Relighting of Facial Images

Peter Csakany   Adrian Hilton
Centre for Vision Speech and Signal Processing, University of Surrey
P.Csakany@surrey.ac.uk   A.Hilton@surrey.ac.uk

Abstract

We present a novel method to relight video sequences given known surface shape and illumination. The method preserves fine visual details. It requires single view video frames, approximate 3D shape and standard studio illumination only, making it applicable in studio production. The technique is demonstrated for relighting video sequences of faces.

1. Introduction

Photo realistic relighting of video sequences is important in studio and film production. Beside the spectacular special effects of high budget movies, putting presenters into virtual environments during live studio broadcasts is becoming an often seen phenomenon in television. Our aim is to create a method which fits into the normal work flow of a television studio and is capable of altering the effect of lighting on the people in the studio. In the case of dynamic scenes this problem is difficult due to movement, requiring relighting at a single time instant. Approaches to re-illumination have generally only considered static scenes with known illumination and/or geometry. For static scenes it is possible to capture the scene under multiple illumination conditions. However, this is not possible for dynamic scenes such as people and faces where the subject moves and changes shape. In this paper, a method is presented for relighting frames of dynamic scenes based on the current frame and illumination only.

In this paper, relighting an objects means estimating reflection parameters and re-rendering the objects under different illumination as opposed to putting additional object into the scene or removing them [1].

There are several techniques which attempt to recover the full bidirectional reflection distribution function (BRDF) of a surface or a point. The technique in [2] requires at least one specular highlight visible on each surface to recover roughness and specular reflectance. The position of specular highlights are detected by knowing the position of the light sources, the camera pose and the geometry. It applies an iterative Monte Carlo simulation to refine the estimated BRDF and requires a high number of cameras. However, it only works in grey scale.

An image based approach is introduced in [3] to handle complex outdoor environments, however it is based on the estimation of the position of the sun and multiple images are needed at different times of the day. An advantage of this approach is that no 3D shape information is required. Other image based techniques [4][5] yield the specular free image of the scene without recovering the BRDF. In the case of uncoloured, diffuse Lambertian objects, it is possible to recover the shape and the direction of illumination at the same time [6].

Shape based techniques require the 3D shape of the object in the scene to be known, however sometimes only diffuse reflection is considered [7] or a reference object of the same material is required to be relit and to be present in the scene [8].

In the case of a complex illumination, the problem of recovering the reflection properties of the objects in the scene is more complicated, hence many approaches restrict the illumination to a single point light source [9] or apply multiple point light sources to illuminate the scene [10][11]. Spherical harmonics are also used in the estimation of reflectance parameters of diffuse surfaces. However these techniques do not consider self shadows and colours [12].

The proposed technique combines the image based and shape based methods capable of handling high frequency textures like human skin. It handles diffuse and specular surfaces alike and accounts for self shadows as well. However, this approach cannot recover the full BRDF, but only the diffuse reflection parameters. Since we aim to apply the proposed method in a standard studio environment we reject the usage of special lighting or high speed cameras [13][14], however we intend to apply it to dynamic scenes.

2. Estimating reflection parameters

There are various reflection models formulated in the vision literature [15][16][17]. We chose to apply the dichromatic [3] reflection model for our research, because it explicitly separates the specular and the diffuse reflection components:

$$I(\lambda) = \rho(\lambda) \int_L I_d(\lambda)\, \bar{n} \cdot \bar{l}_d + k_s \int_L I_s(\lambda, \sigma, \bar{n}, \bar{l}, \bar{v}) \quad (1)$$

The first part of expression 1 represents the diffuse part of the reflected light, where $\rho(\lambda)$ and $\bar{n}$ are the spectral reflectance (colour) and the surface normal at a given point of the surface of the object, respectively. L is the set of *effective light sources* - the light sources

illuminating the surface point. $I_d(\lambda)$ and $\bar{l}_d$ are the spectral intensity and the direction of an elementary illuminant, respectively. The second part of expression 1 represents the specular reflection, where k_s is a scalar scaling factor for non-metallic materials and $I_s(\lambda,\sigma,\bar{n},\bar{l},\bar{v})$ represents the dependence of the specularity on the surface roughness (σ), the surface normal $(\bar{n})$, the illumination $(\bar{l})$ and the view direction $(\bar{v})$.

