RECOVERING INTRINSIC IMAGES USING AN ILLUMINATION INVARIANT IMAGE

M. Farenzena, A. Fusiello

Dipartimento di Informatica - Università di Verona
Strada Le Grazie, 15 - 37134 Verona, IT

ABSTRACT

In this paper a method for the extraction of shading and reflectance intrinsic images from a single uncalibrated image is presented. It is based on the classification of the image derivatives as either caused by shading or reflectance effects, using an illumination-invariant image to guide this classification. Our approach avoids the learning process – which requires ground truth intrinsic images – and obtain results comparable with the state of the art.

Index Terms— Reflectance Recovery, Intrinsic Images

1. INTRODUCTION

The image of a scene depends on many physical characteristics of the surfaces, such as illumination, orientation, depth, reflectance. All these information are confounded by the imaging process into an array of integer values, the sensed image, which reveals the physical parameters only indirectly. An *intrinsic image* [1] is an image that represents one of these physical properties, which are *intrinsic* to the surfaces in the scene. These intrinsic images are extremely useful, for they can be related to the scene far more easily than the sensed image.

The ability to decompose an image into its intrinsic components is a major step toward scene understanding, because algorithms often rely exclusively on one of the intrinsic characteristics of the scene. For example, shape-from-shading techniques require images with no changes in colour (or albedo). In the context of 3D modelling, the reflectance image is used as a texture map: being devoid of illumination effects the model can be re-illuminated without artifacts.

In this paper a new method for decomposing a single image into two intrinsic images – a *shading* image (the illumination at each point) and a *reflectance* image (the colour at each point) – is proposed.

In literature we can find relatively little work on this problem. Some *discriminative* approaches attempt to distinguish the effects of shading and reflectance. Among these, the Retinex algorithm [2] was one of the earliest. It originally described lightness perception for Mondrian images and it worked on lines in the image. Afterwards, Horn [3] proposed a method that extends this on 2D image. Retinex relies on the assumption that the changes in the reflectance of a surface lead to large derivatives, while illumination, which varies slowly, causes small derivatives. Thus, the recovery of reflectance and shading images is obtained by derivatives classification. However, the assumption may not hold in real images. Freeman and Viola [4] add a smoothness prior on the inferred shape in an image to classify such image as either entirely created by shading or all due to reflectance changes. Alternatively, Bell and Freeman [5] trained a classifier to use a set of linear features, instead of derivatives only, based on a steerable pyramid.

A second heuristic adopted to extract intrinsic images is that the shading and reflectance images can be found by filtering the logarithm of the input image [6]. This approach assumes that the shading component is concentrated in the low spatial frequency bands of the log input, while the reflectance image can be found from the high spatial frequencies. However, this assumption, like the one underlying Retinex, also tends not to be true in real images, especially if cast shadows or highlights are present.

Differently from discriminative approaches, *generative* methods create possible surfaces and reflectance patterns that explain the image, then use a model to choose the most likely surface. An approach developed by Sinha and Adelson in [7] works in a domain of painted polyhedral/origami objects, with presegmented junctions and regions. In [8] the authors generate a synthetic world of scenes and their corresponding rendered images, modeling their relationships with a Markov network.

In a different direction, Weiss [9] proposed to use multiple images where the reflectance is constant, but the illumination changes. Using many images ensures that the problem is well-posed, but implies that the application of the method is quite restrictive.

A recent work by Tappens et al. [10] takes a discriminative approach by classifying the derivatives of the image using both classifiers based on the image colour information and classifiers trained to recognize local image patterns to distinguish derivatives caused by reflectance changes from derivatives caused by shading. The method is based on learning estimators that predict filtered versions of the desired image. However, in this problem finding a training set of real data is not trivial. In [11] a set of ground truth intrinsic images is created using colour to measure shading separately from reflectance: test images are pieces of paper coloured with a marker, not visible in the green channel. Even though the resulting images are ground truth decompositions this test set remain quite artificial and it is not clear how it can generalize the behaviour of real cases.

In this paper we build on [12], where an invariant gray-scale image, independent from illumination condition, is derived from one uncalibrated image, without any learning operations. The invariant image is then used to produce a *shadow-free* image.

