma) = \left(I * \frac{\partial^{(n+m)} G}{\partial x^n \partial y^m}\right)(x, y; \sigma) \, . \qquad (5) $$

Image derivatives form the basic components of our descriptors, but we will introduce non-linearity in the features by applying functions of these derivatives.

Common descriptors such as SIFT, HoG and DAISY [26, 12, 30] use first order differential structure in the form of gradient orientation histograms as the basis of the descriptor. In smooth scale space derivatives the gradient orientation may be defined as

$$\theta(x, y; \sigma) = \tan^{-1}\left(\frac{L_y(x, y; \sigma)}{L_x(x, y; \sigma)}\right), \qquad (6)$$

and the scale normalized gradient magnitude as

$$M(x, y; \sigma) = \sigma^2\sqrt{L_x^2(x, y; \sigma) + L_y^2(x, y; \sigma)} \ . \qquad (7)$$

We need to perform this scale normalisation in order to be able to compare M across different scales σ [29].

We also add a representation of the second order differential structure—namely the shape index and the accompanying curvedness measure [21]. The shape index is based on the eigenvalues κ_1 and κ_2 of the Hessian matrix of the image function. It is defined as the angle between the vector of the eigenvalues (κ_1, κ_2) and the first axis in this eigenvalue space. In terms of image derivatives we may express the shape index as

$$S(x, y; \sigma) = \frac{2}{\pi}\tan^{-1}\left(\frac{-L_{xx} - L_{yy}}{\sqrt{4L_{xy}^2 + (L_{xx} - L_{yy})^2}}\right). \qquad (8)$$

The shape index represents the basic second order shapes with dark blobs ($S = -1$), over saddle points ($S = 0$), to bright blobs ($S = 1$), with valley- and ridge-like structure in between. For the detailed geometric interpretation see [21].

The curvedness is simply defined as the length of the eigenvalue vector (κ_1, κ_2) and expresses how pronounced the second order structure is, similar to the role of the gradient magnitude for the first order structure. In terms of image derivatives the scale normalized curvedness may be defined as

$$C(x, y; \sigma) = \frac{1}{2}\sigma^2\sqrt{L_{xx}^2 + 2L_{xy}^2 + L_{yy}^2}\,. \qquad (9)$$

The shape index is rotational invariant by design, contrary to gradient orientation which depends on the choice of coordinate system.

The exact spatial ordering of the texture is not necessarily important, hence it is common (e.g. [26, 12, 30]) to introduce an agglomeration step such as statistical moments or histograms. Here we choose to use smooth histograms inspired by the concept of locally orderless images [22]. This formulation makes the intrinsic parameters of the histogram representation explicit and provides a more robust estimate compared to the traditional histogram formulation.

We define a smooth histogram as a function of the feature f in question and its magnitude F,

$$H(f_i) = \int F(x, y)A(x, y)B(f_i, x, y; f)\,\mathrm{d}x\mathrm{d}y \ , \qquad (10)$$

where f_i denotes the histogram binning variable and will act as the bin center for a specific choice of binning aperture function B. The function A localizes the descriptor to specific parts of the image. We propose to use the Gaussian function of β bin width as smooth bin aperture function for histograms of the shape index $S(x, y; \sigma)$

$$B_{\beta,\sigma}(S_i, x, y; S) = \exp\left(-\frac{(S(x, y; \sigma) - S_i)^2}{2\beta^2}\right). \qquad (11)$$

The Gaussian bin aperture is not a good choice for gradient orientation histograms, since it does not incorporate the fact that θ is periodic. A better choice is to use the von Mises density function as aperture function, since this is the extension of the Gaussian distribution to the unit circle. We therefore propose to use the following smooth bin aperture function for the gradient orientation $\theta(x, y; \sigma)$

$$B_{\beta,\sigma}(\theta_i, x, y; \theta) = \exp\left(\frac{1}{\beta}\cos\left(\theta(x, y; \sigma) - \theta_i - \theta_0\right)\right), \qquad (12)$$

where θ_0 denotes a fiducial orientation.

As feature magnitude F for shape index we will use the curvedness measure C from (9) and for the gradient orientation we will use the gradient magnitude M from (7). The rationale is that we would like local structure with a large magnitude to count more in the histogram. This also has the effect of reducing noise in the histograms caused by noise in the derivative measurements.