The diffuse component is independent of the view point of the observer, therefore it can always be substituted by a virtual point light source.

$$\bar{I}_D(\lambda)=I_D(\lambda)\bar{l}_D=\int_L I_d(\lambda)\bar{l}_d \qquad (2)$$

and

$$I(\lambda)=\rho(\lambda)I_D(\lambda)\bar{n}\cdot\bar{l}_D+k_s\int_L I_s(\lambda,\sigma,\bar{n},\bar{l},\bar{v}) \qquad (3)$$

The colour of the diffuse reflection depends on the spectral reflectance of the surface point and the colour of the illumination. For non-metallic surfaces, k_s can be assumed to be independent of λ, therefore the specular reflection component is independent of the surface colour and only depends on the colour of the illumination.

At a surface point, the effect of the specular reflection changes more rapidly than the effect of the diffuse reflection, since for a given view point $I_s(\lambda,\sigma,\bar{n},\bar{l},\bar{v})$ is vanishing everywhere except in a small neighbourhood of the $\bar{v}$ mirrored on $\bar{n}$, whilst $I_D(\lambda)$ is the integral of all the light sources in a hemisphere visible from a surface point.

The specular reflection component is very sensitive to uncertainty in the surface normal and the light source direction. In our experiments the light sources are part of our studio environment and the objects are real human faces that we represent as a 3D mesh generated by a multi camera stereo capture system [18], hence both the surface normals and the light source directions are inherently noisy.

Our aim in the further sections is to remove the effect of specular reflection from the facial images and to introduce a framework to estimate the diffuse spectral reflection of each point on the surface of an object. In this paper, we introduce a three steps relighting process; first the specular highlights are removed from the image using an image based technique resulting in a *natural image;* second we estimate diffuse saturation for each colour region of the facial image; third we estimate colour albedo at each point of the scanned face.

2.1.Monochrome light assumption

In the case of illumination with multiple coloured light sources and constant view point, the colour of specularly reflecting light changes across the surface with the change of surface normal. The colour of the
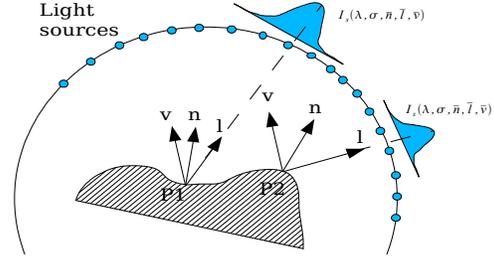

Figure 1 The contribution of light sources to the specular reflection from two different surface points.

diffuse reflection is different from the colour of the specular reflection since the effective kernel of integration in expression 1 is larger. The perceived colour of a surface point is the combination of the diffuse and the specular colour.

In the case of a monochromatic illumination the colour of the specular reflection is constant and equivalent to the colour of the illumination. The colour of the diffuse reflection is the combination of the colour of the object surface and the colour of the illumination

$$I(\lambda)=\rho(\lambda)r(\lambda)I_D\bar{n}\cdot\bar{l}_D+k_s r(\lambda)\int_L I_s(\sigma,\bar{n},\bar{l},\bar{v}) \qquad (4)$$

where $r(\lambda)$ is the spectral distribution of the illumination and I_D and I_s are scalar intensity values. If the colour of the illumination is known, the effect of the coloured illumination can be removed and turned into a white light illumination.

$$I_w(\lambda)=\frac{I(\lambda)}{r(\lambda)}=\rho(\lambda)I_D\bar{n}\cdot\bar{l}_D+k_s\int_L I_s(\sigma,\bar{n},\bar{l},\bar{v}) \qquad (5)$$

The specular reflection of the spectrally normalised illumination, in expression 5, is always white, whilst the diffuse reflection preserves the original material colour. This means that the stronger the specular reflection the whiter the reflected light is. The effect of white light specularity can be better illustrated using the HSV colour model. The hue component of the reflected light remains unchanged by white sepcularity, however its saturation decreases as the intensity of the specular term increases. This is illustrated in figure 2.