In this paper, instead, we use the same invariant image to produce the reflectance intrinsic image, i.e., a *shading-free* image, and the illumination image as well. The results are comparable to those reported in [10], but with no need for a training set.

2. PROBLEM FORMULATION AND OVERVIEW

Assuming a linear response of the camera, the input image $I(x, y)$ is modelled as the product of the shading image $S(x, y)$ and the reflectance image $R(x, y)$:

$$I(x, y) = S(x, y) \cdot R(x, y). \tag{1}$$

The goal is to recover $S(x, y)$ and $R(x, y)$ from $I(x, y)$.

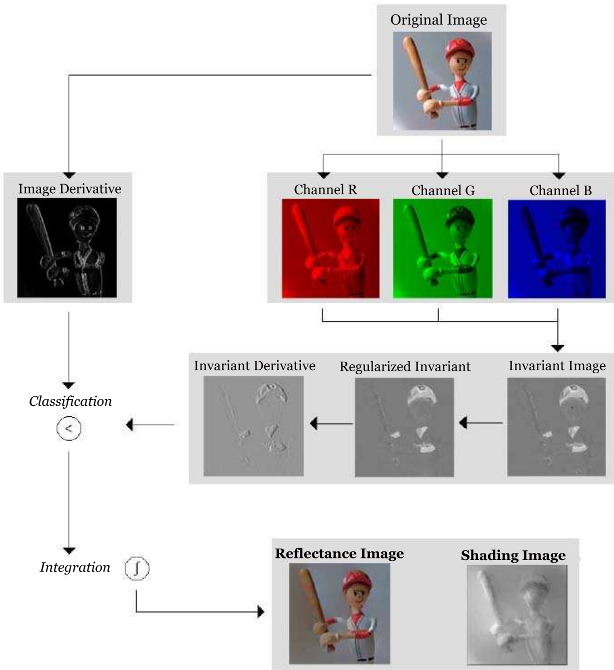


Fig. 1. The overall technique for the extraction of intrinsic images.

This formulation implies the implicit assumption that the surfaces are Lambertian. Although this might be considered a restrictive hypothesis, it offers a tractable starting point from which techniques for image decomposition can be developed.

Another assumption, that will be justified later on, is that there are no sharp illumination changes in the image (i.e., cast shadows). This is coherent with the fact that this paper is complementary with respect to [13], which deals with shadows removal.

As in most of the other approaches in literature, instead of estimating $S(x, y)$ and $R(x, y)$ directly, we attempt to estimate their derivatives: we determine which variations in the original image are due to shading effects and which are due to reflectance changes. We assume that it is unlikely that significant shading boundaries and reflectance edges occur at the same point. This permits to treat every image derivative as either caused by shading or reflectance, thus reducing the problem to binary classification of the image's x and y derivatives.

This classification is realized with the help of the *invariant image* [12]. This image is obtained by projecting the log-chromaticity image in a direction independent from lighting effects. Although it is independent from illumination, it can not be taken as the target reflectance image, for it is only gray-scale and lacks some fine reflectance details as well. However, by construction its derivatives are due to changes in reflectance only, so they provide useful information to guide the classification of derivatives of the original image.

The idea is to compare the derivatives of the original and invariant image and to classify a small difference between the two as a shading effect and a great difference as a reflectance effect. In fact, areas of smooth illumination change in the original image tend to become nearly flat areas in the invariant one (small difference). On the contrary, sharp variations in the original image, due to reflectance, yield great variations between the two derivatives. This is the basis for classification in our approach, and it permits us to avoid a learning process. The overall technique is schematized in Fig. 1 and it will be detailed in the following Sections.

3. EXTRACTION OF THE INVARIANT IMAGE

Consider a fairly narrow-band camera, with a RGB sensor, which images a set of coloured Lambertian surface patches under the daylight. In the log-chromaticity space, given by logarithm of the channel ratios {R/G, B/G}, every pixel in each patch is approximatively collapsed into the same dot. As the illuminant changes, the log-chromaticity points move along an approximately straight line which is independent of the magnitude and the direction of the lighting. Projecting colours perpendicular to this invariant direction due to lighting change produces a 1D gray-scale image that is invariant to illumination. In [12], an algorithm for the extraction of such image, invariant to illuminant colour and intensity, from an uncalibrated image is proposed. For the sake of space we refer to [12] for description and details of this algorithm. An example of invariant image extraction is depicted in Fig. 2.