We propose to construct texture features by combining histograms of gradient orientation with histograms of shape index and to measure these histograms at different scales σ. As a concrete discretization of this representation we choose an equidistant binning in the histograms and fix the number of bins to 8 for gradient orientation and to 9 for shape index histogram features. The bin width β is chosen such that with the specific choice of number of bins, we tile and cover the complete range of the feature. Equation (10) weights each data point that is added to the histogram by its feature magnitude and each bin window, thus each point casts a vote in every bin of the histogram.

For our specific application to galaxy images we set θ_0 in the gradient orientation feature to be the fiducial orientation of the galaxy as defined in § 2. Furthermore, we choose the window function A in (10) to be identical to the galaxy mask as outlined in § 2. This localizes the feature to include features from only galaxy pixels. In addition, a histogram at a specific scale σ is always normalized so that the bin counts H sum to one.

Notice that our gradient orientation histogram is similar to SIFT-like descriptors, except that we do not include a spatial pooling step (i.e. we only employ a single histogram for the region of interest).

Choosing measurement scales. Using the scale space representation we can compute features at a range of scales capturing pixel correlations across these scales. Selected scales should cover the range of characteristic scales for the particular galaxy image. The inner scale is given by the pixel scale, but since we want to compute derivatives up to second order we need to be careful with the numerics. By choosing the smallest inner scale to be $\sigma_i = 0.88$ pixels we will have less than 1% numerical error in the estimation of the second order derivatives [29]. This inner scale will measure geometry at near pixel level corresponding to $0.396''$.

We approximate the effective outer scale for a particular galaxy image with the Petrosian radius (1). For isotropic galaxies this will be a good estimate, however, for elongated ellipsoidal galaxies this will be a poor over-estimate. We have opted for the simple heuristics of picking the effective outer scale as a function of the Petrosian radius. Let w be the smallest of the image width and height measured in pixels. We then use the Petrosian radius as outer scale $\sigma_o = R_p$ if $4R_p/w \leq 1$, and otherwise choose σ_o such that $4\sigma_o/w = 1$. In order to avoid artifacts in the computed scale space derivatives introduced by boundary effects, it is common to discard pixels that are close to the boundary. The heuristic ensures at least a one σ_o distance from the galaxy to the image boundary. This definition of the outer scale will measure the geometry at galaxy scale. If $\sigma_i > \sigma_o$, we discard the image from the analysis.

We sample the range of effective scales $[\sigma_i; \alpha\sigma_o]$ in exponentially growing steps. We found experimentally that $\alpha = 0.2$ is a good value for the fraction of the outer scale, which focuses the descriptor on the range of scales where relevant structure occurs in galaxy images. We note that this specific choice is application dependent. We choose to use 8 scale levels in the interest of minimizing the computational effort and at the same time achieving good results.

4. SSFR Prediction Experiments

We use regression to predict specific star formation rate (sSFR) from combinations of the texture descriptors outlined above.

Evaluation. We consider different models and feature combinations to predict the sSFR value for each galaxy image. We perform 10-fold cross validation (CV) on our subset of the SDSS dataset. As measure of the prediction error we report the root mean square error (RMSE) averaged over the 10 CV folds. We also report the standard deviation of the RMSE computed from the RMSE on each fold (when interpreting these values it has to be kept in mind that the CV folds are strictly speaking not fully independent).

Models. Because scatter plots indicated a near linear relation between our features and the sSFR, we consider a standard linear least squares regressor as predictor (*Linear*). To further improve the performance, we employ non-linear regression techniques using the Shark machine learning library [18]. We initially considered Gaussian process regression with radial Gaussian kernels, where the bandwidth parameter of the kernel and the precision of the noise were adapted by grid-search as well as gradient-based optimization of the logarithmic marginal likelihood function (or evidence) [27]. However, because the Gaussian processes did not significantly improve over the linear regression, we apply multi-layer perceptron neural networks (*MLP*). Each MLP has a single hidden layer with 100 units with logistic activation functions and a linear output unit. We add short-cut connections linking the inputs directly to the output unit. The training data of each CV fold was further split into an MLP-training and an MLP-validation set using a 9:1 split ratio. The network was trained starting from small weights by minimizing the squared error on the MLP-training set using the iRProp$^+$ first-order optimization algorithm [19]. The weight configuration with the smallest squared error on the MLP-validation set was considered to be the final hypothesis. This "early stopping" of a training process that increases the complexity starting from an (almost) linear model typically fosters good generalizing hypothesis (note that the actual number of hidden units is of lesser importance if chosen large enough, see [5]). In the following, we only report the linear regressor and MLP results.