The intensity of the reflected light in a combined diffuse-specular case:

$$I=k_d I_D+k_s I_S \qquad (6)$$

where k_d is the diffuse reflection coefficient. Expression 6 can be expanded for each colour channel:

$$I^c=\frac{c}{3}k_d I_D+\frac{1}{3}k_s I_S \qquad (7)$$

where $c\in\{R,G,B\}$. We define saturation as :

$$S=1-3\,M/I \qquad (8)$$

where $M=min(R,G,B)$ and R,G,B are the red, green blue pixel values, respectively. By substituting expression 7 into expression 8 the relationship between intensity and saturation can be derived:
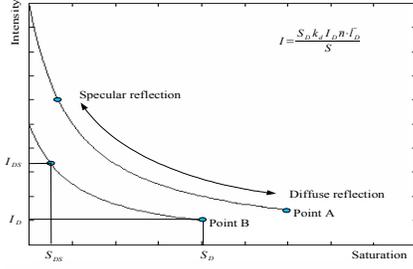
Figure 2 *The effect of white light specularity to the intensity and saturation of the reflected light for two different materials.*
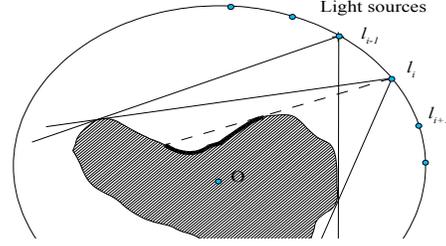

Figure 3 *The highlighted region is occluded from light source i, hence rendering the object from the point of view of light source i the highlighted area is invisible.*

$$S = 1 - 3\left(\frac{M}{3} k_d I_D + \frac{1}{3} k_s I_S\right)/I$$
$$= 1 - (M k_d I_D + (I_{DS} - k_d I_D))/I \qquad (9)$$
$$= k_d I_D (1 - M)/I$$

$$S I = k_d I_D (1 - M) \qquad (10)$$

Equation 10 shows that, under the assumption of monochromatic illumination, the saturation-intensity product is independent of the specularity. Hence it must have the same form under diffuse only illumination $I_S = 0$,

$$S_D = 1 - 3\frac{M}{3} k_d I_D / I_{diffuse} \qquad (11)$$

where $I_{diffuse} = k_d I_D$ and in the case of a white light illumination S_D is equivalent to the saturation of the surface material.

$$S_D I_{diffuse} = k_d I_D (1 - M) \qquad (12)$$

Substituting expression 10 into expression 12 we get

$$S I = S_D I_{diffuse} = S_D k_d I_D = S_D k_d I_d \bar{n} \cdot \bar{l}_d \qquad (13)$$

Since the right hand side of expression 13 depends only on the diffuse parameters, the *SI* product is independent of the viewpoint of the observer. If the observer moves around a specular object the amount of specular reflection from the point varies, hence the perceived colour of the point changes, however the *SI* product stays invariant to view direction.