(a) (b)

Fig. 2. Extraction of invariant image example: (a) original image; (b) gray-scale invariant image.

If the original image does not match with the assumptions the resulting invariant image turns out corrupted by noise. Then, before proceeding any further, we try to restore such image by applying a regularization process based on a Markov Random Field (MRF) [14]. Our a-priori model for the invariant image is piecewise constant. The minimization is performed by a simulated annealing algorithm using Metropolis sampler [15, 14].

4. CLASSIFICATION

Let $\rho_k(x, y)$ denote the logarithm of the grey-scale image corresponding to a single channel of the sensed colour image. We first calculate the gradients:

$$\begin{aligned} \nabla_x \rho_k(x, y) &= \frac{\partial}{\partial x} \rho_k(x, y) \\ \nabla_y \rho_k(x, y) &= \frac{\partial}{\partial y} \rho_k(x, y) \end{aligned} \tag{2}$$

Then, in order to take into account the variations derived from the three channel, we consider their mean $\bar{\nabla}_x$ and $\bar{\nabla}_y$.

Similarly, we calculate the derivatives of the regularized invariant image $\mathcal{I}(x, y)$:

$$\begin{aligned} \nabla_x \mathcal{I}(x, y) &= \frac{\partial}{\partial x} \mathcal{I}(x, y) \\ \nabla_y \mathcal{I}(x, y) &= \frac{\partial}{\partial y} \mathcal{I}(x, y) \end{aligned} \tag{3}$$

The objective now is finding a binary classification of $\bar{\nabla}_x$ and $\bar{\nabla}_y$, by comparing them with $\nabla_x \mathcal{I}$ and $\nabla_y \mathcal{I}$. The idea is that smooth variations in illumination are flat regions in the invariant image. So, if the difference between the two derivatives is small, the variation must be classified as due to shading, as due to reflectance otherwise. This is realized with the following thresholding operation:

$$q_i(x,y) = \begin{cases} 1 & \text{if } \|\bar{\nabla}_i(x,y)\| > \|\nabla \mathcal{I}_i(x,y)\| \\ & \text{and } \|\bar{\nabla}_i(x,y)\| - \|\nabla \mathcal{I}_i(x,y)\| \leq \tau \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

The function q_i is equal to 1 when $\bar{\nabla}_i(x,y)$ is due to shading, 0 in the other case. q_i is a classification of the derivatives of the original image, and from that we can obtain $\mathcal{F}_{sh,x}$ and $\mathcal{F}_{sh,y}$, the derivatives of the shading image, as well as $\mathcal{F}_{ref,x}$ and $\mathcal{F}_{ref,y}$, the derivatives of the reflectance image.

5. INTEGRATION

Once the derivatives of the shading and reflectance images are estimated, they can be used to recover the actual images. Each derivative represents a set of linear constraints on the image, and using both derivative images results in a overconstrained system. We recover each intrinsic image from its derivatives with the same method used by Weiss in [9] to find the pseudoinverse of the unconstrained system of derivatives. If f_x and f_y are the filters used to compute the x and y derivatives and $\mathcal{F}_{sh,x}$ and $\mathcal{F}_{sh,y}$ are the estimated derivatives of shading image, then the solution for $S(x,y)$ is:

$$S(x,y) = g*[(f_x(-x,-y)*\mathcal{F}_{sh,x}) + (f_y(-x,-y)*\mathcal{F}_{sh,y})], \quad (5)$$

where $*$ is convolution, $f(-x,-y)$ is a reversed copy of $f(x,y)$, and g is the solution of

$$g*[(f_x(-x,-y)*f_x(x,y)) + (f_y(-x,-y)*f_x(x,y))] = \delta. \quad (6)$$

In this work, f_x and f_y are $[-1,1]$ filters. The reflectance image is found in the same fashion.

