A Survey of Specularity Removal Methods

Alessandro Artusi,[1] Francesco Banterle[2] and Dmitry Chetverikov[3]

[1]Cyprus Institute CaSToRC, Cyprus
artusialessandro4@googlemail.com
[2]Visual Computing Lab. ISTI-CNR, Italy
frabante@gmail.com
[3]MTA SZTAKI, Budapest, Hungary
csetverikov@sztaki.hu

Abstract
The separation of reflection components is an important issue in computer graphics, computer vision and image processing. It provides useful information for the applications that need consistent object surface appearance, such as stereo reconstruction, visual recognition, tracking, objects re-illumination and dichromatic editing. In this paper we will present a brief survey of recent advances in separation of reflection components, also known as specularity (highlights) removal. Several techniques that try to tackle the problem from different points of view have been proposed so far. In this survey, we will overview these methods and we will present a critical analysis of their benefits and drawbacks.

Keywords: Specularity Removal, Specular-free image, diffuse and specular reflections, polarization, intrinsic images, multi-flash, inpainting, reflection model

ACM CCS: I.4.8 [Scene Analysis]: Colour shading, shadowing, diffuse and specular reflections, highlights; I.4.3 [Enhancement]: Specularity removal; General; [I.4.0]: Dichromatic reflection model

1. Introduction

Specularity removal can be viewed as the general problem of extracting information contained in an image and transforming it into certain meaningful representation. This representation is able to describe the intrinsic properties of the input image, and it is well-known under the name of intrinsic images introduced by Barrow *et al*. [BT78]. Several characteristics of the original input image can be defined as intrinsic images: illumination colour, illumination geometry, surface reflectance, surface geometry and view-point [TI04]. In our case, the two intrinsic characteristics that must be extracted are the *diffuse* and the *specular* reflection components.

Several applications in computer graphics, computer vision and image processing can benefit from using this meaningful information in the form of intrinsic images. On the one hand, the presence of specular reflection is inevitable since in the real world we have many materials that show both diffuse

and specular reflections. On the other hand, many algorithms in computer graphics, computer vision and image processing assume perfect diffuse surfaces and consider locations of specular reflection as outliers. This simplification reduces the robustness of these algorithms when used in applications where specular surfaces should be considered. As a result, the applicability of these algorithms is limited.

In particular, stereo reconstruction, visual recognition and tracking need to have a consistent surface appearance of an object in different images. The appearance of a surface can significantly vary in the presence of highlights. Specular highlights may cover surface details and appear as additional features that are not intrinsic to the object [TLQ06]. Acquired textures with the presence of highlights are a classical example of how details may be lost and the simulation of the changes of the light source position will be affected by the highlights and shadows generated by the light source used during the acquisition process.

In this paper, we describe recent work on separation of reflection components, also known as Specularity (Highlights) Removal. These techniques can be used to separate the diffuse and specular reflection components of an image.

In the following sections we will introduce the specularity removal problem, followed by methods classification based on their characteristics and also some basic concepts, used through the survey, will be introduced (Section 2). Afterwards the techniques (Sections 3 and 4) will be presented, discussing their advantages and drawbacks (Section 5).

2. Overview

The information contained in the image of an object or a scene can be extracted in form of different representations or descriptions. These representations or descriptions are also called intrinsic image properties, and the result of separation is an intrinsic image that describes a specific intrinsic property.

In our case, the intrinsic image properties are the two kinds of reflection components: the *interface* one, also called specular, and the *body* one, also called diffuse.

Why do we need to extract this information?

Most algorithms used in numerous tasks of computer vision, computer graphics and image processing, such as stereo matching, photo-consistency, segmentation, recognition and tracking work under the assumption of perfect Lambertian surfaces (perfect diffuse reflection); they consider specular pixels (highlights) as outliers or noise.

Unfortunately, in the real world we have many materials that show both diffuse and specular reflections. When these algorithms are used directly on surfaces that show also specular reflection, this can lead to problems: stereo mismatching, false segmentations and recognition errors. For example, the photo-consistency based registration often used in three-dimensional (3D) model reconstruction assumes Lambertian surfaces; otherwise, the measure of how consistent the views are will fail because of specularities.

In this case, the separation of specular and diffuse components is useful, and having a specular-free (diffuse) map is often desired. Figure 1 shows an example of how the two reflection components can be processed separately and afterwards recombined to produce particular visual effects [MZK*06]. This process is called *dichromatic editing*. By modulating the two reflection components, visual effects such as wetness, make-up, additional light or changes in the appearance, can be produced.

Figure 2 shows an image where details and colours are completely washed out in the highlights region. This case is typical in the texture acquisition process, where the presence of a source light may generate the presence of highlights. To

(a) (b) (c) (d)
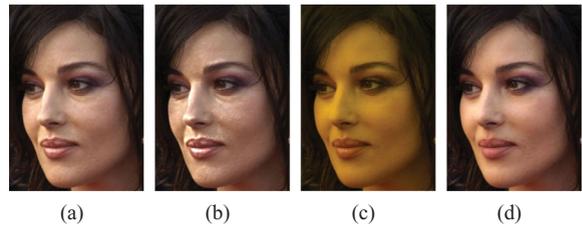
Figure 1: *Example of different visual effects produced by dichromatic editing [MZK*06]: (a) is the input image, (b) wetness effect, (c) skin colour change and (d) effect of make-up. Input image courtesy of S.P. Mallick.*
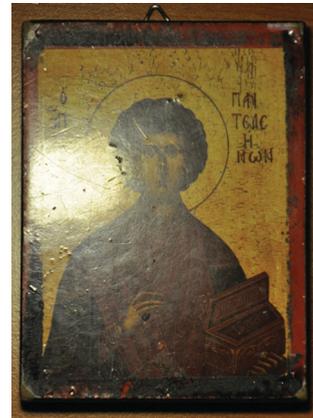
Figure 2: *Example of an image where the presence of highlights generates the loss of details and colour information. Details and colours are completely washed out in the highlights region.*

recompose the details and the colour information, highlights removal techniques are required. Moreover, the simulation of light movements will not be realistic enough if the highlights are not removed.

On the other hand, specular reflectance can play a role in human perception, and some algorithms rely on this information [MZK*06]. For example, Blake *et al.* [BB88] reconstruct geometry from specularity. In a shape-from-specularity system, Healey *et al.* [HB88] derive and use relationships between the properties of a specular feature in an image and the local properties of the corresponding surface. Osadchy *et al.* [OJR03] have shown how to use the highlights information to recognize challenging shiny transparent objects.

Why is the component separation problem difficult?

In general, recovering two intrinsic images from an input image is a classical ill-posed problem [Wei01]. In fact, the number of unknown variables is larger than the number of equations. We have only one equation that defines the total radiance as the sum of different terms as will be described in

Table 1: *The classification of highlights removal methods. These methods are classified based on the number of images used as input and the type of information used.*

Tech. - Sec.	Type	Single image		Multiple images		
		CSA	NA	IS	MFI	POL
CRM - 3.1.1	G	x	–	–	–	–
TDA - 3.1.2	L	x	–	–	–	–
BM - 3.1.3	G	x	–	–	–	–
USFI - 3.2.1	L	–	x	–	–	–
PDE - 3.2.2	L	–	x	–	–	–
IT - 3.2.3	L	–	x	–	–	–
TS - 3.2.4	L	–	x	–	–	–
CC - 3.2.5	L	–	x	–	–	–
FR - 3.2.6	G	–	x	–	–	–
HM - 4	G	–	–	x	–	–
HFLF - 4	G	–	–	x	–	–
MBS - 4.1	G	–	–	x	–	–
MII - 4.2	L	–	–	x	–	–
CP - 4.3	G	–	–	–	–	x
MF - 4.4	L	–	–	–	x	–

L is used for local and *G* is used for global approach. The number in the first column, indicates the paragraph where the technique appears in the survey.

Section 2.2. If one analyses this equation, the only term that can be known is the chromaticity of the illumination (colour associated with the specular component); all the other terms are unknown.

Highlights removal can be reduced in complexity when geometry is known. In this case, the problem is straightforward to solve and it can be efficiently solved, for example see Dellepiane *et al.*'s work [DCC*10]. However, in most of the cases this information is not available making the solution of the separation problem difficult.

Various specularity removal techniques are available in literature; they differ in the information they use and in how this information is used. Table 1 summarized how these techniques may be classified and in which paragraph of this survey they are explained. The notation used in Table 1 for the categories is the following: CSA, colour space analysis; NA, neighbourhood analysis; POL, polarization; IS image sequences; MFI, multiple-flash images. The notation used in Table 1 for the techniques is the following: CRM, using colour reflection model [KSK87, KSK88]; TDA, 2D diagram approach [ST95b, SK00]; BM, Bajcsy *et al.* method [BLL96]; USFI, use of specular-free image [TI05b, YCK06, SC09]; PDE, PDE approach [MZK*06]; IT, inpainting technique [TLQ06]; CC, use of colour information and classifier [TFA05]; TS, separation of highlight reflections on textured surfaces [TLQ06]; FR, Fresnel methods [Ang07]; HM, histogram methods [CGS06]; HFLF, high-low frequency separation [LPD07, NGR06]; MBS, multi-baseline-stereo [LB92, LYK*03, LS01]; MII,
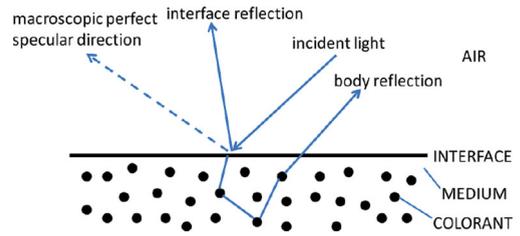
Figure 3: *A graphical representation of the reflection process when light strikes a surface. (After Shafer [Sha85].) Two kinds of reflection are generated: specular and diffuse.*

deriving intrinsic images from image sequences with illumination changes [Wei01]; CP, colour and polarization methods [NFB97, KLHS02, MPC*07, USGG04]; MF, multi-flash methods [FRTT04, ARNL05].

Based on the type of data used as input, there are two main categories: *single-image* and *multi-image* methods. The first category performs the separation of the reflection components using only a single image. The second category makes use of a sequence of images benefiting from the fact that under varying viewing direction diffuse and specular reflections show different behaviours.

In the single-image category, we can further subdivide the techniques based on the information they use: *neighbourhood analysis*, *colour space analysis*. The first group uses the information of the neighbourhood pixels to compute the pixel diffuse colour. The second group considers colour space to analyse the distribution of the diffuse and specular components and uses this information for the separation.

Besides this classification, these approaches can be categorized as *local* or *global* depending on how they use the information contained in the image data. The local methods utilize local pixel interactions to remove the specular reflection, while the global methods estimate diffuse colours of image regions, which implies segmentation.

In the next section, we will review the reflection models and the definitions used by the techniques presented in this survey.

2.1. Basic definitions

As we said earlier, we have two types of reflections: *specular* and *diffuse*. The specular reflection is the light reflected at the interface between the air and the surface. We can notice in Figure 3, that the direction of the specular reflection is different at the macroscopic and microscopic levels. This is because the direction is related to the surface normal that is different at the macroscopic level (reference surface normal) and at the microscopic level (local surface normal). The specular reflection is governed by the well-known Fresnel's law that relates the specular reflectance to the angle of incidence,
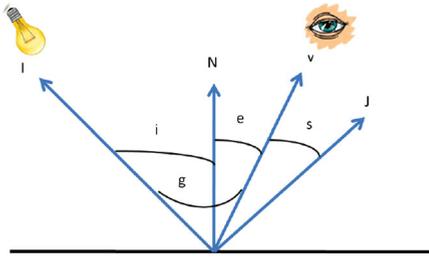
Figure 4: *The reflectance geometry used by Shafer [Sha85] and by Cook and Torrance [CT82] for describing their reflection models.*

the index of refraction of the material, and the polarization of the incoming illumination [Sha85].

The diffuse reflection is generated by the light that penetrates through the interface and passes through the medium, where it undergoes scattering from the colourant. As it can be noted from Figure 3, in this case the light is either transmitted through the material (when the latter is not opaque), absorbed by the colourant, or reflected through the same interface by which it entered [Sha85, NS92]. The reflected light represents the diffuse reflection. The colour of the diffuse reflection is, in general, different from that of the illumination, and the diffuse reflection is usually considered to be unpolarized [Sha85]. This helps the reflectance separation process since the degree of light polarization can be considered as a strong indicator of specular reflection.