2.2. The natural image

The *SI* product depends on the distribution of the illuminant, the shape and the diffuse reflectivity of the material making up the surface. The shape of the object is represented in two levels. On the local level, the shape of the surface is represented by the surface normal. On the global level, the 3D shape modifies the illumination by masking parts of the incident light field out, hence changing the kernel of the integral in expression 2. The *SI* product for each object point can be calculated from a colour image of the object. If the illumination field and 3D object shape are known, the diffuse reflectance of the object can be recovered form a single image.

$$S_D k_d = S I / (I_D \bar{n} \cdot \bar{l}_D) \qquad (14)$$

where *S* and *I* are the saturation and intensity of a point calculated as a result of the HSV conversion of the colour image. The $S_D k_d$ product is a illumination independent material property. The image of the object replaced by the $S_D k_d$ product is an illumination independent representation of the material of the object, therefore we call it the *natural image*.

The accuracy of the natural image depends on the accuracy of the 3D description (shape and normal) of the object to be modelled and the accuracy of the position of the light sources. For capturing 3D shape of dynamic objects, such as the human face, we developed an active light multiple stereo capture system, that uses infra red patter projection to simultaneously capture shape and colour at video rate and produces detailed triangular mesh representation. The illumination environment is captured by taking a series of images of a mirror ball with different shutter speed and reconstructing the *light source image* as a high dynamic range image. The light source image maps a sphere to the image plane so that each pixel in the light source image acts as a light source, located on the sphere. Due to self occlusions, not all the light sources illuminate all the points of the object. The set of light sources illuminating a point make up the list of effective light sources. To establish the set of effective light sources *(L)* for each surface point, the object must be rendered from the view point of each light source.

Since the light source image contains a large number of pixels, the generation of the list of effective light sources is a very time consuming process. For computational efficiency, we sub-sample the original light source image and consider only the highest intensity pixels. This approximates the original light field with a low number, typically 20-50, light sources. Since the applied sampling is non-uniform, these light sources occupy different solid angle in the spherical light field changing the radiant energy reaching a point.

$$I_d = I_{average} \cdot \omega \qquad (15)$$

where $I_{average}$ is the average intensity within a sub-sampled region of the light source image, ω is the solid angle this region occupies. The direction of an elementary light source is equivalent to the vector from

the middle of the object to the centre of the area of the light source.

$$l_d = \overline{Ol} \tag{16}$$

where l is an elementary light source and O is the centre of the object in figure 3.

2.3. The colour albedo and modelling material variation

The natural image represents properties related to the reflectivity of the material, however it does not contain any information about the spectral distribution of the reflected light. The object surface does not reflect light uniformly in all frequencies. Using the RGB colour model the spectral reflectivity of a material is:

$$\rho(\lambda) = k_d \cdot (r_{\lambda_R}, r_{\lambda_G}, r_{\lambda_B}) \tag{17}$$

where

$$\sum_{i=R,G,B} r_{\lambda_i} = 1 \tag{18}$$

and r_{λ_R}, r_{λ_G} and r_{λ_B} are the *normalised spectral reflectance values* in the frequency range of the red, green and blue sensors, respectively. As discussed in section 2.1, due to the white light assumption the hue component of the colour of the reflected light is equivalent to the hue component of the colour of the surface point, whilst the saturation of the reflected light is dependent on the amount of specularity. In the case of diffuse only reflection, the saturation of the reflected light, S_D is equivalent to the saturation of the colour of the surface point. It means that the surface points sitting on e hyperbola, similarly to figure 2, with their respective intensity and saturation values cannot be turned back to their diffuse state as long as their S_D diffuse saturation values are not known. According to expression 14, S_D cannot be estimated directly, since it cannot be decoupled from k_d, without some extra information. This extra information is acquired by a region based approach.