As a baseline, we use the constant model predicting the sSFR value to be the average sSFR value of the training set (later referred to as *Average*).

We also include a color-based model of the sSFR which was provided together with the SDSS dataset (*Color*). The method is based on the approach described in [16, 15, 28, 10], which employs a physical model of the relations between sSFR and spectrum of a galaxy.

Finally, we augment the color-based physical model by our texture features. This is done by fitting the residuals of *Color*. We refer to the resulting *additive models* [17] as either *Linear-AM* or *MLP-AM* depending on whether linear regression or our neural network approach was used.

Features. As input features, we consider gradient orientation (*GO*) and shape index (*SI*) features as well as their combination (referred to as *All*). Each feature consists of histograms at 8 scale levels.

Furthermore, for reference we include the best results achieved using a feature set consisting of histograms of filter responses for second order directional derivatives and the Laplacian (*2nd*), i.e. the filters used in [31]. These features were implemented using the smooth histograms defined by (10)-(11), and computed at multiple scales using the same choices as for our features.
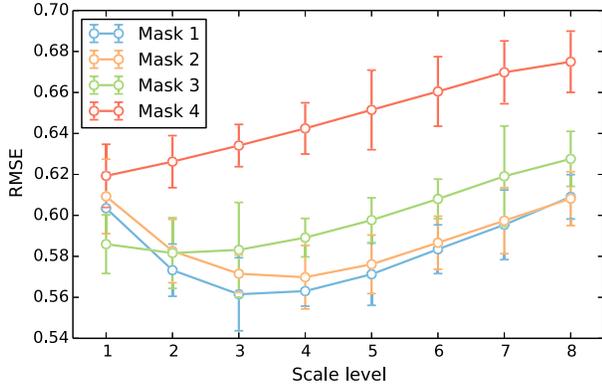
Figure 2. Plot of RMSE (error bars indicate 1 standard deviation of the CV error) of Linear *gri* (SI) across the 8 scale levels for the four masks. Notice for masks 1–2 the curve has a dip indicating that for single scale features an optimal scale exists.
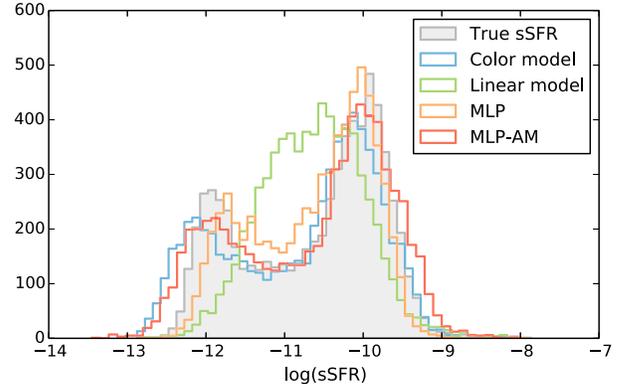

Figure 3. Plot of the distributions of predicted sSFR values for different predictors and the ground truth for mask 1, using the *gri* and shape index (SI) features. It is seen that all models but the linear recover the bimodal sSFR distribution.

We use 4 different mask sizes in decreasing size with mask 4 being the smallest. The amount of galaxy images that passes all inclusion criteria outlined in § 3 for all masks can be found in Table 1.

5. Results and Discussion

Table 1 summarizes our results for different combinations of features extracted from either a single band (g, r, and i) or all bands (gri) and different regressors. The additive models (AM) yield more accurate predictions (2 standard deviations better) than the standard color-based predictor. Thus, the texture features provide information orthogonal to the color model.