The terminology used by Shafer [Sha85], for describing the reflectance geometry is illustrated in Figure 4 . The directions are denoted as follows: I is the illumination direction, N is the normal surface, V is the viewing direction, and J is the perfect specular direction at the macroscopic level. The angles are denoted as follows: i is the angle of incidence, e the angle of emittance, g the phase angle, and s the off-specular angle.

Mainly, two reflection models have been used in the context of specularity removal: *Dichromatic* [Sha85] (Section 2.2) and *Cook and Torrance* [CT82] (Section 2.3).

2.2. The dichromatic reflection model

The dichromatic reflection model [Sha85] is a simple model of reflectance in which the total radiance is the sum of two independent terms: the radiance L_s of the light reflected at the interface and the radiance L_d of the light reflected from the surface body:

$$L(\lambda, i, e, g) = L_d(\lambda, i, e, g) + L_s(\lambda, i, e, g). \quad (1)$$

Equation (1) can be decomposed for better understanding of the components:

$$L(\lambda, i, e, g) = w_d(i, e, g)c_d(\lambda) + w_s(i, e, g)c_s(\lambda). \quad (2)$$

That is, each of the two terms in Equation (1) is formed by different terms c and w. The term c called *composition* is the relative spectral power distribution (SPD), which depends only on the wavelength λ. The term w called *magnitude* is a geometric scale factor which depends only on the geometry [Sha85]. These two terms relate the reflectance geometry to the wavelength of the light that strikes the material surface.

The main assumption is about the surface that must be an opaque inhomogeneous medium with one significant interface. There is another assumption about the material surface made by the dichromatic reflection model [Sha85]: The surface is not optically active. This means that the surface does not have fluorescent and thin film properties, and the colourant is uniformly distributed.

The assumptions about the reflection are the following:

- The reflection from the surface is invariant with respect to rotation around the surface normal.
- There are no inter-reflections among surfaces.
- The body reflection is Lambertian, which means that the brightness is independent from the viewing direction.
- The specular reflection has the same colour of the illumination and tends to be polarized. As it will be shown later in this survey, considering the specular reflection having the same colour of the illumination will simplify the reflectance separation process. However, this is not always true in the real world as it has been demonstrated by Angelopoulou [Ang07] (Section 3.2.6). The assumption of polarization is used by the polarization based techniques (Section 4.3) to determine the colour of the specular component, simplifying the reflectance separation process.

Despite these assumptions, the model [Sha85] has a high level of generality. In fact, no assumptions are made about the *imaging geometry*. The model can be applied to either curved or planar surfaces, as well as to textured surfaces. No specific functions are assumed by the model for describing the terms w and c. This means that no specific geometric model of highlights is used.

No assumptions are made about the light source either: the model can be applied to pointwise, extended, and infinitely distant light sources. The model does not assume that the amount of illumination is the same everywhere in the scene.

The model [Sha85] suffers from some flaws, but they have a minimal impact on the usefulness of the model. In reality, both the specular reflection and the diffuse reflection exhibit an interdependence between wavelength and geometry. In [Sha85], this effect is estimated to be negligible.

How is it possible to relate the SPD of the light to its colour coordinates?

The spectral projection, as proposed by Shafer [Sha85], is a way to compute the colour values from the SPD of the measured light.

Before introducing the concept of spectral projection, it is necessary to recall shortly how a colour value is obtained in a colour camera

$$I_X = \begin{pmatrix} r_X \\ g_X \\ b_X \end{pmatrix} = \begin{pmatrix} \int_{\lambda} X(\lambda)\bar{r}(\lambda)\,d\lambda \\ \int_{\lambda} X(\lambda)\bar{g}(\lambda)\,d\lambda \\ \int_{\lambda} X(\lambda)\bar{b}(\lambda)\,d\lambda \end{pmatrix}, \qquad (3)$$

where $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ and $\bar{b}(\lambda)$ are the responsitivities in the three channels, and $X(\lambda)$ is the SPD. The spectral projection is a linear transformation [Sha85], which means that in the case of a mixture of two or more SPDs the resulting colour is the sum of the corresponding pixel colours taken in the same proportion:

$$I_{\alpha X + \beta Y} = \alpha I_X + \beta I_Y. \qquad (4)$$

Applying this result to the Dichromatic Reflectance Model (2), we obtain

$$I = w_{\mathrm{d}} B + w_{\mathrm{s}} G, \qquad (5)$$

where I is the pixel colour, w_{s} and w_{d} the magnitudes of reflection at the point, and B and G the colours of the diffuse and specular reflections of the material, respectively. One can observe that the scale factors w_{d} and w_{s} vary from point to point, whereas the colours B and G are the same in all points of the surface, because they are the spectral projections of $c_{\mathrm{d}}(\lambda)$ and $c_{\mathrm{s}}(\lambda)$ that do not vary with the geometry under the assumption of uniform illumination. Equation (5) does not assume any specific G, or any specific colour responses of the camera.

It is easy to extend the model (2) to the case of *ambient* light. The reflected light of ambient illumination contains specular reflection and diffuse reflection. This light is incident from, and reflected into, all directions equally. From these two considerations we can modify the dichromatic model adding a single term $L_{\mathrm{a}}(\lambda)$ that represents the reflection caused by the diffuse illumination. In this case, Equation (2) is modified as

$$L(\lambda, i, e, g) = w_{\mathrm{d}}(i, e, g)c_{\mathrm{d}}(\lambda)$$
$$+ w_{\mathrm{s}}(i, e, g)c_{\mathrm{s}}(\lambda) + L_{\mathrm{a}}(\lambda), \qquad (6)$$

whereas Equation (5) is modified as

$$I = w_{\mathrm{s}} B + w_{\mathrm{d}} G + I_{\mathrm{a}}, \qquad (7)$$

where I_{a} is the colour associated with the light reflected from the diffuse illumination $L_{\mathrm{a}}(\lambda)$.

2.3. Cook and Torrance reflectance model

A model similar to the Dichromatic one has been introduced by Cook and Torrance in the 1982 [CT82]. This model also considers the total radiance associated to the total light reflected (bidirectional reflectance) at a certain point p of a surface material as the sum of the two terms specular and diffuse reflections. Their model can also be extended to take into account the ambient light; it assumes the diffuse and the ambient components to reflect light equally in all directions, that is, independently from the location of the viewer. In this way, only the specular reflection depends on the viewer position. To describe the specular reflection the authors assume that the surface consists of microfacets that reflect specularly, and only those that have the normal in the direction of H contribute to the specular component that can be modelled as

$$L_{\mathrm{s}} = \frac{F}{\pi} \frac{D}{(N \cdot I)} \frac{T}{(N \cdot V)}, \qquad (8)$$

where F represents the Fresnel term and describes how the light is reflected from each microfacet; it is a function of angle and wavelength. T represents the geometrical attenuation factor and D is the microfacets distribution term. As observed by Angelopoulou [Ang07], the Cook–Torrance model states that the specular reflection varies over light spectrum due to the fact that it depends on the index of refraction, which is function of wavelength. On the other hand, the model assumes that for dielectric materials this aspect is negligible, so the colour can be considered the same as the colour of the light source. In Section 3.2.6, we will show that this is important and can be used for reflectance component separation [Ang07].

2.4. Fresnel term computation

The Fresnel term F, as explained in Section 2.3, describes the reflection and transmission of electromagnetic waves at an interface. F is defined as the ratio of reflected light over incident flux densities at different wavelengths

$$F = \frac{1}{2}\left(r_{\mathrm{N}}^2 + r_{\mathrm{L}}^2\right), \qquad (9)$$

where r_N and r_L are the reflection coefficients with respect to the incident plane and the plane perpendicular to the incident one, respectively. They are given by the following Fresnel equations [Hec98]:

$$r_{\mathrm{N}}(\lambda) = \frac{n_{\mathrm{t}}(\lambda)\cos\Theta_{\mathrm{i}} - n_{\mathrm{i}}(\lambda)\cos\Theta_{\mathrm{t}}}{n_{\mathrm{i}}(\lambda)\cos\Theta_{\mathrm{t}} + n_{\mathrm{t}}(\lambda)\cos\Theta_{\mathrm{i}}}, \qquad (10)$$

$$r_{\mathrm{L}}(\lambda) = \frac{n_{\mathrm{i}}(\lambda)\cos\Theta_{\mathrm{i}} - n_{\mathrm{t}}(\lambda)\cos\Theta_{\mathrm{t}}}{n_{\mathrm{i}}(\lambda)\cos\Theta_{\mathrm{i}} + n_{\mathrm{t}}(\lambda)\cos\Theta_{\mathrm{t}}}, \qquad (11)$$

where Θ_{i} and Θ_{t} are the angle of incidence and transmittance respectively. n_{i} and n_{t} are the indices of incident and transmitting media respectively. As it can be seen these two

terms depend on the wavelength λ, which means that F is also a function of wavelength. Combining Equations 10 and 11 with Equation (9) and using the Snell's Law, we obtain:

$$F(\lambda, \Theta_i)$$

$$= \frac{1}{2}\left(\frac{n_i^2(\lambda)\cos\Theta_i^2 + J - 2n_i(\lambda)\cos\Theta_i\sqrt{J}}{n_i^2(\lambda)\cos\Theta_i^2 + J + 2n_i(\lambda)\cos\Theta_i\sqrt{J}}\right)$$

$$+ \frac{1}{2}\left(\frac{n_t^2(\lambda)\cos\Theta_i^2 + n_{it}^2(\lambda)J - 2n_i(\lambda)\cos\Theta_i\sqrt{J}}{n_t^2(\lambda)\cos\Theta_i^2 + n_{it}^2(\lambda)J + 2n_i(\lambda)\cos\Theta_i\sqrt{J}}\right),$$
(12)

where $n_{it}(\lambda) = n_i(\lambda)/n_t(\lambda)$, $J = J(\lambda, \Theta_i) = n_t^2(\lambda) - n_i^2(\lambda) + n_i^2(\lambda)\cos\Theta_i$.

2.5. Polarization Model

Based on the assumption that the specular component tends to be polarized, the colour of the specular component can be determined using a polarized filter. The polarization filter is placed in front of the sensor that measures the light reflected by the material surface. One important aspect is that when the material is perfectly diffuse, rotating the polarization filter does not alter the image brightness. In other words, the diffuse component is perfectly unpolarized. Because the specular component is highly polarized, it varies as a cosine function. The specular component is formed by two terms: a constant term I_{sc} and a cosine term with amplitude I_{sv} [NFB97]

$$I(\theta) = I_d + I_{sc} + I_{sv}\cos 2(\theta - \alpha),$$
(13)

where θ is the angle of the polarization filter and α the phase angle. The values of I_{sc} and I_{sv} depend on the material properties and the angle of incidence. The Fresnel ratio relates the two specular component terms to the material properties and the angle of incidence.

2.6. Flash image model

For a static scene, the scene radiance of an image captured using a flash can be viewed as a linear combination of the radiance due to the flash and the ambient illumination [ARNL05]. The irradiance map of a linear-response camera can be modelled as

$$I = F + A = \phi P + \alpha E,$$
(14)

where ϕ and α are the radiance maps of the flash and ambient images, respectively, P the flash intensity, and E the exposure time.

3. Single-Image Methods

Several approaches have been proposed in this main class of component separation techniques. These methods use different information for the separation, and all of them have the common advantage of easily reproducing the input. The methods rely only on a single input image. They do not require several images of the same scene obtained from different points of view, which needs time and may cause difficulties in the image acquisition process.

3.1. Colour space analysis

The analysis of the colour space allows to understand the distribution of diffuse and specular components in a colour image, and utilize this information in the separation process. Klinker *et al.* [KSK87, KSK90, KSK88] linked the *RGB* colour space to the dichromatic model and proposed to use colour information to separate reflection components. A fast approach based on a 2D diagram has been presented by Schluns *et al.* [ST95b, ST95a]. Bajcsy *et al.* [BLL96] proposed a method that through colour image segmentation is able to detect diffuse and specular reflections.

3.1.1. *Using a colour reflection model to separate highlights from object colour*

Colour pixels form a dense cluster in the dichromatic plane. Therefore, a relationship between the shape of the colour cluster and the geometric properties of the diffuse and specular reflections can be used to determine the characteristic features of the colour clusters. To relate the properties of the dichromatic model to the shape of the colour cluster, Klinker *et al.* [KSK88] classify colour pixels as *matte*, *highlight* and *clipped*. A matte pixel exhibits only body reflection, a highlight pixel both body and specular reflections. It is worth to notice that the colours in a highlight area that lie on a line of constant body reflection vary only in their respective amounts of specular reflection. The colours of these pixels form a straight highlight line in the colour space parallel to the specular reflection vector. A clipped colour pixel is a highlight pixel at which light reflection exceeds the dynamic range of the camera [KSK87].