If the object surface is made of a small number of uniformly saturated patches and the spatial distribution of the effective light sources is sparse, then it can be assumed that the reflected light from most of the surface area of a patch is diffuse, hence the saturation of the reflected light is equivalent to S_D. The peak of the histogram (S_{peak}) of all saturation values of a uniformly saturated patch provides a good approximation of the saturation of the diffuse reflection for that surface patch.

$$S_D \approx S_{peak} \tag{19}$$

Points with lower saturation value than the S_{peak} must be partly specular, however points with higher saturation value than S_{peak} cannot be specular. Their increased saturation must either be the result of noise or variation of the saturation of the colour of the material, but not the result of specularity, hence the
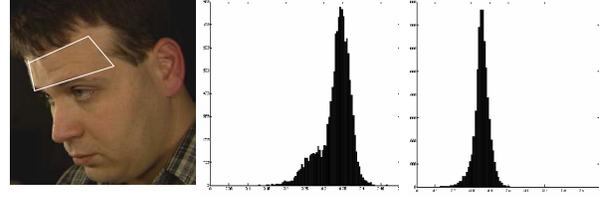

Figure 4 top, the saturation histograms of the boxed area and the entire face shown left, respectively; The left hand side of the boxed area is specular, whilst most of the face is diffuse, hence diffuse saturation is approximated by the peak of the histogram.

diffuse saturation for each point of a near uniformly saturated surface patch is

$$S_D = \begin{cases} S_{peak} & if \quad S < S_{peak} \\ S & if \quad S \geq S_{peak} \end{cases} \tag{20}$$

Note, that the constraint of uniformly saturated patches is less restrictive than the constraint of uniformly coloured patches. The hue, perceptually the most salient parameter of colour, can vary within a patch whilst the saturation of the patch is constant. A patch can turn form red to green without violating the assumption of uniform saturation but cannot turn from dark blue to light blue.

By substituting S_D into expression 14, the value of the diffuse reflectance, k_d, is obtained. The estimated value of the diffuse saturation and the point based hue value of the surface colour make up the normalised spectral reflectance values of each surface point. The colour triplet of the spectral reflectances act as colour albedo.

$$colour\ albedo = \rho(\lambda) = (k_d r_{\lambda_R}, k_d r_{\lambda_G}, k_d r_{\lambda_B}) \tag{21}$$

If the object surface is detailed, the assumption of a small number of uniformly saturated surface patches does not hold, then the histogram technique cannot be applied directly. However, many materials can be approximated as near uniformly saturated with high frequency disturbances, such as the human skin. Skin has a dominant colour riddled with small anomalies caused by wrinkles, hair or discolouration. In the case of such material a modified technique is applied. If the total area of all disturbances are small compared to the area of the surface patch, the histogram technique still gives a good approximation of the diffuse saturation of the patch, however S_D cannot be applied to the disturbed points. If the correct diffuse saturation of the disturbed point is larger than estimated, then its reflectance will be lower, whilst if the estimated saturation value is lower than the correct value the reflectance will be higher than its correct value. After a subsequent relighting, this results in dark or bright points giving salt and pepper noise errors.

To eliminate this problem, first the anomalies are removed by applying a median filter to the colour images. This removes small deviations from the

material colour whilst minimising the effect of blurring.

Second, the diffuse reflectance is estimated for each point by finding the peak of the saturation histogram, approximating S_D similarly to expression 20:

$$S_D = \begin{cases} S_{peak} & if \quad S < S_{peak} \\ S_{median} & if \quad S \geq S_{peak} \end{cases} \qquad (22)$$

and substituting S_D into expression 14. This sets the diffuse saturation near uniform. This value does not contain the small variation caused by the disturbances in the material colour. In this case, the material is a good approximation of the uniformly saturated patch assumption.

Finally, to preserve surface details features removed by median filtering are reintroduced by modifying the value of the diffuse reflection coefficient:

$$k_d = k_d^{median} I / I_{median} \qquad (23)$$

In this way, the original colour disturbances are represented as intensity variation, however their colour is lost.

2.4. Relighting

The colour albedo and the 3D shape of an object provides enough information to generate a diffuse only rendering of the object under arbitrary illumination applying the first part of expression 1. However, we need the specular reflection parameters to produce a photo realistic rendering. As explained in section , the specular parameters are very sensitive to variations in both the 3D shape and the distribution of illuminants, therefore our method reduces the effects of specularity rather than estimates it. To enhance the visual effect we chose to acquire the specular parameters, k_s and $I_s(\lambda, \sigma, \bar{n}, \bar{l}, \bar{v})$, from a pre-recorded library of materials. This allows relighting under complex and arbitrarily coloured illumination.