Even in single bands the texture information is correlated with the sSFR value, see the Linear and MLP (All) results. Notice that we obtain slightly better accuracies in the g-band. However, the best texture-only results are obtained on the combined gri-bands.

Using gradient orientation features alone does not provide enough information in this particular application. Instead we need to include the shape index feature or use the shape index feature alone. We only include results for the Linear *gri* predictor, but the tendency is the same for the single bands and the MLP predictor. This is consistent with similar observations made in [24], in which it is argued that increasing differential order of the features can be beneficial for discriminability. The results on the second order features *gri* (2nd) are comparable to the (all) and (SI) results for mask 1 but with an increased variance, and for masks 2–3 these features are inferior to the shape index (SI) results.

Fig. 2 show the RMSE of the linear regressor based on shape index (SI) features using single scale levels applied to the combined *gri* features. Remember that, due to our scale range selection procedure (§ 3) for each image the exact scale used at each scale level will vary as a function of the galaxy size. Notice for masks 1–2 the curve has a dip, indicating that for single scale features an optimal scale exists. However the results of Table 1 show that by combining information at several scales simultaneously we are able to obtain better predictions than with a single scale.

The reason for the generally poor results on mask 4 is that these masks tend to only include the galaxy nuclei which usually appears as a bright saturated blob of light. Our texture features does therefore not provide much information at this part of the galaxy.

Our results also indicate that a linear model actually does a good job of fitting the data, but we do get a slight improvement by introducing the non-linear MLP.

To provide some additional insight Fig. 3 show histograms of the spectroscopic sSFR values together with the results of the predictors Linear *gri* (SI), MLP *gri* (SI), and MLP-AM *gri* (SI). All predictors but the linear are able to recover the two known classes of star-forming and quiescent galaxies seen by the two modes in the histograms. Notice how the color-based predictor systematically underestimates the sSFR value (seen by the shift of the histogram to the left) and that the MLP has a tendency to push the modes towards the mean of the dataset. It is evident that the MLP does a better job at recovering the true sSFR distribution than the linear predictor. It can be nicely seen how fitting the residual (MLP-AM) corrects the Color model.

6. Conclusions

We propose to combine gradient orientation and shape index histograms measured at several scales to describe image texture. SIFT-like descriptors include a spatial pooling

Table 1. Summary of our results for different model-feature pairs applied to either single bands (g, r, and i) or all bands (gri) using four different masks (in decreasing size). The results are based on 6880 images passing the inclusion criteria. The numbers in the table indicate RMSE and cross validation standard deviation. *Average* refers to predicting the training data mean (i.e. an estimator of the data variance) and *Color* is the current state-of-the-art physical model (see § 4). *Linear* and *MLP* denote linear and non-linear regression. *Linear-AM* and *MLP-AM* are the additive models combining *Color* with *Linear* and *MLP*, respectively. Gradient orientation (*GO*) and shape index (*SI*) features as well as their combination *All* and second order features (*2nd*) are considered. For more results see the supplementary material.

Method	Band (features)		Mask 1	Mask 2	Mask 3	Mask 4
Average			0.88 ± 0.02	0.88 ± 0.01	0.88 ± 0.01	0.88 ± 0.01
Color			0.33 ± 0.01	0.33 ± 0.02	0.33 ± 0.02	0.33 ± 0.02
Linear	g	(all)	0.61 ± 0.01	0.62 ± 0.02	0.62 ± 0.01	0.65 ± 0.01
	r	(all)	0.65 ± 0.02	0.63 ± 0.02	0.63 ± 0.01	0.67 ± 0.02
	i	(all)	0.65 ± 0.02	0.64 ± 0.02	0.64 ± 0.02	0.67 ± 0.01
	gri	(all)	0.53 ± 0.02	0.54 ± 0.02	0.55 ± 0.02	0.59 ± 0.02
Linear	gri	(SI)	0.53 ± 0.02	0.54 ± 0.02	0.55 ± 0.02	0.59 ± 0.01
Linear	gri	(GO)	0.81 ± 0.02	0.83 ± 0.01	0.84 ± 0.01	0.85 ± 0.02
Linear	gri	(2nd)	0.53 ± 0.03	0.57 ± 0.05	0.68 ± 0.31	0.64 ± 0.05
MLP	g	(all)	0.55 ± 0.01	0.57 ± 0.02	0.58 ± 0.02	0.61 ± 0.01
	r	(all)	0.61 ± 0.02	0.59 ± 0.02	0.61 ± 0.02	0.63 ± 0.01
	i	(all)	0.61 ± 0.02	0.60 ± 0.02	0.61 ± 0.01	0.64 ± 0.02
	gri	(all)	0.49 ± 0.02	0.50 ± 0.01	0.52 ± 0.01	0.55 ± 0.02
MLP	gri	(SI)	0.50 ± 0.02	0.50 ± 0.01	0.52 ± 0.01	0.56 ± 0.01
Linear-AM	gri	(SI)	0.29 ± 0.02	0.29 ± 0.01	0.29 ± 0.02	0.29 ± 0.01
MLP-AM	gri	(SI)	0.29 ± 0.02	0.29 ± 0.02	0.29 ± 0.02	0.29 ± 0.02