This information can be used to detect and remove highlights from colour images. The algorithm by Klinker *et al.* [KSK87] starts with normalizing the input image. This requires the estimation of the illumination colour. Afterwards, the input image is manually subdivided into different areas, and the pixels belonging to the selected areas are projected onto the colour space, where a dichromatic plane is fitted to the colour data from each image area. In each dichromatic plane, a search is performed to find the matte, the brightest highlight and the clipped colour lines. A recursive line splitting algorithms [Pav77] is used to extract these lines, then the lines are classified as matte, highlights and clipped colour vectors.

The specular reflection is removed by projecting the colour of every pixel onto the respective dichromatic plane. When the projected colour is close to a clipped colour line, it is replaced by the colour at the end of the highlight vector.

Then the colour of every pixel along the highlight vector is projected onto the matte line [KSK87]. The result of this process is the diffuse image; a specular (highlight) image can also be generated. Similarly to other methods, in this case the selection of the specular and diffuse regions is done manually by the user.

3.1.2. *Two-dimensional diagram approach*

This fast approach uses a 2D colour representation to perform the separation of reflection components. In particular, it works in the *uv-space*, that is, of the normalized values of the $Y'U'V'$ colour space (*uv*-chromaticities). In addition, a 1D space (hue space, or *h-space*) is used containing the number of pixels per interval of α, the polar angle in the *uv*-space. The approach has the following steps:

- The matte colours are estimated. For this, the *RGB* data is transformed to the *uv*-space and then to the *h*-space. The *h*-space is processed using a morphological filter [ST95b, SK00]. The processed data is searched for maxima. For each maximum in the *h*-space corresponding to a certain value of α, a maximum in the *uv*-space is searched along the line l_α. Each maximum in the *uv*-space is a matte colour. In this way the diffuse components are determined.

- Segmentation is done in the *h*-space to identify the correct diffuse colour for each pixel. The specular component is represented by the illumination colour, and it can be estimated from the input image.

3.1.3. *Approach by Bajcsy et al.*

A novel colour space, called *S-space* was introduced by Bajcsy *et al.* [BLL96] to analyse the colour variations on objects in terms of brightness, hue and saturation. This colour space can help to describe various optical phenomena such as shading, highlight, shadow and inter-reflection [BLL96]. The *S-space* colour space is based on three orthogonal basis functions (axes) S_0, S_1 and S_2. They are chosen to have S_0 aligned with the orientation of the neutral spectrum (i.e. white or grey) in the visible range. This corresponds to the specular reflection. Typical set of basis functions are the Fourier basis functions but other types can be chosen. The algorithm converts the RGB colour space of the input scene to the *S-space*, then the appearance correlates brightness, hue and saturation are derived from it. During the colour space conversion the scene illumination is discounted using a white reference plate as an active probe. The resulting image will look as if the scene illumination was white. Analysing the surface reflections for a surface with uniform body reflectance in hue and saturation in the *S-space*, Bajcsy *et al.* [BLL96] have observed that the interface reflection always decreases the saturation of body reflection, and increases the brightness. Moreover, they have observed that the shad-

ing does not influence the hue or the saturation of the body reflection. Based on these observations, the authors have proposed an image segmentation technique to detect and separate interface reflection from body reflection. It is based on a decomposition of a three-dimensional space into a set of one-dimensional data such as hue and saturation. The result of discounting the scene illumination is that the interface reflection clusters in the *S-space* are all aligned with the brightness direction. As a consequence of this, the interface reflections have the same hue as the underlying body reflection clusters. Another important observation is that under neutral illumination, the planes formed by the interface and body reflections are parallel to the S_0-axis and perpendicular to the $S_1 - S_2$ plane. Since the interface and body reflections from the same object are aligned at the same hue, under neutral illumination they can be segmented by hue segmentation.

The authors noticed that in each hue segmented plane in the *S-space*, body reflection forms a linear cluster, that has uniform saturation. Where interface reflection exists in addition to body reflection, the saturation is smaller. Once the saturation of the body reflection has been found, it is straightforward to detect the interface reflection. Working in the *S-space* and using multiple views with a spectral differencing algorithm, Lee *et al.* [LB92] avoid the use of segmentation.

3.2. Neighbourhood analysis

Different techniques have been proposed in this category. All of them perform local operations and use colour information in the separation process. Tan *et al.* [TI05b], Yoon *et al.* [YCK06] and Shen *et al.* [SC09] presented fast approaches based on the use of specular-free, or specularity-invariant, image. Mallik *et al.* [MZK*06] introduce a partial differential equation (PDE) that iteratively erodes the specular component at each pixel. Also, inpainting techniques have been applied to achieve the separation [TLQS03] [LM06] [OTS*05]. Tan *et al.* [TLQ06] presented a technique that makes use of texture data from outside a highlight, demonstrating a possibility to overcome problems typical for colour space analysis methods. Tappen *et al.* [TFA05] presented methods that use colour information and image derivative classifiers to recover the diffuse and specular intrinsic properties of an image.

3.2.1. *Use of specular-free image*

These methods are based on the idea of initially generating a pseudo-diffuse component image. This provides a partial separation of the specular component, which is later used to complete the reflection component separation of the original image. The pseudo-diffuse component image is called *Specular-Free Image* because it is essentially a specularity-invariant representation of the input image. This new image is free from specular reflection and it has the same geometrical profile as the diffuse reflection component of
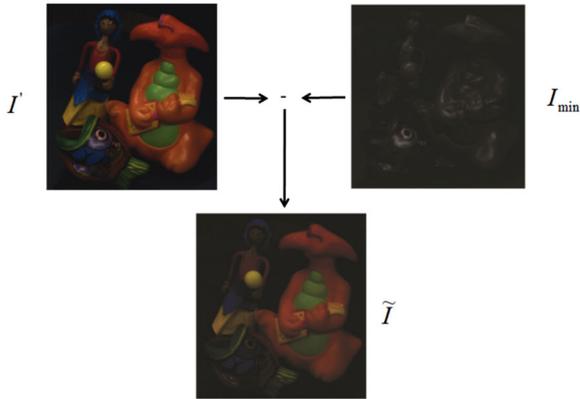
Figure 5: *An intuitive scheme for computing the specular-free image.*

Figure 6: *Specular-to-diffuse mechanism describes how a specular pixel can be purified from its specular component. (After Tan et al. [TI05b].)*

the input image [YCK06]. The meaning of the geometrical profile is the ratio of intensities (or colours) of neighbouring pixels within each diffuse colour region. For instance, suppose that two adjacent pixels have diffuse reflection components like (A, kA), where A denotes a diffuse colour and k is a constant. Then, they will have colours like (B, kB) in the specular-free image. This means that, the ratio of the diffuse reflection components does not change, although the colours of the specular-free image change.

Figure 5 shows an intuitive scheme on how a *specular-free image* can be computed. As input it receives a normalized input image I' and $I_{\min}$, that is the minimum intensity image, and the output is the specular-free image $\tilde{I}$. More details about the different methods used to compute a *specular-free image* are given in the following sections. In these sections, we make use of the dichromatic model's notation (Section 2.2) and in particular the one presented in Equation (5).

Specular-Free Image as Defined by Tan *et al.* [TI05b]

The specular-free image was introduced by Tan *et al.* [TI05b] [TI05a]; it is produced by the so-called *specular-to-diffuse* mechanism. This mechanism is based on the maximum chromaticity and intensity values of diffuse and specular pixels. Chromaticity, or normalized *RGB*, is defined as

$$\sigma(x) = \frac{I(x)}{I_r(x) + I_g(x) + I_b(x)}. \qquad (15)$$

If we consider only one of the two components (diffuse or specular) in Equation (5), the corresponding chromaticity will be independent from the corresponding weighting factor. In this way, we can define *diffuse chromaticity* as

$$\lambda(x) = \frac{B(x)}{B_r(x) + B_g(x) + B_b(x)}, \qquad (16)$$

and *specular*, or *illumination chromaticity* as

$$\Gamma = \frac{G}{G_r + G_g + G_b}. \qquad (17)$$

Equation (5) can be re-written as

$$I(x) = m_{\rm d}(x)\lambda(x) + m_{\rm s}(x)\Gamma, \qquad (18)$$

where

$$m_{\rm d}(x) = w_{\rm d}(x)(B_r(x) + B_g(x) + B_b(x)), \qquad (19)$$

$$m_{\rm s}(x) = w_{\rm s}(x)(G_r + G_g + G_b). \qquad (20)$$

From these definitions we have

$$\sigma_r + \sigma_g + \sigma_b = \lambda_r + \lambda_g + \lambda_b = \Gamma_r + \Gamma_g + \Gamma_b = 1. \quad (21)$$

This method requires that the colour of the specular component be pure white ($\Gamma_r = \Gamma_g = \Gamma_b$). Since it is practically impossible to find a pure white specular component in the real world, we need to normalize the input image. This requires either the knowledge or the estimation of the illumination chromaticity Γ. Introducing $I'(x) = I(x)/\Gamma$ and $\lambda'(x) = \lambda(x)/\Gamma$, we obtain

$$I'(x) = m_{\rm d}(x)\lambda'(x) + m_{\rm s}(x). \qquad (22)$$

The *maximum chromaticity* is defined as

$$\sigma'_{\max}(x) = \frac{\max\{I'_r(x), I'_g(x), I'_b(x)\}}{I'_r(x) + I'_g(x) + I'_b(x)}. \qquad (23)$$

The intensity versus maximum chromaticity space can be used to explain the specular-to-diffuse mechanism as illustrated in Figure 6 . Taking two pixels, a specular x_1 and a diffuse x_2 with the same diffuse chromaticity λ', project them to the maximum chromaticity intensity space. The location of

the diffuse pixel will be to the right of the specular pixel since diffuse maximum chromaticity is larger than specular maximum chromaticity. Tan *et al.* [TI05b, TNI04] discovered that subtracting a small scalar number from Equation (22) (specular pixel) and projecting the subtracted value onto the maximum chromaticity-intensity space forms a curve in the space as shown in Figure 6. The curve obeys the following equation

$$I'_{\max}(x) = \frac{m_{\mathrm{d}}(x)(\lambda'_{\max}(x) - 1/3)\sigma'_{\max}(x)}{\sigma'_{\max}(x) - 1/3} \qquad (24)$$

The intersection point between the curve and the vertical line representing the diffuse pixel represents the diffuse component of the specular pixel $(m_{\mathrm{d}}(x_1)\lambda'_{\max})$. At this point m_{s} is equal to zero. To obtain this point, we first calculate $m_{\mathrm{d}}(x_1)$ derived from Equation (24)

$$m_{\mathrm{d}}(x_1) = \frac{I'_{\max}(x_1)(3\sigma'_{\max}(x_1) - 1)}{\sigma'_{\max}(x_1)(3\lambda'_{\max}(x_1) - 1)}. \qquad (25)$$

In Equation (25), the only unknown entry is the value of the diffuse maximum chromaticity $\lambda'_{\max}(x_1)$. Tan *et al.* [TI05b] proposed to set the diffuse maximum chromaticity to an arbitrary scalar value for all pixels regardless of their colour. To avoid negative values, the authors proposed to choose this value as the smallest value of the maximum chromaticity. Often, this gives noisy results; to avoid them, the value must be selected in the interval [0.5, 1.0].

Finally, the specular-free image defined by Tan *et al.* [TI05b] can be described as

$$\tilde{I}(x) = m_{\mathrm{d}}(x)\tilde{\lambda}(x). \qquad (26)$$

Specular-Free Image as Defined by Yoon *et al.* [YCK06]

Here, the main idea is to find a specularity-invariant quantity for each pixel. In this case, it is also required to either normalize the input image or have the illumination chromaticity as pure white. Let define $I_{\min}(x)$ the smallest of the three values for r, g and b in the input pixel x. Introducing $\lambda_{\min}(x)$ in a similar way, we can write the relation between $I_{\min}(x)$ and $\lambda_{\min}(x)$ as

$$I_{\min}(x) = m_{\mathrm{d}}(x)\lambda_{\min}(x) + \frac{1}{3}m_{\mathrm{s}}(x), \qquad (27)$$

so the specular-free image is obtained as

$$\tilde{I}(x) = I'(x) - I_{\min}(x). \qquad (28)$$

This new image is characterized by having only two bands since one of the three components becomes zero by definition.