3. Results

Figure 5 depicts the original colour image. The 3D model of the face is represented as a triangular mesh consisting 24123 triangles. The surface of the face is very detailed, however it is undulated due to noise during the data acquisition and the subsequent processing steps. The effect of white light illumination and the high number of skin features in the facial image are clearly visible in figure 5. The aim of the proposed process is to preserve these features while radically changing the effect of the illumination.

The middle and right images in figure 5 show the natural image and the recovered colour albedo. Due to the filtering steps applied during the recovery process the natural image is visibly blurred. This is compensated in the colour albedo image. The colour albedo image is free from the effects of specularity and depicts the diffuse colour and reflectivity of each pixel.

The colour albedo serves as a basis for further relighting.

Figure 6 depicts the final relighting of the face using three different light fields. The left image is the result of a diffuse relighting with a single point light source placed at a 45º angle to the right of the centre of the face. The middle and right images in figure 6 were rendered under a point illumination from the opposite side of the face, however the skin model in the right image contains added specularity visible in the forehead and cheek.

The images in figure 6 were generated by applying the uniformly saturated surface assumption to the whole of the image including the eyes, lips and the hair. The results verify that the assumption is valid for most parts of the face. The eyes show significant deviation. The colours of the iris are less saturated than the rest of the face, whilst the saturation of the rest of the eye is very low since it is close to white. The decreased saturation of the eye is mistaken for specularity resulting in decreased intensity values in the colour albedo image. This can be eliminated by segmenting the face and handling the eye differently.

Another visible effect in figure 6 is the elevated intensity values in the neck region. This is caused by self-illumination from the shoulders and lower body and ambient light. These are not taken into consideration because only the face is captured, therefore no such information is available.

Figure 5 left, original colour image; middle, natural image, right, colour albedo image

Figure 6 left, diffuse rendering using a point light source illumination placed in the right of the image. Middle, diffuse rendering using a point light source in the centre of the image. Right, same as middle with added specularity visible on the forehead and cheek.

Figure 7 left, the original image; middle and right the face rendered in a beach and in a forest environment.
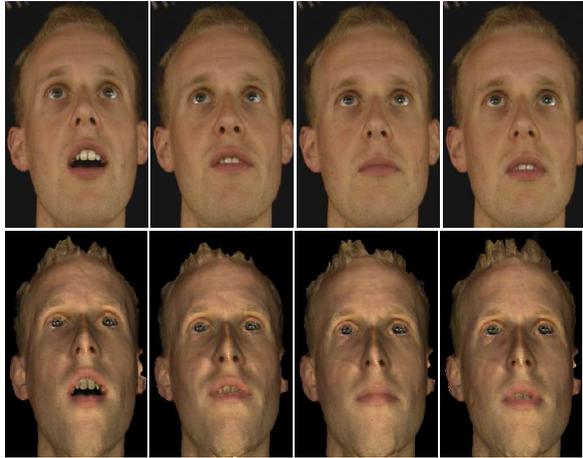

Figure 8 top, the original images of a sequence; bottom, the same images relit by a point light source offset to the right.

Figure 7 illustrates relighting with complex light fields. The original face is relit in two different environments, on a beach and in a forest, respectively. In the beach image, the main illuminant is the sun low above the horizon, whilst in the forest image the light of the sun is filtered by the forest resulting in a irregular illumination.

Figure 8 shows four frames of a dynamic sequence and the same frames relit using a point light source placed in front of the face slightly offset to the right simulating candle light. Self shadowing is visible around the nose. The face deforms during speech resulting in changes of the relit model. This effect is visible in the left of the images over the lips and also below the nose.