step collecting information from a grid of histograms tiling the region of interest (ROI). This allows SIFT descriptors to some extend code spatial structure in the ROI beyond first order differential structure. Our gradient orientation feature can be thought of as a single histogram SIFT descriptor. Contrary to general SIFT-like descriptors, we have the luxury of having a segmentation of the object of interest. Instead of applying a spatial pooling step we choose to increase the differential order.

The descriptor introduced in this paper is tuned towards the specific application, predicting the specific star-formation rate (sSFR) from galaxy images, by confining the descriptor to only include information from the galaxy pixels mask. Based on the mask we fix the outer scale used in the scale-space as well as the dominating orientation used in the gradient orientation histogram. However, the descriptor can easily be reconfigured to be constrained to a local image patch and even be extended to a collection of histograms extracted from a spatial pooling scheme such as used in descriptors such as SIFT, HoG and DAISY [26, 12, 30]. The dominating orientation may be estimated following the same approach as in SIFT. Fixing the scale range is application dependent and requires an analysis of the concrete problem under consideration.

The power of the new descriptor is demonstrated in the application of predicting sSFR from imaging data. We obtain good results when using the texture features alone. By combining the color-based physical model with texture information, we outperform the state-of-the-art for sSFR prediction.

The success of the shape index feature can be explained by realizing that what distinguishes a quiescent galaxy from a star-forming one is the distribution of stars, gas, and dust. This leads to the presence or absence of blob-like structures, as well as the occurrance of ridge-like structures caused by spiral arms and stripe patterns formed by the distribution of gas and dust—the shape index is tuned to this type of second order structure.

A current limitation of the approach is that we extract features independently from each band image ignoring the natural correlation across bands. A future extension would be to extract color descriptors by extending the shape index descriptor to be based on the Hessian matrix of the 2D intensity manifold embedded in the spatio-color space. This strategy would also be readily applicable on other types of color image data.

One of the challenges for computer vision and machine learning in astrophysics is to take models and knowledge

gained from one training set (i.e. a particular survey) and apply it to data taken using different telescopes, instruments and techniques. For our current efforts, the primary difference will be the absence of the spectroscopic ground truth for current and future galaxy surveys. Many of the largest planned surveys are indeed imaging-only and while some spectroscopic follow-up will be done, it will be impossible to obtain complete spectroscopic coverage of the more numerous (and often fainter) galaxies being imaged. Against this background, this study is the first step towards enabling the quantification of physical galaxy properties from imaging data alone. We expect that this mapping of galaxy appearance and properties will prove extremely useful when applied to future large scale imaging-only surveys such as the Large Synoptic Survey Telescope (LSST).

Acknowledgements

References

[1] http://www.gama-survey.org.

[2] http://www.sdss.org.

[3] R. G. Abraham, P. Nair, P. J. McCarthy, et al. The Gemini Deep Deep Survey. VIII. When Did Early-Type Galaxies Form? *Astrophys. J.*, 669:184–201, 2007.

[4] M. Banerji, O. Lahav, C. J. Lintott, et al. Galaxy zoo: Reproducing galaxy mophologies via machine learning. *MNRAS*, 406:342–353, 2010.