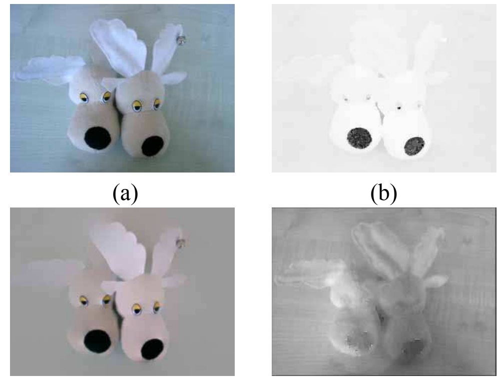
6. EXPERIMENTAL RESULTS

We applied the approach to real data. We firstly calculated the invariant image, and restored it using MRF; we classified derivatives of the original images by comparison with the derivatives of such invariant image; finally, we obtained the actual shading and reflectance image by integration.

The first example, the *Reindeers* image, shows the effectiveness of the approach (Fig. 3). As the reader can notice, the shape and shading effects, on the ears for example, are correctly in the shading image, while the reflectance image contains only colour information.

The *Baseball* image is taken from Tappens et al. [10]. Fig. 4 shows the result obtained with our method. Comparing them with [10], reflectance images are good approximations of a purely flat coloured image, while in both cases the shading images have some defects: in our case, some parts are too smooth, while in Tappens et al. some reflectance effects are present.

The *Lego* image is taken from [10] too, and in Fig. 4 the results are shown. Using our method, the reflectance image is clear and sharp, while the one obtained by Tappens et al. [10] is quite blurred, meaning that some edge points due to reflectance are missed. On the contrary, our shading image is less refined, with lack of some details.



(a) (b) (c) (d)

Fig. 3. Example of extraction of intrinsic images: (a) original image; (b) invariant image restored by MRF; (c) reflectance image; (d) shading image.

Results from an outdoor image, *Child*, are also depicted in Fig. 4. As the reader can notice, the invariant image correctly does not contain illumination effects, and this leads to an almost correct separation into reflectance and shading images. In [10], the shading image appear quite artificial, with over-emphasized details; the one obtained with our method, instead, correctly recover shape and shading information.

7. COMMENTS AND FUTURE WORK

Imaging is a many to one mapping which confounds intrinsic images into the sensed image. It is no surprise, then, if the problem of extracting intrinsic images is a tough one, and it has not been solved satisfactorily, yet. We presented here a technique for recovering shading and reflectance intrinsic images. The results are comparable with the state of the art, but they are obtained without any learning processes to classify derivatives. This work is motivated by the fact that training set of reflectance and shading real data is very difficult to obtain: the invariant image we use proves to be a good instrument to guide classification. However, such invariant image can be obtained only under quite restrictive assumptions.

Moreover, the assumption that a variation in an image is due to reflectance *or* to shading does not always hold. In fact, the outline of an object, for example, is usually both a change of reflectance *and* of shading. One could circumvent this problem if the geometry of the scene is known. In the context of 3D reconstruction, where the intrinsic reflectance image is used as a texture maps, this is the case. The 3D model can be projected into the image and used as additional information to extract intrinsic images.

The plan for future work is to add geometric information and to further investigate other strategies to obtain images that are invariant to illumination, working in less restrictive conditions and possibly with no loss of reflectance information.

Acknowledgments

Images used in the experiments are courtesy of Marshall Tappen. Thanks to Alessandro Pontone for providing part of the code.