4. Conclusion and future work

The results demonstrate realistic relighting with preservation of fine visual detail. The approach developed is suitable for relighting moving objects such as faces from the captured image sequence. We have shown that relighting of scenes from a single colour image is possible if the illumination field and the 3D object shape are known. The presented technique fits into the standard work flow of studio production, however further enhancements are needed to achieve photorealism.

5. References

[1] C. Loscos and G. Drettakis and L. Robert, "Interactive Virtual Relighting of Real Scenes", IEEE Trans. Vis. Comput Gr., 2000, pp. 289-305

[2] Y. Yu, P. Debevec, J. Malik, and T. Hawkins, "Inverse Global Illumination: Recovering Reflectance Models of Real Schenes from Photographs", Proc. SIGGRAPH 99, 1999, pp. 215-224

[3] Y. Yu and J. Malik, "Recovering Photometric Properties Of Architectural Scenes From Photographs", Proc. SIGGRAPH 98, 1998, pp. 207-217

[4] R.T. Tan, K. Nishino and K. Ikeuchi, "Separating Reflection Components based on Chromaticity and Noise Analysis", IEEE T. Pattern Anal., 2004, pp. 1373-1379

[5] R.T. Tan and K. Ikeuchi, "Separating reflection Components of textured Surfaces using a Single Image", IEEE T. Pattern Anal., 2005, pp. 178-193

[6] D. Samaras and D. Metaxas, "Incorporating Illumination Constraints in Deformable Models for Shape from Shading and Light Direction Estimation", IEEE T. Pattern Anal., 2003, pp. 247-263

[7] C.B. Madsen and R. Laursen, "A Model-Based Approach to Image Relighting with a Potential for Real-time Implementation", International Conference on Vision, Video and Graphics, 2005, pp. 237-243

[8] A. Hertzmann and S.M. Seitz, "Example-Based Photometric Stereo: Shape Reconstruction with General, Varying BRDFs", IEEE T. Pattern Anal., 2005, pp. 1254-1264

[9] Y. Sato, M.D. Wheeler and K. Ikeuchi, "Object shape and reflectance modelling from observation", International Conference on Computer Graphics and Interactive Techniques, 1997, pp. 379 - 387

[10] S. Tominaga and B.A. Wandell, "Standard surface-reflection model and illumination estimation", J.Opt.Soc.Am., 1989, pp. 576-584

[11] HC Lee, "Method for computing the scene-illuminant chromaticity from specular highlights", J.Opt.Soc.Am, 1986, pp. 1694-1699

[12] R. Basri and D.W. Jacobs, "Lambertian Reflectance and Linear Subspaces", IEEE T. Pattern Anal., 2003, pp. 218-232

[13] T. Hawkins, A. Wenger, C. Tchou, A. Gardner, F. Goransson and P. Debevec, "Animatable Facial Reflectance Fields", Eurographics Symposium on Rendering, 2004, pp.

[14] A. Wenger, A. Gardner, C. Tchou, J. Unger, T. Hawkins and Paul Debevec, "Performance Relighting and Reflectance Transformation with Time-Multiplexed Illumination", Proc. SIGGRAPH 05, 2005, pp. 756-764

[15] S. Tominage, "Surface Identification Using the Dichromatic Reflection Model", IEEE T. Pattern Anal., 1991, pp. 658-670

[16] S.K. Nayar, K. Ikeuchi and T. Kanade, "Surface Reflection: Physical and Geometrical Perspective", IEEE T. Pattern Anal., 1991, pp. 611-634

[17] X.D. He, K.E. Torrace, F.X. Sillion and D.P. Greenberg, "A Comprehensive Physical Model for Light Reflection", Computer Graphics, 1991, pp. 175-186

[18] I.A. Ypsilos, A. Hilton, and S. Rowe, "Video-rate Capture of Dynamic Face Shape and Appearance", IEEE Int. Conf. Face and Gest. Recog., 2004, pp. 117- 122