[5] P. L. Bartlett. The sample complexity of pattern classification with neural networks: the size of the weights is more important than the size of the network. *IEEE T Inform Theory*, 44(2):525–536, 1998.

[6] E. Bertin and S. Arnouts. SExtractor: Software for source extraction. *Astron. Astrophys. Supp.*, 117:393–404, 1996.

[7] M. R. Blanton and J. Moustakas. Physical Properties and Environments of Nearby Galaxies. *Ann. Rev. Astron. Astrophys.*, 47:159, 2009.

[8] M. R. Blanton, D. J. Schlegel, M. A. Strauss, et al. New York University Value-Added Galaxy Catalog: A Galaxy Catalog Based on New Public Surveys. *Astron. J.*, 129:2562, 2005.

[9] S. Bridle, S. T. Balan, M. Bethge, et al. Results of the GREAT08 challenge: An image analysis competition for cosmological lensing. *MNRAS*, 405(3):2044–2061, 2010.

[10] J. Brinchmann. Private communication.

[11] J. Brinchmann, S. Charlot, S. D. M. White, et al. The physical properties of star-forming galaxies in the low-redshift Universe. *MNRAS*, 351:1151, 2004.

[12] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *CVPR*, volume 1, pages 886–893, 2005.

[13] S. P. Driver, D. T. Hill, L. S. Kelvin, et al. Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release. *MNRAS*, 413:971, 2011.

[14] G. Fasano, E. Vanzella, A. Dressler, et al. Morphology of galaxies in the WINGS clusters. *MNRAS*, 420(2):926–948, 2012.

[15] A. Gallazzi, J. Brinchmann, S. Charlot, et al. A census of metals and baryons in stars in the local Universe. *MNRAS*, 383:1439–1458, 2008.

[16] A. Gallazzi, S. Charlot, J. Brinchmann, et al. The ages and metallicities of galaxies in the local universe. *MNRAS*, 362:41–58, 2005.

[17] T. J. Hastie and R. J. Tibshirani. *Generalized Additive Models*. Chapman & Hall, 1990.

[18] C. Igel, T. Glasmachers, and V. Heidrich-Meisner. Shark. *JMLR*, 9:993–996, 2008.

[19] C. Igel and M. Hüsken. Empirical evaluation of the improved Rprop learning algorithm. *Neurocomputing*, 50(C):105–123, 2003.

[20] J. J. Koenderink. The structure of images. *Biol. Cybern.*, 50:363–370, 1984.

[21] J. J. Koenderink and A. J. van Doorn. Surface shape and curvature scale. *Image Vision Comput.*, 10(8):557–564, 1992.

[22] J. J. Koenderink and A. J. van Doorn. The structure of locally orderless images. *IJCV*, 31(2/3):159–168, 1999.

[23] J. J. Koenderink and A. J. van Doorn. Local structure of Gaussian texture. *IEICE Trans Inf Syst*, E86-D(7):1165–1171, 2003.

[24] A. B. L. Larsen, S. Darkner, A. L. Dahl, and K. S. Pedersen. Jet-based local image descriptors. In *ECCV 2012*, volume LNCS 7574, pages 638–650. Springer, 2012.

[25] T. Leung and J. Malik. Representing and recognizing the visual appearance of materials using three-dimensional textons. *IJCV*, 43(1):29–44, 2001.

[26] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004.

[27] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2005.

[28] S. Salim, R. M. Rich, S. Charlot, et al. UV Star Formation Rates in the Local Universe. *Astrophys. J. Suppl. Ser.*, 173:267–292, 2007.

[29] B. M. ter Haar Romeny. *Front-End Vision and Multi-Scale Image Analysis*. Kluwer Academic Publishers, 2003.

[30] E. Tola, V. Lepetit, and P. Fua. DAISY: An efficient dense descriptor applied to wide-baseline stereo. *IEEE TPAMI*, 32(5):815–830, 2009.

[31] M. Varma and A. Zisserman. A statistical approach to texture classification from single images. *IJCV*, 62(1/2):61–81, 2005.

[32] S. C. Zhu, Y. N. Wu, and D. Mumford. Filters, random fields, and maximum entropy (frame): Towards a unified theory for texture modeling. *IJCV*, 27(2):107–126, 1998.