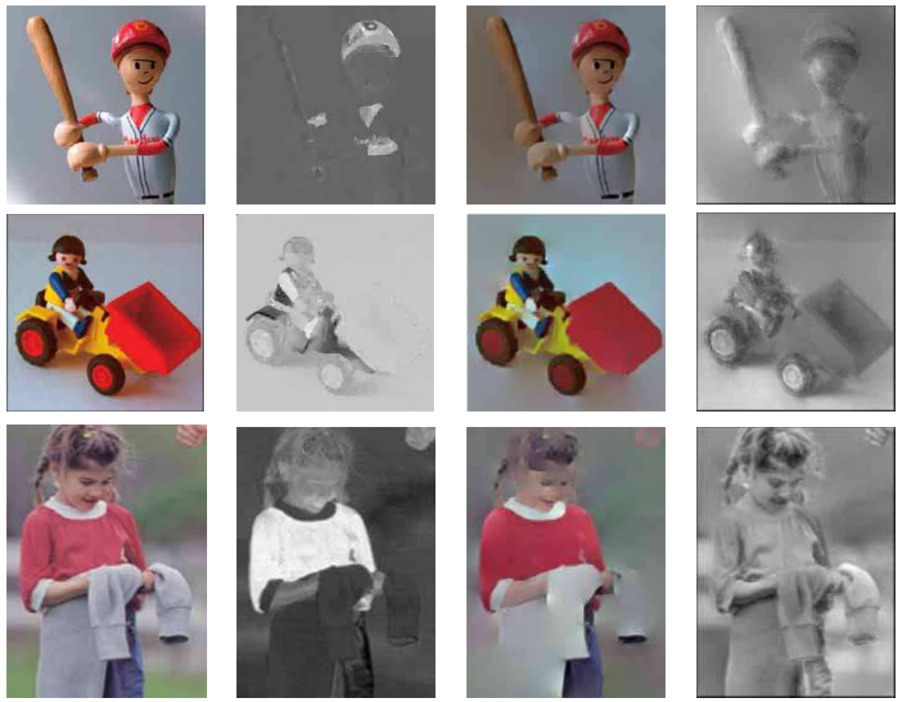


Fig. 4. Examples of extraction of intrinsic images: (a) original image; (b) invariant image restored by MRF; (c) reflectance image; (d) shading image.

8. REFERENCES

[1] H.G. Barrow and A.P. Pentland, "Recovering intrinsic scene characteristics from images," in *Computer Vision Systems*, D. Knill and W. Richards, Eds. Academic Press, 1978.

[2] E.H. Land and J.J. McCann, "Lightness and retinex theory," *J. Optical Soc. Am.*, vol. 61, pp. 1–11, 1971.

[3] B.K.P. Horn, *Robot Vision*, chapter 9, Cambridge Mass: MIT Press, 1986.

[4] W.T. Freeman and P.A. Viola, "Bayesian model of surface perception," in *Advances in Neural Information Processing Systems*, 1998.

[5] M. Bell and W.T. Freeman, "Learning local evidence for shading and reflection," in *Proceedings of the IEEE International Conference on Computer Vision*, 2001.

[6] J. Thomas A.V. Oppenheim, R.W. Shafer and G. Stockham, "Nonlinear filtering of multiplied and convolved signals," *IEEE Trans. Audio and Electroacoustics*, vol. 16, no. 3, pp. 437–466, 1968.

[7] P. Sinha and E.H. Adelson, "Recovering reflectance in a world of painted polyhedra," in *Proceedings of the IEEE International Conference on Computer Vision*, 1993, pp. 156–163.

[8] E.C. Pasztor W. T. Freeman and O.T. Carmichel, "Learning low level vision," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 25–47, 2000.

[9] Y. Weiss, "Deriving intrinsic images from image sequences," in *Proceedings of the IEEE International Conference on Computer Vision*, 2001.

[10] W. T. Freeman M. F. Tappen and E. H. Adelson, "Recovering intrinsic images from a single image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 9, pp. 1459–1472, 2005.

[11] W. T. Freeman M. F. Tappen and E. H. Adelson, "Estimating intrinsic component images using non-linear regression," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2006, pp. 1992–1999.

[12] M.S. Drew G. D. Finlayson and C. Lu, "Intrinsic images by entropy minimization," in *Proceedings of the European Conference on Computer Vision*, 2004, pp. 582–595.

[13] C. Lu G. D. Finlayson, S.D. Hordley and M.S. Drew, "On the removal of shadows from images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 59–68, 2006.

[14] J. K. Marroquine, *Probabilistic Solution of Inverse Problem*, Ph.D. thesis, Massachusetts Institute of Technology, 1985.

[15] C.D. Gellat Jr S. Kirkpatrik and M.P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4, pp. 671–680, 1983.