Both of the above methods either require that the specular component be pure white or need the knowledge of illumination chromaticity for the normalization of the input image.
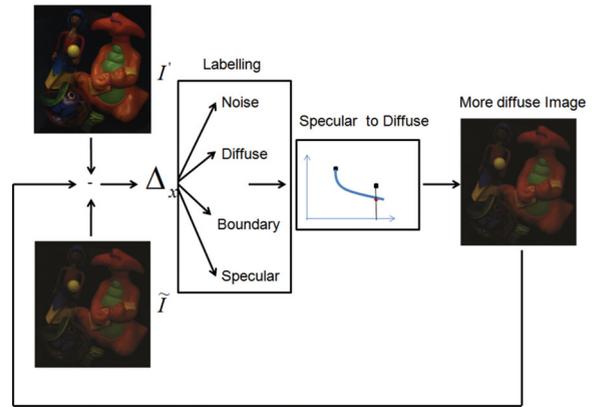
Figure 7: *The scheme of the technique of Tan et al. [TI05b].*

Tan *et al.* [TNI03, TI03] proposed a method to estimate the illumination chromaticity using the Hough Transform. As used by Yoon *et al.* [YCK06], Shen *et al.* [SC09] define the specular-free image as in Equation (28) and have introduced a new step of specular-free image called modified specular-free (MSF) image that is more close to the diffuse component of the input image. This is achieved adding to the specular-free image an offset. The offset may be either constant for all pixel [SZSX08] or being pixel dependent [SC09].

On one hand, these methods are simple and fast because they apply single-pixel operations. On the other hand, the original surface colour is not preserved. This can lead to problems when this colour information is essential. Once the specular-free image has been obtained, it can be used to complete the separation process. Tan *et al.* [TI05b] and Yoon *et al.* [YCK06] developed two different iterative frameworks to perform this operation. Below, we describe the two frameworks separately.

Method of Tan *et al.* [TI05b]

The authors have presented an iterative framework that applies local operations involving two neighbouring pixels. The method assumes uniform surface colour of three neighbouring pixels: one diffuse, c, and two specular, a and b. Figure 7 shows the scheme of the specularities removal technique of Tan *et al.* [TI05b]. The method is based on the difference of logarithmic differentiation of the normalized input and specular-free images

$$\Delta_x = \mathrm{dlog}(I'(x)) - \mathrm{dlog}(\tilde{I}(x)). \qquad (29)$$

In discrete domain, the logarithmic differentiation is computed using [TI05b]

$$\mathrm{dlog}(I(x)) = \log\left(\sum_{i \in (r,g,b)} I_i(x+1)\right) - \log\left(\sum_{i \in (r,g,b)} I_i(x)\right). \qquad (30)$$
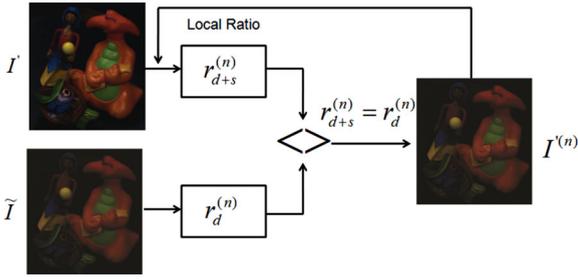
Figure 8: *The scheme of the technique by Yoon et al. [YCK06].*

The algorithm works as follow:

- As input receive two images, the input image normalized I' and the *Specular-Free* image $\tilde{I}$. The difference of logarithmic differentiation is computed using Equation (30).
- The labelling step is used for categorising the pixels of the normalized image I'. In the case Δ_x is equal zero, the pixel is diffuse otherwise is specular. In the case of specular pixels we may have ambiguity. We need to verify if we are in the case of noise or boundary pixels. Boundary pixels are characteristic of multicoloured images, in this case the Δ_x is not equal to zero even if the two neighbouring pixels are both diffuse. This is due to the fact that two neighbouring diffuse pixels have a different surface colour in the boundary condition. This ambiguity is resolved using a simple chromaticity based method, where the chromaticity values of two of the colour channels of the neighbouring pixels are analysed. To identify if a pixel is a noisy one or not, the maximum chromaticity values of two neighbouring pixels are compared. Two specular pixels never have the same maximum chromaticity.
- To the pixels labelled specular, the specular-to-diffuse mechanism is applied to produce a more diffuse image.
- Using this new image, the entire process is repeated as is shown in Figure 7.

In the case of multi-colour surfaces, a colour segmentation procedure is applied, and in each colour region the above method is used for specular separation [TI04].

Method of Yoon *et al*. [YCK06]

Yoon *et al*. [YCK06] also proposed an iterative framework based on the comparison of local ratios. Figure 8 shows the scheme of their technique.

Considering two adjacent pixels, x_1 and x_2, in an image I, the local ratio is defined as

$$r = \frac{\sum\limits_{i\in(r,g,b)} I_i(x_1)}{\sum\limits_{i\in(r,g,b)} I_i(x_2)}. \tag{31}$$

Two kind of local ratios can be defined, the local ratio for the two-band specular-free image, r_d, and the local ratio for the input specular image, r_{d+s}

$$r_d = \frac{m_d(x_1)}{m_d(x_2)}, \tag{32}$$

$$r_{d+s} = \frac{m_d(x_1) + m_s(x_1)}{m_d(x_2) + m_s(x_2)}. \tag{33}$$

A diffuse image can be generated separating its specular component by making r_{d+s} equal to r_d for every pixel. One condition is assumed, namely that at least one diffuse pixel ($m_s = 0$) exists in each colour region. The authors proposed an algorithm that iteratively decreases the specular reflection coefficients to obtain $r_{d+s} = r_d$.

This is achieved by the following steps (*n*th iteration)

$$r_{d+s}^{(n)} = \frac{m_d(x_1) + m_s^{(n)}(x_1)}{m_d(x_2) + m_s^{(n)}(x_2)}. \tag{34}$$

The local ratio of the specular input image can be either bigger or smaller than that of the two-band specular-free image. When it is bigger, we have the following

$$m_s^{(n)}(x_1) > r_d m_s^{(n)}(x_2). \tag{35}$$

In this case, we achieve $r_{d+s}^{(n)} = r_d$ by decreasing $m_s^{(n)}(x_1)$ with the value of $r_d m_s^{(n)}(x_2)$. Considering the normalized image I' (22), for pixel x_1 we have that, at the *n*th iteration, the new normalized image after decreasing the specular coefficients will be equal to

$$I'^{(n)}(x_1) = I'^{(n-1)}(x_1) - \frac{m}{3}, \tag{36}$$

where m is derived as

$$m = \sum_{i\in(r,g,b)} I_i'^{(n)}(x_1) - r_d \sum_{i\in(r,g,b)} I_i'^{(n)}(x_2). \tag{37}$$

When the local ratio of the specular input image is smaller than that of the two-band specular-free image, pixel x_2 is updated as follows:

$$I'^{(n)}(x_2) = I'^{(n-1)}(x_2) - \frac{m}{3}, \tag{38}$$

where

$$m = \sum_{i\in(r,g,b)} I_i'^{(n)}(x_2) - \frac{1}{r_d} \sum_{i\in(r,g,b)} I_i'^{(n)}(x_1). \tag{39}$$

Differently from Tan *et al*. [TI05b] and Yoon *et al*. [YCK06], Shen *et al*. [SC09] proposed a simpler solution to the separation of the reflection components that is based on the assumption that the MSF image is similar to the diffuse one, and this allows to assume that the chromaticity of the two images is close. In this way, the specular component can be seen as related to a scalar value. The authors found that a single scale is sufficient to adjust all pixels in an image. This specular scale can be determined using a least-square technique [SC09].

3.2.2. *The PDE Approach*

The main idea of this approach is to achieve the separation by iteratively eroding the specular component at each pixel. The approach starts with a partial separation provided by illumination-dependent colour space (*SUV*) [MZK*06], and afterwards the separation is completed by multi-scale erosion using partial differential equations (PDEs).

The definition of illuminant-dependent colour space aims at defining transformations that exploit knowledge of the illuminant colour to provide more direct access to the diffuse information in an image. Once we have a direct access to the diffuse information, the separation process is not anymore an ill-posed problem as it was defined in Section 2.

The main advantage of an illuminant-dependent colour space is that it can be derived as a linear combination of the three colour channels of an *RGB* image to obtain one or two *diffuse channels*. The *SUV* colour space is characterized by having two channels, *UV*, purely diffuse, while the *S*-channel contains both specular and diffuse components and it is aligned with the illuminant colour [MZK*06].

This partial specular/diffuse separation is completed using a family of non-linear PDEs that define multi-scale erosion. This method may be used on different images sources such as still and sequence (video) images.

Before defining the PDE, it is necessary to re-formulate the problem through a re-parametrization of the *SUV* colour space using a combination of cylindrical and spherical coordinates.

If I is the input *RGB* image and I_{SUV} is its *SUV* representation with the components I_S, I_U and I_V, the parametrization of *SUV* is

$$\rho = \sqrt{I_U^2 + I_V^2}, \quad \theta = \arctan\left(\frac{I_U}{I_V}\right), \quad \phi = \arctan\left(\frac{I_S}{\rho}\right). \tag{40}$$

The properties of this parametrization are simple to derive; because the components U and V are purely diffuse, the new components ρ and θ are also purely diffuse, that is, independent from the specular component. θ is called the *generalized hue*. Finally, ϕ can be viewed as a linear combination of specular and diffuse components: $\phi = \phi_s + \phi_d$.

Now the problem is reduced to the estimation of $\phi_d(x, y)$, the diffuse contribution to ϕ at each pixel. This is achieved by solving a PDE that iteratively erodes the specular contribution to ϕ and converges to an estimate of ϕ_d.

The PDE governing the evolution of ϕ, $\varepsilon(x, t)$ at scale t, is defined as [MZBK06]

$$\varepsilon_t = -s(\rho, \nabla \rho)(\nabla \varepsilon^T M \nabla \varepsilon)^{1/2}, \tag{41}$$

where M is a matrix that is different for each case, and s is a stopping function depending on ρ and its gradient.

Given an input image parametrized according to Equation (40), at the beginning for scale $t = 0$ we have $\varepsilon(x, 0) = \phi(x)$. The solution to Equation (41) obtained at scale t corresponds to the erosion of ϕ.

This means that the value of ϕ at each image pixel is replaced by the minimum value of the neighbourhood region.

The parameter t defines the shape and size of the region. Since $\phi_d(x) \leq \phi(x)$, if the region has at least one purely diffuse pixel, then $\varepsilon(x, t)$ evaluated at t will converge to $\phi_d(x)$.

3.2.3. *Inpainting techniques*

Inpainting is a procedure that modifies an image in an undetectable way. It is often used to restore damaged paintings and photographs and add or remove elements and objects. Usually, an inpainting technique fills in a damaged region by propagating information from the region boundaries.

Tan *et al*. [TLQS03] noticed that highlights pixels contain useful information for guiding the inpainting process. In fact, in highlights regions, we have typically a single illumination colour and making use of chromaticity analysis we can recover information about the illuminant colour. Another important information, in the highlights regions, is that the colour is the sum of diffuse and specular reflection components. As noticed by Tan *et al*. [TLQS03], an inpainting technique can be subject to illumination constraints.

This is illustrated in Figure 9 where the illumination constraints force the diffuse components of highlight pixels to lie on corresponding parallel lines whose orientations are specified by the illumination colour C_s, and whose positions depend on the observed radiance values I_o. The distance that separates the parallel lines represents shading differences between the pixels.

An illumination-constrained inpainting technique should not be applied across texture edges, where a more appropriate inpainting technique should be used. For this purpose a stopping function is introduced that approaches zero for large gradient preventing inpainting across edges [TLQS03].
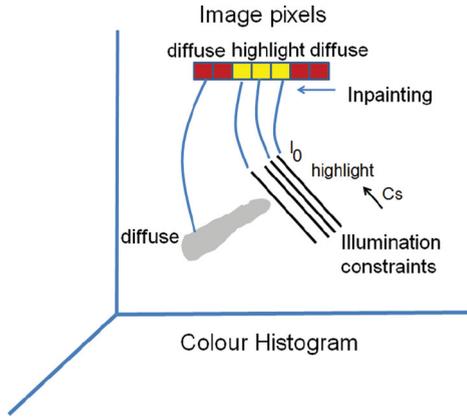
Figure 9: *A graphical representation of the illumination constraints. These constraints are used in the separation process. They define the position of the diffuse component of the specular pixels. (After Tan et al. [TLQS03].)*

Inpainting calculates the energy function E in the highlight region Ω

$$E = \int_{(x)\varepsilon\Omega} s(\|\nabla N(x)\|)$$
$$(\gamma(\nabla r(x))^2 + \gamma(\nabla g(x))^2 + (\nabla\|I_d(x)\|)^2)$$
$$+ (1 - s(\|\nabla N(x)\|))\|\nabla (I(x) - I_d(x))\| \, dx. \quad (42)$$

The energy function consists of two parts. The first part, governed by the stopping function s, is the part used outside texture edges. The second part, governed by the negated stopping function $1 - s$, is used across texture edges.

γ is a weight to tune the chromaticity smoothness. $N(x)$ is the normal direction of the plane defined by the origin and the illumination constraint line of (x), which passes through the image $I(x)$. $I_d(x)$ represents the diffuse component. It can be written, from Equation (18), as $I_d(x) = I(x) - m_s(x)\Gamma$, where Γ is the illuminant colour. The components r and g are the chromaticities. The illuminant colour is obtained by determining the value of Γ that results in the smallest energy. Once the illuminant colour has been found, the pixels are inpaintend by $I_d(x) = I(x) - m_s(x)\Gamma$. When the pixel to be painted is saturated, a standard TV inpainting is used [TLQS03].

In Figure 10 are shown the steps for an initial inpainting estimation that allows to reduce the computation time to achieve the minimum energy equation (42). This technique is not completely automatic because user indication of highlight regions is needed.
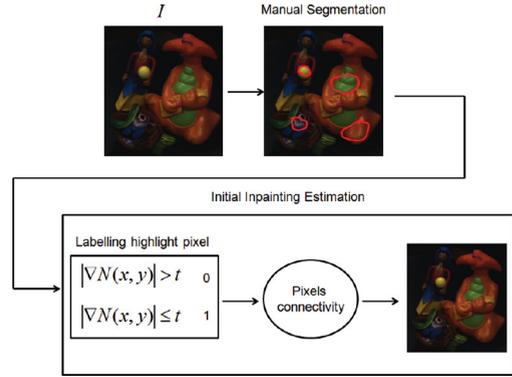
Figure 10: *Initial Inpainting estimation Tan et al. [TLQS03].*

3.2.4. *Separation of highlight reflections on textured surfaces*

Tan *et al.* [TLQ06] introduced a method based on a simple idea to overcome the limitations, discussed in Section 5 of the colour-space analysis and neighbour-based methods. The idea is based on the following concept: if a diffuse colour outside the highlight region is used as representative of the diffuse colour inside, diffuse colour textures can serve as exemplars of their spatial distribution. When no distinct texture data is identified for a pixel, the method does the separation in the same way as the conventional colour-space techniques.

How is the texture constraint formulated? Let p be a highlight pixel in a highlight region and p' a diffuse pixel candidate lying outside the highlight region and having a diffuse colour D. An $n \times n$ window centred on p is compared with the windows centred on all possible candidate diffuse pixels p' outside the highlight region. The texture distance between the windows of p and p' defines the texture constraint. This distance is computed as the average of the distances between their corresponding elements [TLQ06]

$$E(p, p') = \frac{1}{n^2} \sum_{k=1}^{n^2} \text{dist}[D_{p'}(k), I_p(k), \vec{S}], \quad (43)$$

where k is the pixel index within the window, and dist is the minimum angular distance between $D'_p(k)$ and the illumination constraint line defined by $I_p(k)$ and the illumination colour direction $\vec{S}$.

The texture constraint is applied in the following way (Figure 11). For each highlight pixel, starting with texture scale 1×1, a set of diffuse pixel candidates along the illumination constraint line is obtained. The ambiguity level of the candidates is measured as the maximum chromaticity distance between pairs of diffuse colours within the candidate set. If the ambiguity level is below the ambiguity threshold (thr1), the set is declared consistent. In this case, the candidate with the smallest angular distance is selected.
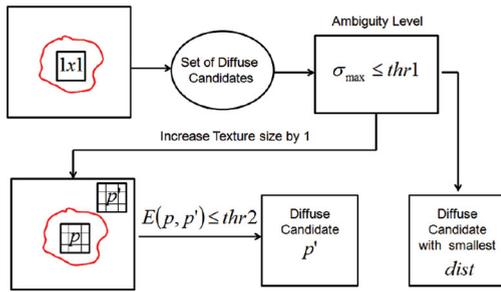
Figure 11: *Texture constraint application [TLQ06].*

In case the set of the possible candidates is not consistent, that is, the ambiguity level is above the ambiguity threshold (thr1), the texture scale (size) of the candidate windows is increased by 1. To determine if a diffuse image area is a match for the highlight pixel, the distance according to Equation (43) is calculated. If this distance is below a certain threshold (thr2), a match is found [TLQ06].

With this operation, the original set of diffuse pixel candidates is iteratively pruned, and the ambiguity computation is repeated with the new set. The method is not completely automatic as the candidate regions (specular and diffuse) are selected by the user.

3.2.5. *Use of colour information and classifier*

Tappen *et al.* [TFA05] decompose an image into shading and reflectance images by classifying each image derivative as being caused by a shading or a reflectance change. The assumption made here is that the input image I can be expressed as the product of the shading and reflectance images. In the logarithmic domain the derivative of the input image is the sum of the derivatives of these two images. To avoid the case when both shading and reflectance occur at the same point, the authors use the assumption that every image derivative is caused by either shading or reflectance. The problem is thus reduced to binary classification of image derivatives. The method of Weiss [Wei01] is then used to recover each intrinsic image from its derivatives. Two different kinds of classifiers were used to classify the derivatives, utilizing colour and grey-scale information, respectively. Using only colour information is not sufficiently robust and precise. Changes in colour intensity could be caused by either shading or reflectance variations [TFA05]. Using only local colour information, colour intensity changes can not be classified properly. Because the shading patterns have a unique appearance which can be discriminated from reflectance patterns. This allows the use of the local grey-scale information; so the classifier that uses grey-scale information yields more robust results.

After the derivative classification, some areas may still contain ambiguous local information. To solve this problem, a mechanism is used to propagate information from the areas where the classification is clear, into the areas where the local evidence is ambiguous. To propagate the evidence a Markov Random Field (or Markov Random Network) is applied, where each node is represented by a derivative with two possible states representing shading and reflectance changes [TFA05].

3.2.6. *Specular highlights detection based on fresnel reflection coefficient*

To simplify the separation of diffuse and specular components, a common practice is to approximate the colour of highlights for dielectric material with the colour of incident light. On the other hand, Equation (12) expresses the reflectance ratio as a function on wavelength, and the index of refraction is as well wavelength dependent. This implies that the colour of highlights and the colour of the incident light are not necessarily the same. Angelopoulou [Ang07] has shown that the Fresnel coefficient changes with wavelength, and this change can be significant for specularity detection techniques. Moreover, the dependence of the Fresnel term on the wavelength affects the colour of specular highlight making it distinct from that of the incident light. Based on this observation, the ability to extract the Fresnel coefficient allows to describe the colour of the specular highlight. The author has shown that for the specular reflection the spectral derivative is a measure of how the Fresnel term changes with wavelength. In fact, for specularities the spectral derivative of an image is primarily a function of the Fresnel term, and this is independent of the particulars of its derivation. The spectral gradient is obtained by computing the spectral derivative for a multispectral image, in a logarithmic scale for each colour band, and subtracting pairs of consecutive bands. This encodes the Fresnel term changes over a range of wavelengths, and it can be considered a good descriptor for the colour of specular highlights. In this way, an easy segmentation step can be performed based on the spectral gradient computation to a pixel level and making use of a high-dimensional mean-shift segmentation algorithm.

4. Multi-Image Methods

These techniques use information contained in an image sequence of the same scene taken either from different points of view or with different light information. Such sequence contains much more information on specularity than a single image since the specular reflection varies through the images. In the case of a sequence of images taken from different points of view, scene points showing specular reflection in a view can exhibit purely diffuse reflection in other views. By matching specular pixels to their corresponding diffuse points in other views, it is possible to determine the diffuse components of the specularities. Early work using images from different points of view was done by Lee *et al.* [LB92],

where a spectral differencing algorithm allows to avoid segmentation.

More recently, Lin *et al.* [LYK*03] presented a method based on colour analysis and multi-baseline stereo that makes use of sequence of images to achieve the separation of specular reflection. Lin *et al.* [LS01] presented a technique that makes use of two photometric images without calibrated lighting. This technique does not assume any dependencies among pixels, such as regionally uniform surface reflectance. Using recent work on the statistics on natural images, Weiss [Wei01] formulated the separation problem as a maximum likelihood estimation problem. Some techniques proposed in literature use colour and polarization [NFB97, KLHS02, USGG04, MPC*07]. In this case, the scene needs to be captured under different polarization orientations. These methods rely on the fact that the specular reflection component tends to be polarized.

Using the intensity histogram and a thresholding strategy, Chen *et al.* [CGS06] were able to reconstruct the specular field. They defined the gradient histogram as the difference between the histograms of two different intensities. Using the gradient threshold, the associated threshold for each intensity is then found and used for the threshold step. To reconstruct the whole specular field almost 200 images are required making the method unpractical. Lamond *et al.* [LPD07] have introduced a separation technique based on the work by Nayar *et al.* [NGR06]. As explained in [NGR06], the direct and indirect illumination can be separated by shifting high-frequency illumination and afterwards computing the maximum (peak) and minimum (average) pixel values, the specular reflection can be computed as difference of these two values. Finally, the diffuse reflection is computed as the difference between the full-lit image (summation of the all four patterns) and the specular image of the scene.

Moreover, the use of flash images have been explored to separate specular and diffuse reflections. In particular, Feris *et al.* [FRTT04] presented a method that makes use of a sequence of *n* images taken from the same point of view but with different flash position. Agrawal *et al.* [ARNL05] instead analyse the use of flash/no-flash images to achieve the reflections separation goal.

4.1. Diffuse-specular separation from image sequences

Using multi-baseline stereo, Lin *et al.* [LYK*03] defined correspondence constraints between specular and diffuse pixels based on the following assumptions: diffuse reflection satisfies the Lambertian property; specular reflection varies in colour from view to view; scene points having specular reflection exhibit purely diffuse reflection in some other views. The algorithm can be summarized in two steps: the identification of the specular pixels and finding stereo correspondences for the specular pixels.
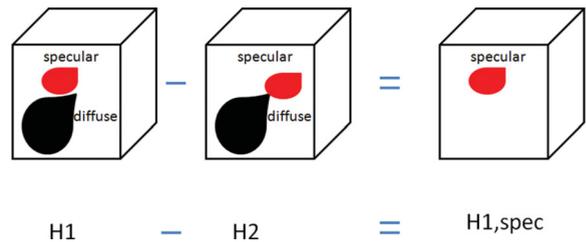
Figure 12: *A graphical representation of the operation of the colour histogram difference (CHD). H_1 and H_2 are colour histograms for different views. The colour of the specular reflection is changing from view 1 to view 2, and this helps to localize the specular regions as the difference of the two histograms. (After [LYK*03].)*

To identify the specular pixels, the colour histogram difference (CHD) is used in the context of multi-baseline stereo. This is based on colour changes of specular reflection from view to view.

As is illustrated in Figure 12, these view-dependent colour variations of specular reflection can be detected by histogram differencing [LYK*03]. Two major obstacles may arise when using the standard CHD: histogram clutter and colour occlusion resulting in diffuse pixels detected as specular. Lin *et al.* [LL97], integrated the use of CHD with polarization to reduce the correlation of specularities between views. To solve these problems, the standard CHD has been modified [LYK*03] to make use of the Epipolar geometry, which allows to reduce the number of histogram points by differencing in image rows (scanlines) rather than the entire image. The colour occlusion problem is solved using three images (left I_L, central I_C, right I_r) to obtain a *tri-view* CHD. In this way, if a diffuse reflection in a scanline of the central image I_C is geometrically or specularly occluded in either I_L or I_r, it is still present in the combined histogram $H_L \cup H_r$. The set of the histogram points in H_C that contains specular reflections is easily determined. These points are back-projected to the image to locate the specular pixels. They can be represented as a binary mask image S_C, where 0 indicates a diffuse pixel and 1 a specular pixel.

Using a longer stereo sequence, it is possible to obtain a more robust detection procedure called *multi-CHD*. For the central reference image I_C, it is possible to form *n* different uniform-interval triplets (I_L^k, I_C^k, I_R^k) for $k = 1, 2, \ldots, n$. The tri-view CHD can be computed on each of these triplets producing *n* specular point sets S_C^k. These sets are used to vote for the final detection results. A pixel is selected as specular if the number of votes from $S_{C,k}$ exceeds a given threshold [LYK*03].

Once the specular pixels have been detected in an image, the separation is done by associating them to their corresponding diffuse points in other images. The

correspondences cannot be computed directly because of large differences in intensity and colour. However, they can be obtained under the constraint of their disparity relationship to the specular pixel in the reference image [LYK*03]. Once the correspondence between the specular pixel P_S and the diffuse pixel P_D has been obtained, the separation process is simple. Often, it is necessary to take into account noise and mismatches. In this case, for each P_S all possible corresponding diffuse points in the data set are found, then the value of P_D is obtained as the average mean of the corresponding colour measurements.

4.2. Deriving intrinsic images from image sequences

As discussed by Weiss [Wei01], the relationship between the input image $I(x)$, the reflectance image $R(x)$ and the luminance image $L(x)$ can be written as

$$I(x) = L(x)R(x). \tag{44}$$

The author focuses on a problem that is easier than the complete separation problem: Given a sequence of K images where the reflectance is constant over time, and only the illumination changes, recover K illumination images and a single reflectance image. In the logarithmic domain, the logarithm of the input image, $i(x)$, is the sum of the logarithm of the illumination, $l(x)$, and the logarithm of the reflectance, $r(x)$. Weiss used a robust property of natural images, namely, that when derivative filters are applied to such images, the filter outputs tend to be sparse. This observation is valid for a wide range of natural images [OF96, Sim97].

For two different images the histograms of the horizontal derivative filter outputs have similar shapes closely resembling the Laplacian distribution [OF96, Sim97]. Using this observation, Weiss formulates the separation problem as a maximum-likelihood (ML) estimation problem. For a bank of N filters f_n, denote the sequence of output images by $o_n(x, k) = i * f_n$, where $*$ is the convolution operation and $k = 1, 2, \ldots, K$. The reflectance image filtered by the nth filter will be $r_n = r * f_n$. The ML estimation gives the ML estimate of the filtered reflectance image $\hat{r}_n$. In [Wei01], this estimate is shown to be the median of the filter outputs.

The logarithmic reflectance image $r(x)$ is then recovered by simply solving a linear problem using a pseudo inverse solution. Once $r(x)$ has been estimated, an estimate of $l(x)$ is easy to obtain.

4.3. Polarization

Following the model in Section 2.5, the separation problem can be viewed as that of solving a linear system [NFB97]

$$I_i = f_i v, \tag{45}$$

where f_i is the ith filter setting and v the unknown vector:

Figure 13: *Decomposition of $I_{\min}$ in each pixel into I^g and $I^l_{\min}$ in RGB colour space. Image courtesy of Steven Lin [KLHS02].*

$$v = (I_C, I_{sv} \cos 2\alpha, I_{sv} \sin 2\alpha) \tag{46}$$

with $I_C = I_d + I_{sc}$.

From the above equation it is easy to see how to determine I_c, I_{sv} and α applying three different filters. Now it is necessary to separate the diffuse component from the specular constant term I_{sc} contained in I_C. The minimum and maximum values of the image brightness are determined as

$$I_{\min} = I_C - I_{sv}, \tag{47}$$

$$I_{\max} = I_C + I_{sv}, \tag{48}$$

and the degree of polarization ρ is measured as

$$\rho = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}, \tag{49}$$

ρ can be used to classify points into those that are only diffuse and those that contain a specular component. In the latter case, further processing is necessary to extract the specular component. This procedure uses constraints on the reflection component in each image point [NFB97].

As an extension of the work in [NFB97], Kim *et al.* [KLHS02] presented an approach to separate the specular reflection. Based on Equation (13), in each pixel they divide the colour space into two subspaces: the specular line space and the diffuse plane space, as illustrated in Figure 13 .

It can be observed in Figure 13 that, as predicted by Equation (13), the colour of a specular pixel seen through different polarized angles will lie on a line in the *RGB* space (line space). The line space is parallel to the constraint line of polarization in each pixel, whereas the diffuse plane space is perpendicular to the specular direction. For a user-defined threshold t, a pixel is considered to be specular when

$$I_{\max} - I_{\min} = 2I_{sv} > t. \tag{50}$$

The specular constraint line in Equation (13) can be written in parametric form $L(\text{p})$ (Figure 13). The half line space I^l is the specular direction line and I^g is a vector in the plane

orthogonal to I^1 and $L(p).I^g$ is a constant vector independent of the polarization angle θ in each pixel. To separate the specular reflection from I_{min}, it is necessary to locate p_d

$$p_d = \left| I^1_{min} \right| - \left| I^1_d \right| = |I_{min} - I_d|. \qquad (51)$$

This allows to determine I^1_d, then specular and diffuse reflections can be separated using simple vector calculations. To locate I^1_d, the I^1 space is employed with the initial value of I^1_{min} at each pixel.

An energy function is used to smooth the spatial variation of the specular component [KLHS02]. The direction of smoothness is controlled by gradient information in the diffuse plane. The narrow dynamic range of the camera can cause saturation in some areas of the input image and can generate erroneous line constraints. This problem is solved using an inpainting technique [CS01].

Umeyama *et al.* [USGG04] make use of polarization and statistical analysis of images. As explained earlier in this section, an image obtained through a polarizer is a linear sum of diffuse and specular reflection components. The coefficient of the specular component depends on orientation of the polarizer, the geometric configuration of light source, object and camera. As observed in [USGG04], if one of these parameters is unmeasurable, the value of the coefficient is also unknown. The problem of separating diffuse and specular reflections requires isolating signals from mixture of signals. The authors proposed to use the *Independent Component Analysis* to separate the two types of reflections.

Ma *et al.* [MPC*07] make use of linear (similar to [DHT*00]) and circular polarization filters placed on the camera, blocking all the specularly reflected light, and on the light sources. This will held an image that is the linear combination of diffuse and specular reflections. Placing the polarization filter on the camera, is equivalent to acquiring an image as $I_1 = 1/2I_D$ free of the specular component. Placing the polarization filter on the light sources, is equivalent to acquiring an image as $I_2 = 1/2I_D + I_S$, in case of linear polarization, and as $I_2 = 1/2(I_D + I_S)$, in case of circular polarization. Then, with simple arithmetic operations, the specular and diffuse reflections can be easily separated.

4.4. Multi-flash methods

These methods rely on the observation that the shift of the highlights in a sequence of images, taken from fixed viewpoint with varying light source positions, is generated by the shift of the light source that produces the highlights. Figure 14 shows a diagram that illustrates the technique by Feris *et al.* [FRTT04]. Given an image sequence taken by a fixed camera for varying flash-light positions, an image with reduced highlights is built; then, making use of the I_{max} the final output image is improved during the matting process.
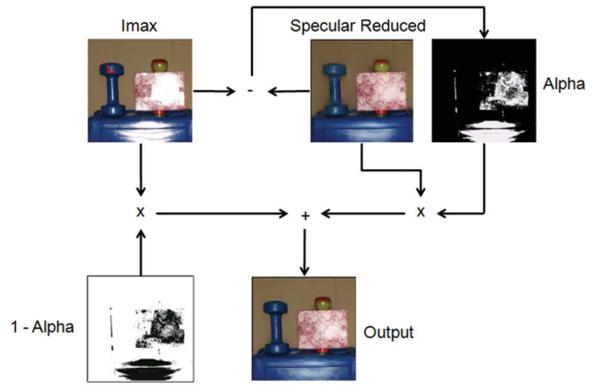
Figure 14: *Diagram of the multi-flash method presented by Feris et al. [FRTT04].*
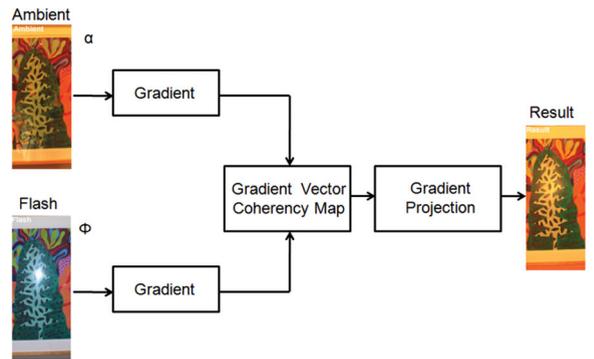
Figure 15: *Diagram of the multi-flash method presented by Agrawal et al. [ARNL05].*

I_{max} is computed as the maximum composite of the input images.

The median of the intensity gradients of the individual images is used to reconstruct an image with the specularity reduced compared to the original sequence. A Poisson solver is applied for this step. The matting process uses the difference between the reconstructed image and the I_{max} (as alpha channel) to replace the specular regions of I_{max} by the corresponding specularity-reduced regions in the reconstructed image. If the highlights do not move among images, the method cannot remove them [FRTT04].

In Figure 15 is depicted the diagram that describes the technique by Agrawal *et al.* [ARNL05].

The authors use the flash imaging model, as described in Section 2.6, Equation (14), to capture the illumination-invariant parts of the input image. A coherence model is applied based on the orientation of the image gradient vector. As observed by the authors, the gradient orientation in the flash image should be coherent with the gradient

orientation in the ambient image. The only exception are the regions with flash artefacts and ambient shadow edges. The method calculates the gradient magnitude and orientation for the two images (flash and ambient), then a gradient orientation coherency map is obtained as

$$M = \frac{|\nabla\phi \cdot \nabla\alpha|}{(|\nabla\phi||\nabla\alpha|)}, \qquad (52)$$

where M encodes the angular similarity between the two gradient images at pixel level. On the other hand, the two gradient magnitudes are different, and they are related by an unknown scalar k

$$\nabla\phi = k\nabla\alpha. \qquad (53)$$

In an ideal case no artefacts appear in the ambient and flash images, but in a typical case artefacts may appear in both images. This can be modeled as unknown noise, and the new gradient is computed as

$$\nabla\alpha' = \nabla\alpha + \nabla\eta_A, \qquad (54)$$

where η_A is the unknown noise in the ambient image. For the flash image, the new gradient is

$$\nabla\phi' = \nabla\phi + \nabla\eta_F = k\nabla\alpha + \nabla\eta_F. \qquad (55)$$

where η_F is the unknown noise in the flash image. This decomposition of the flash image ϕ is an ill-posed problem. It is solved by analysing the gradient coherence and using gradient projections to compute the new gradient field for the flash image corrupted by the unknown noise η_F.

5. Discussion

In this section we present a detailed comparison of the techniques surveyed in this paper, highlighting their advantages and disadvantages.

5.1. Characteristics analysis

Several factors may influence which approach is superior to another one, such as the number of images to be captured, automatic operation versus manual help, light constraints and the reflectance model used, merits of quality, and the hardware used during the acquisition phase. Table 2 provides a comparative summary of the specular removal methods discussed in this survey. Later, we discuss each of the factors in more detail.

5.1.1. *Number of input images*

Comparing the single-image techniques to the multiple-image ones, we observe that the latter can be limited by the high number of input images needed to achieve satisfactory results. However, some of these methods need only a limited number of input images, and this may be a good compromise between feasibility and quality (e.g. polarization [NFB97, KLHS02, MPC*07, USGG04], high low frequency separation [LPD07, NGR06] and the method by multi-flash [FRTT04, ARNL05]). Requiring large number of input images and/or using a special hardware reduce the feasibility of the method making it complicated to implement (to reproduce the setup). Also, it may increase the acquisition time which is limited in certain applications. The single-image methods need just one image which can make them more attractive.

Table 2: *Comparison of the highlight removal techniques by their characteristics.*

Technique	Images	User interaction	Light requirement	Hardware
Colour Space [KSK87, KSK88, ST95b, SK00, BLL96]	1	Manual segmentation	IC–DM	Single camera
Specular-Free Image [TI05b, YCK06, SC09]	1	Automatic	IC–DM	Single camera
Inpainting [TLQ06]	1	Manual segmentation	IC–DM	Single camera
PDE [MZK*06]	1	–	IC–DM	Single camera
Textured Surfaces [TLQ06]	1	Manual segmentation	IC–DM	Single camera
Colour Classifier [TFA05]	1	–	–	Single camera
Fresnel Coefficient [Ang07]	1	Manual segmentation	No IC–DM	Single camera
Multi-Baseline Stereo [LB92, LYK*03, LS01]	50–70	No segmentation	DM	Multiple cameras
Deriving Intrinsic Images [Wei01]	40–70	No segmentation	Illumination changes	Single camera
Polarization [NFB97, KLHS02, MPC*07, USGG04]	6–10	No segmentation	DM	Polarized filters
Histogram [CGS06]	200	No segmentation	–	Single camera
High-Low Frequency Separation [LPD07, NGR06]	4, 32	No segmentation	–	Single camera
Multi-flash [FRTT04, ARNL05]	4–8, 2	No segmentation	FM	Flash system

The methods are grouped into the single-image category (top) and the multiple-image category (bottom). IC, illuminant compensation; DM, dichromatic reflectance model and FM, flash model.

5.1.2. *Hardware and user interaction*

Typical problems in the setup of multiple-image methods are as follows. In polarization techniques, the spatial shift between images due to the motion of polarization filter, the chromatic abberation of lenses and the camera noise of charge-coupled device sensors can generate errors in polarization fitting, which leads to artefacts in scene boundaries. In multi-baseline methods, the positions of the cameras in the acquisition step play an important role in the ability to detect specular regions. Specular colour may not vary substantially for very small intervals, while visibility may often vary for large baselines. This makes it difficult to build a universal setup that can be used in many different cases; this is even more difficult for data acquisition outside laboratory.

In some cases, user intervention is not considered as a weak point. It is mainly used for segmentation step in the neighbour-based and colour space methods. Accurate segmentation is hard to achieve automatically when specularities are present. When doing this step manually may help increase the selection precision for the regions to be processed. A typical example is the colour space analysis techniques discussed in Section 3.1. These techniques need to group each highlight pixel with non-highlight pixels having the same diffuse colour. This is a difficult task because one can have colour-blended pixels at texture boundaries and texture colours mixed with highlight components. Also, it often happens that some highlights do not have corresponding diffuse pixels outside of the highlight region. Polarization methods can also be difficult to tune due to the large number of parameters, thresholds, which may influence the ability to separate the reflection components. Selecting the thresholds must be done carefully because a bad threshold may label specular pixel as diffuse and may cause other pixels to be labelled wrongly as well. This may also influence the computation time, which usually happens for scenes with very large highlights in planar regions [NFB97].

5.1.3. *Light constraints*

The light constraints and the reflection model used by a separation technique play a major role in the ability of detecting specularities in the input image. The assumptions made by the reflection model may not be satisfied in some cases making the technique unable to detect the specularities regions properly. One typical assumption is that the spectral distribution of the incident light is the same as that of the highlights. However, Angelopoolou [Ang07] has shown that the error introduced by this approximation is not negligible. This is a typical assumption of most of the methods presented in this survey and based on the dichromatic reflection model. For example, this applies to the neighbour-based methods presented in Sections 3.2.1 and 3.2.3 as well as the colour-space analysis described in Section 3.1. In the neighbour-based

methods, the diffuse information is mainly propagated from outside to inside a highlight. Such approach may face problems for discontinuities in surface colours, across which diffuse information cannot be accurately transferred [TLQ06]. As discussed in Section 3.2.1, the main advantage of these techniques is their simplicity and fast computation. On the other hand, for the above reasons the original surface colour is not preserved, and this can lead to problems when it is essential to preserve colour. The methods based on colour-space analysis (Section 3.1) can suffer from ambiguity that arises during the separation process. The main idea of these techniques is to project the highlight colour along the illumination colour direction onto a point having the same diffuse component.

Often, the projection point becomes ambiguous because several possible points can be selected while only one is correct. The factors that cause this ambiguity are image noise, colour blending at edges, roughnesses and highly textured surface, as discussed in the study of Tan *et al.* [TLQ06].

For these reasons, the earlier methods have been presented only for smooth, textureless surfaces. Extending these methods to rougher and textured surfaces would require segmentation of the surface into different diffuse colours, the existence of purely diffuse counterparts for each highlight pixel, and a different approach to illumination colour estimation. Image geometry and material roughness can make the highlight cluster skewed, which makes it difficult to estimate the illuminant colour by vector fitting. Noise and multiple diffuse colours can make it difficult to obtain the colour distribution, which affects the derivation of the geometry properties and surface roughness from the histogram shape. Methods to estimate the illuminant colour have been developed, but they are sensitive to noise and often are required assumptions that are not acceptable in practice of highlight removal applications. Moreover, the estimation cannot be made for a highlight that lies on a surface of a uniform colour.

The assumption that the illumination colour is uniform throughout the highlight is also not always true. This assumption is valid only when the highlight is generated by a single-source light; however, it can be generated by multiple interreflections making the illuminant colours non-uniform.

To resolve the ambiguities arising from illumination and surface characteristics, one has to impose constraints that can be additionally used in the separation problem. These constraints can be illumination-based, such as those proposed by the inpainting techniques (Section 3.2.3). This helps better recover textures obscured by highlights, which are distorted or eliminated by traditional methods, and improve shading in the area where the diffuse intensities within the highlight exceed the diffuse shading at the borders [TLQS03]. A disadvantage with these methods is that they need segmentation to resolve ambiguity of texture with the same appearance as the highlights. In the case of texture-based techniques,

the constraints are based on the higher-order colour data (texture data) in the form of spatial colour distribution (Section 3.2.4). This use of spatial information in the image does not suffer from surface colour discontinuity problems that is characteristic of the neighbour-based and colour-space analysis approaches. In employing texture-based constraints, an issue that may arise is texture variation, where some local diffuse texture within a highlight area may not be fully consistent with outside textures. This is solved using traditional separation methods for this type of situations.

The PDE approaches (Section 3.2.2) also rely on the dichromatic model. During the erosion process, the pixels that violate the dichromatic model are treated as outliers. The resulting artefacts can be reduced using a post-processing inpainting technique.

The primary limitation of the colour classifier technique (Section 3.2.5) is related to the fact that the classifier needs to incorporate knowledge about the surface structure as well as the appearance when illuminated. This relates to the choice of the training set of images that show reflectance and shading changes. Choosing a training set that includes enough information may be difficult.

If a proper set of images, for the training set, is not chosen artefacts may be accentuated.

In case of the polarization methods, if the data are not consistent with the cosine model Equation (13) it is necessary to consider outlier pixels that have a root-square-mean error (RMSE) above a certain threshold in the polarization parameter fitting problem. Polarization approaches may misidentify specularities in diffuse reflection areas near occluding boundaries. This is due to the fact that in these regions the surface normals are nearly perpendicular to the viewing directions, and in this situation the diffuse reflection exhibits polarization [LL97]. As for the colour space analysis techniques, similarly to polarization it is necessary to impose constraints to provide consistency among the examined pixels.

The assumption that diffuse reflection follows the Lambertian model may not be valid for some types of surfaces/materials, which is a typical problem for stereo matching. Two typical assumptions of the multi-baseline methods are that the specular colour changes among viewpoints, and the specular points are moving with the views. In areas with high curvature and large roughness, it may happen that the highlights do not shift enough between the views. In this case the two assumptions are violated. This problem can also be typical for multi-flash methods. Other characteristic problems of multi-baseline methods are colour saturation, image noise and colour blending. These problems may reduce the accuracy in the detection of the specular regions. The CHD performs well in detecting specularities for surface with spatially varying colour under complex illumination, not requiring any segmentation. However, the CHD-based methods can face problems when the ambient and the direct illumination

colours approximately match the surface colour, or when they have the same spectral distribution and the surface colour is neutral. This is because in this case the specular reflections may move insignificantly in colour space between two views, and as a consequence the specular reflection is misidentified. Some of these problems have been solved by either integrating the CHD with polarization [LL97] or extending the traditional CHD to the *multi-CHD* [LYK*03].

5.2. Comparison of images

From the discussion presented in the previous section, one can see that several techniques have constraints that limit their usability in real applications. For example, the applicability of many of the multi-images approaches is limited by the large number of input images required and the use of specific hardware and careful setup. Concerning the single image approaches, it is clear that the colour space analysis approaches fails when highlights are on highly textured or rough surfaces. The presence of noise and colour blending at edges leads to the ambiguity problem described in Section 5.1. Also, in the single-image category, the Colour Classifier techniques [TFA05] may result in visible artefacts such as colour degradation, lost of texture and edge information when compared with the original input image. Later, we will present the results of an image comparison test for the techniques discussed in this survey. However, based on these observations, we have excluded from this comparison the techniques earlier.

The experiment is performed on various input images starting from single-colour, multi-colour and textured surfaces, increasing the texture complexity.

We start with comparing the techniques that require a single image as input. Figure 16 shows a comparison of two techniques that make use of the concept of free-specular image as introduced in Section 3.2.1.

These techniques work well on images with single-colour and multi-colour surfaces, but they may face problems for images with highly textured surfaces. In Figure 16, second row, the diffuse output image obtained with the Yoon *et al.* [YCK06] technique shows visible artefacts. Also, when the brightness of the input images is not proportional to the flux of the incoming light of the camera, these techniques have difficulties in separating the highlights [Figure 16, 3rd row (b) and (c)].

Techniques that try to be robust in case of highly textured surfaces are the PDE [MZK*06] and inpainting [TLQS03]. In Figure 17, we compare them using an input image with a complex texture (PDE 1st row, inpainting 2nd row). The two techniques provide similar results. However, when the input is a complex texture pattern, such as the one in Figure 18, the inpainting technique cannot propagate through efficiently.
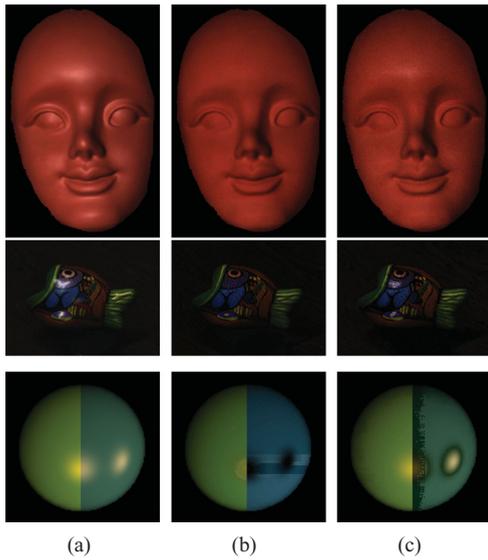
Figure 16: *R. Tan et al. [TI05b] vs. Yoon et al. [YCK06]:(a) Input images, (b) and (c) diffuse output images obtained by Tan et al. [TI05b] and Yoo et al. [YCK06] technique respectively. Input Images courtesy of R. Tan.*
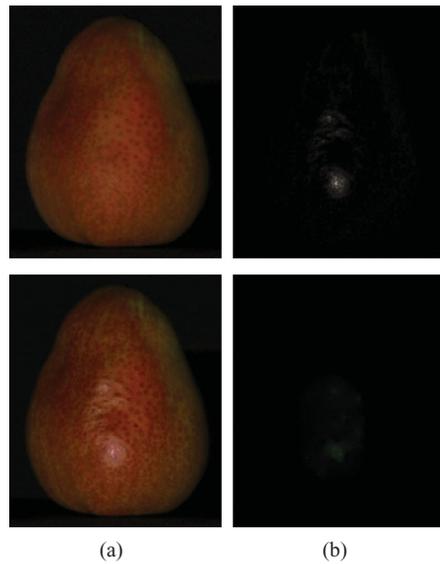
Figure 18: *Comparison of the PDE [MZK*06] (1st row) and inpainting [TLQS03] (2nd row) techniques: (a) diffuse component and (b) specular component. Images courtesy of Mallick and P. Tan.*

Figure 19: *Results by Feris et al. [FRTT04]. (a–d) The four input images; (e) the output diffuse image.*

The highlights are not completely removed generating visible artefacts [Figure 18, 2nd row (a)].

Among the multi-image approaches, the most reliable techniques, in terms of the reduced number of input images and hardware setup required, are the Multi-Flash techniques. As example of these techniques, we show in Figure 19 a sequence of four input images used by the technique Figure 19(a–d) and the output diffuse image Figure 19(e).

Particular attention must be given to the way the input images are taken, because this will influence the ability to remove the highlights. A few examples when the Feris *et al.* [FRTT04] technique fails are shown in Figure 20. In

Figure 17: *Comparison of PDE [MZK*06] (1st row) and inpainting [TLQS03] (2nd row): (a) Diffuse component and (b) Specular component. Images courtesy of S. P. Mallick and P. Tan.*

(a)   (b)

Figure 20: *Examples when the technique by Feris et al. [FRTT04] fails. (a) Input and (b) diffuse output image.*

(a)   (b)   (c)
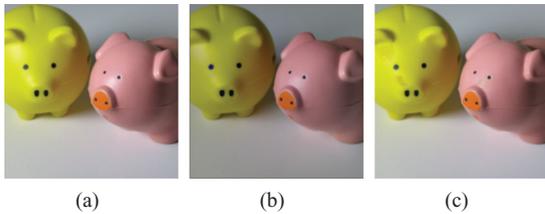
Figure 21: *Comparison of a multi-image technique to single-image one: (a) input image; (b) result by Feris et al. [FRTT04]; (c) result by the inpainting technique [TLQS03].*

this case, the visible artefacts are due to the fact that the highlights are not completely removed.

Finally, we compared the technique by Feris *et al.* [FRTT04] to the inpainting technique [TLQS03]. Figure 21 (c) demonstrates that, when the illuminant colour is badly estimated, the single-image inpainting method [TLQS03] is uncapable to completely remove the highlights.

6. Conclusions

In this paper, we have reviewed the specular removal methods and presented a useful classification for convenient selection of a proper method for a specific application. Specular reflection is inevitable due to the characteristics of materials existing in the real world. Despite this, approaches in many applications rely on a pure diffuse reflection model. Such approaches can completely fail when facing the specular reflection. So far, specularity removal methods have been used to improve the photo-consistency based image-to-surface registration as well as for 3D model reconstruction of surfaces with specular reflection component. These methods can be also used in photo editing to produce special effects, for object recognition and tracking and, in general, in image processing applications where it is desirable to remove the specular component. Although the currently available methods achieve good component separation results, they are limited by the conditions of their applicability. In particular, most of the techniques rely on a specific reflection model and assume that the specular reflectance varies insignificantly with wavelength, which means that its colour is essentially the same as that of the light source. This, together with their noise sensitivity, reduce the range of applications where the current methods can be used. A more general and robust method that overcomes the limitations of the current methods is highly requested. The recent physics-based method by Angelopoulou [Ang07], which does not use the dichromatic model and does not assume the wavelength independence, is a step in this direction.

Acknowledgements

References

[Ang07] ANGELOPOULOU E.: Specular highlight detection based on the fresnel reflection coefficient. In *Proceeding IEEE International Conference on Computer Vision ICCV* (Rio de Janeiro, Brazil, 2007), pp. 1–8.

[ARNL05] AGRAWAL A., RASKAR R., NAYAR S. K., LI Y.: Removing photography artifacts using gradient projection and flash-exposure sampling. *ACM Transactions on Graphics 24*, 3 (2005), 828–835.

[BB88] BLAKE A., BRELSTAFF G.: Geometry from specularities. In *Proceeding of the IEEE International Conference on Computer Vision ICCV* (Bombay, India, 1988), pp. 394–403.

[BLL96] Bajcsy R., Lee S. W., Leonardis A.: Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation. *International Journal of Computer Vision 17*, 3 (1996), 241–272.

[BT78] Barrow H., Tanenbaum J.: Recovering intrinsic scene characteristic from images. In *Proceedings of the Computer Vision System* (1978), pp. 3–26.

[CGS06] Chen T., Goesele M., Seidel H.: Mesostructure from specularity. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (New York, USA, 2006).

[CS01] Chan T., Shen J.: Mathematical models for non-texture inpainting. *Journal of Applied Math SIAM 62* (2001), 1019–1043.

[CT82] Cook R. L., Torrance K. E.: A reflectance model for computer graphics. *ACM Transaction on Graphics 1* (1982), 7–24.

[DCC*10] Dellepiane M., Callieri M., Corsini M., Cignoni P., Scopigno R.: Improved color acquisition and mapping on 3D models via flash-based photography *Journal on Computing and Cultural Heritage 2* (2010), 9, DOI: 10.1145/1709091.1709092.

[DHT*00] Debevec P., Hawkins T., Tchou C., Duiker H. P., Sarokin W., Sagar M.: Acquiring the reflectance field of a human face. In *Proceedings of SIGGRAPH* (New Orleans, Louisiana, USA, 2000).

[FRTT04] Feris R., Raskar R., Tan K.-H., Turk M.: Specular reflection reduction with multi-flash imaging. In *Proceedings of the IEEE Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI04)* (Brazil, 2004), vol. 42, pp. 316–321.

[HB88] Healey G., Binford T. O.: Local shape from specularity. In *Proceeding of the CVGIP* (1988), vol. 42, pp. 62–86.

[Hec98] Hechts E.: *Optics*, 3rd edition. Addison Wesley Longman, 1998.

[KLHS02] Kim D., Lin S., Hong K., Shum H.: Variational specular separation using color and polarization. In *Proceedings of the IAPR Workshop on Machine Vision Applications* (2002), pp. 176–179.

[KSK87] Klinker G. J., Shafer S. A., Kannade T.: Using a color reflection model to separate highlights from object color. In *Proceedings of the 1st International Conference on Computer Vision, IEEE London* (1987), pp. 145–150.

[KSK88] Klinker G. J., Shafer S. A., Kannade T.: The measurement of highlights in color images. *International Journal of Computer Vision 2* (1988), 7–32.

[KSK90] Klinker G. J., Shafer S. A., Kannade T.: A physical approach to color image understanding. *International Journal of Computer Vision 4* (1990), 7–38.

[LB92] Lee S. W., Bajcsy R.: Detection of specularity using color and multiple views. In *ECCV '92: Proceedings of the Second European Conference on Computer Vision* (London, UK, 1992), Springer-Verlag, pp. 99–114.

[LL97] Lin S., Lee S. W.: Detection of specularity using stereo in color and polarization space. *Computer Vision Image Understanding 65*, 2 (1997), 336–346.

[LM06] Li Y., Ma L.: Metal highlight spots removal based on multi-light-sources and total variation inpainting. In *Proceedings of the ACM International Conference on Virtual Reality Continuum and its Applications VRCIA* (New York, NY, USA, 2006), ACM Press, New York, NY, pp. 323–326.

[LPD07] Lamond B., Peers P., Debevec P.: *Fast image-based separation of diffuse and specular reflections.* Tech. Rep., 2007.

[LS01] Lin S., Shum H.-Y.: Separation of diffuse and specular reflection in color images. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Kauai, Hi, USA, 2001), vol. 1, pp. 341–346.

[LYK*03] Lin S., Yuanzhen S. L., Kang S. B., Tong X., Yeung Shum H.: Diffuse-specular separation and depth recovery from image sequences. In *Proceedings of European Conference on Computer Vision (ECCV)* (2003), pp. 210–224.

[MPC*07] Ma W. C., Peers P., Chabert C. F., Weiss M., Debevec P.: Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination. In *Proceedings of the Eurographics Symposium on Rendering (EGSR)* (Inria, Grenoble, France, 2007).

[MZBK06] Mallick S. P., Zickler T., Belhumeur P. N., Kriegman D. J.: Dichromatic separation: specularity removal and editing. In *Proceedings of the SIGGRAPH Sketches* (Boston, USA, 2006).

[MZK*06] Mallick S. P., Zickler T., Kriegman D. J., Belhumeur P. N.: Specularity removal in images and videos: A PDE approach. In *Proceedings of the European Conference on Computer Vision* (Graz, Austria, 2006), pp. 550–563.

[NFB97] Nayar S. K., Fang X., Boult T.: Separation of reflection components using color and polarization. *International Journal of Computer Vision 21* (1997), 163–186.

[NGR06] Nayar S. K., Grossberg G. K. M. D., Raskar R.: Fast separation of direct and global components of a scene

using high frequency illumination. *ACM Transaction on Graphics 25*, 3 (2006), 935–944.

[NS92] NOVAK C., SHAFER S.: Anatomy of a color histogram. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* (1992), pp. 599–605.

[OF96] OLSHAUSEN B. A., FIELD D. J.: Emergence of simple cell receptive field properties by learning a sparse code for natural image. *Nature 381* (1996), 607–608.

[OJR03] OSADCHY M., JACOBS D., RAMAMOORTHI R.: Using specularities for recognition. In *Proceeding IEEE International Conference on Computer Vision ICCV* (Nice, France, 2003), pp. 1512–1519.

[OTS*05] ORTIS F., TORRES F., SAMEER S., MANEESHA S., CHID A., PERNER P.: A new inpainting method for highlights elimination by colour morphology. In *Pattern Recognition and Image Analysis* (2005), vol. 3687/2005, Spring-Verlag, Berlin/Heidelberg, pp. 368–376.

[Pav77] PAVLIDIS T.: *Structural Pattern Recognition*. Spring-Verlag, Berlin, Heidelberg, New York, 1977.

[SC09] SHEN H. L., CAI Q. Y.: Simple and efficient method for specularity removal in a image. *Applied Optics 48*, 14 (2009), 2711–2719.

[Sha85] SHAFER S. A.: Using color to separate reflection components. *Color Research and Applications 10*, 4 (1985), 210–218.

[Sim97] SIMONCELLI E.: Statistical models for images: compression, restoration and synthesis. In *Proceedings of the Asilomar Conference on Signals, Systems and Computers* (1997), pp. 673–678.

[SK00] SCHLÜNS K., KOSCHAN A.: Global and local highlight analysis in color images. In *Proceedings of the 1st International Conference on Color in Graphics and Image Processing (CGIP)* (2000).

[ST95a] SCHLÜNS K., TESCHNER M.: Analysis of 2d color spaces for highlight elimination in 3d shape reconstruction. In *Proceedings of the Asian Conference on Computer Vision II* (1995), pp. 801–805.

[ST95b] SCHLÜNS K., TESCHNER M.: Fast separation of reflection components and its application in 3d shape recovery. In *Proceedings 3rd Color Imaging Conference* (1995), pp. 48–51.

[SZSX08] SHEN H. L., ZHANG H. G., SHAO S. J., XIN J. H.: Chromaticity separation of reflection component in single images. *Pattern Recognition 41* (2008), 2461–2469.

[TFA05] TAPPEN M. F., FREEMAN W. T., ADELSON E. H.: Recovering intrinsic images from a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence 27* (2005), 1459–1472.

[TI03] TAN R., IKEUCHI K.: Estimating Chromaticity of Multicolored Illuminations. In *Proceeding of the IEEE International Workshop on Color and Photometric Methods in Computer Vision* (2003).

[TI04] TAN R., IKEUCHI K.: Intrinsic properties of an image with highlights. In *Meeting on Image Recognition and Understanding (MIRU)* (2004).

[TI05a] TAN R., IKEUCHI K.: Illumination color and intrinsic surface properties physical-based color analysis from a single image. *Transactions of Information Processing Society of Japan 46* (2005), 17–40.

[TI05b] TAN R., IKEUCHI K.: Separating reflection components of textured surfaces using a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence 27*, 2 (2005), 178–193.

[TLQ06] TAN P., LIN S., QUAN L.: Separation of highlight reflections on texture surfaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (New York, USA, 2006), pp. 1855–1860.

[TLQS03] TAN P., LIN S., QUAN L., SHUM H. Y.: Highlight removal by illumination-constrained inpainting. In *Proceeding IEEE International Conference on Computer Vision (ICCV)* (Nice, France, 2003), vol. 1, pp. 164–169.

[TNI03] TAN R., NISHINO K., IKEUCHI K.: Illuminant chromaticity estimation using inverse-intensity chromaticity space. In *Proceedings of the IEEE Computer Vision and Pattern Recognition Conference* (Wisconsin, USA, 2003), pp. 18–20.

[TNI04] TAN R., NISHINO K., IKEUCHI K.: Separation of reflection components based on chromaticity and noise analysis. *IEEE Transaction on Pattern Analysis and Machine Intelligence PAMI 10* (2004), 1373–1379.

[USGG04] UMEYAMA, S., GODIN, G.: Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images. *IEEE Transaction on Pattern Analysis and Machine Intelligence 26*, 5 (2004), 639–647.

[Wei01] WEISS Y.: Deriving intrinsic images from image sequences. In *Proceeding IEEE International Conference on Computer Vision (ICCV)* (Vancouver, Canada, 2001), vol. 2, pp. 68–75.

[YCK06] YOON K. J., CHOI Y., KWEON I. S.: Fast separation of reflection components using a specularity-invariant image representation. In *Proceedings of the IEEE International Conference on Image Processing ICIP* (Atlanta, USA, 2006), pp. 973–976.