Estimation of Point Light Source Parameters for Object–Based Coding

Jürgen Stauder

Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung
Universität Hannover
Appelstraße 9A
D–30167 Hannover
F. R. Germany

Mailing address:
Jürgen Stauder
Institut für Theoretische Nachrichtentechnik
und Informationsverarbeitung
Appelstraße 9A
D–30167 Hannover
F. R. Germany

Signal Processing: Image Communication,
Vol. 7, Nos. 4–6, November 1995, pp. 355–379. January 1995

Abstract: In this paper, the source model of moving rigid 3D objects of an Object–Based Analysis–Synthesis Coder (OBASC) is extended from *diffuse* to *non–diffuse* illumination introducing the explicit illumination model of a distant point light source and ambient diffuse light. For each image of a real image sequence containing moving objects, first, shape and 3D motion parameters describing the objects are estimated assuming an ellipsoid–like smooth shape. Then, the illumination parameters are estimated by a fast iterative Maximum–Likelihood Gauß–Newton estimation method. Typically, the illumination parameters converge after very few images close to the true ones. The accurateness depends on the amount of object rotation and the correctness of the shape assumptions. For a real image sequence showing a textured ball covering 20% of image area, rotating about 10^0 per frame, and illuminated by spot and ambient light, the extension of the source model reduces the model failures from 9.9% of the image area to 6.7%. In the area of model failures, the image synthesized from the source model parameters differ significantly from the real image. In this early experiment, source model parameters are coded losslessly. Since model failures are expensive by means of bit rate, a significant reduction of bit rate can be expected.

1. Introduction

For transmission of moving images at very low bit rates (e.g. video telephony), Object–Based Analysis–Synthesis Coding (OBASC) [12] based on the source model of moving rigid 3D objects [14] has been proposed. In an OBASC scheme, changes between two successive images of an image sequence are described by a source model. The source model defines a model world consisting of model objects, model illumination and a model camera. The model world corresponds to the real world consisting of real objects, real illumination and a real camera. The illumination is assumed to be diffuse and a perspective camera model is used. Parameters for the shape, motion and colour of the model objects are estimated by image analysis techniques. The parameters are coded, transmitted to the decoder, stored and used to synthesize a model image. The image quality of the model image is evaluated according to a receiver model. Areas of the model image where the image quality is not sufficient enough are called model failures. Only in these areas, colour parameters are coded and transmitted. Because coding of colour parameters is expensive by means of bit rate, the number of picture elements (pels) in areas of model failures is held to be small and can be used as an efficiency measure for the coding technique.

The illumination of a video telephony scene can consist of directed light from light sources and ambient light from walls or furniture. In case of such a non–diffuse illumination, both, object motion and illumination have an impact on the change between two successive images of an image sequence. If a real object rotates under non–diffuse illumination, temporal luminance variations on its surface can be observed, e.g. a part of the surface may get brighter or darker. The source models of known coding techniques [3][5][14] assume that changes from one image to the next are due to geometric effects of object or camera motion. Photometric effects due to object motion under non–diffuse illumination are neglected. Thus, model failures due to these photometric effects may occur. To handle non–diffuse illumination in an OBASC scheme, a more sophisticated illumination model has to be chosen, its parameters have to be estimated, and the parameters have to be used to synthesize the temporal luminance variations.

Pearson and Lee [9][15] were the first making use of illumination parameters in an OBASC scheme. They propose an irradiance compensated image synthesis. Irradiance is the measure of incident light on an object surface. By irradiance compensation, temporal luminance variations on object surfaces can be synthesized. Pearson and Lee used manually measured illumination parameters. For an OBASC scheme, illumination parameters have to be estimated automatically.

First approaches to illumination estimation [25] are based on a simple source model assuming uniformly–coloured still objects and a sole distant point light source without ambient light. The distant point light source is defined by an illumination direction. For instance, Pentland [16] estimates the illumination direction from one image assuming objects with equally distributed surface normals. For coding of moving images, such a source model has too many restrictions. Moving and textured objects have to be considered.

In case of moving textured objects under non–diffuse illumination, object motion estimation and illumination estimation are joint problems [17]. From two images, Moloney and Dubois [10] estimate for each pel a displacement vector and a temporal luminance variation. They

apply a global optimization method with a simple smoothness constraint for the temporal luminance variations. Mukawa [11] introduces the illumination model of a sole point light source for displacement estimation. From two images, he estimates for each pel a displacement vector and a temporal luminance variation. He applies a global optimization method smoothing the temporal luminance variations using a constraint based on the illumination model. From the displacement vector field, Mukawa estimates the shape of one object assuming rigidity. Finally, he obtains the illumination direction from the 3D object shape and the estimated temporal luminance variation field by a global optimization method.

In this paper, the source model of moving rigid 3D objects of an OBASC scheme will be extended from *diffuse* to *non–diffuse* illumination to improve the coding efficiency. In order to cover a wide range of possible lighting situations, the chosen illumination model consists of a distant point light source combined with diffuse ambient light [22]. The illumination model is defined by two illumination parameters for the direction of the distant point light source and one for its relative irradiance compared to 100% total irradiance of point source and ambient light.

The illumination parameters will be estimated from the shape and motion of the model objects provided by OBASC image analysis as well as from the temporal luminance variations observed between the current model image and the current real image. The model objects are 3D, but in opposite to Mukawa, instead of a motion vector field, a motion description by only one set of 3D rotation and translation parameters per object will be used allowing for efficient coding of object motion. In order to restrict the implementation complexity, a fast local Gauß–Newton optimization method will be applied. To avoid a divergence of the Gauß–Newton optimization method, a start point for the optimization will be evaluated from the illumination parameter estimates of former images.

The illumination estimation method will be incorporated in an OBASC scheme. Therefore, image synthesis and analysis techniques shall be extended. In image synthesis, an irradiance compensation based on the illumination model will be introduced. For image analysis, 3D motion and 3D shape will be estimated as in [14], but the motion estimation technique shall be adapted to the illumination model. The coding efficiency of the OBASC scheme will be assessed using the size of model failure as criterion.

The paper is organized as follows. In Section 2, the principle of OBASC is explained. In Section 3, the applied extended source model of moving rigid 3D objects with *non–diffuse* illumination is developed. In Section 4, the proposed illumination estimation technique is presented. Furthermore, it is shown how the method is embedded in the image analysis of an OBASC scheme. Results of simulations examining the coding efficiency of the OBASC scheme based on the extended source model are discussed in Section 5.

2. Principle of Object–Based Analysis–Synthesis Coding

OBASC [12] subdivides the scene visible in an image of a sequence into uniformly moving objects and describes each object by parameters defining its motion, shape, and colour. Motion parameters define position and motion of the objects, shape parameters their shape. Colour parameters include luminance as well as chrominance reflectance of the object surfaces. The illumination of the scene is described by illumination parameters. In Fig. 1, the concept of OBASC is shown. The parameters are saved in a parameter memory. The parameter memories in coder and decoder contain the same information. From theses parameters, at time instant k, the image synthesis generates a model image s_k' which is displayed at the decoder. The stored parameters and the current image s_{k+1} are used as input for the image analysis. First, the image analysis estimates new shape, motion and illumination parameters in order to reuse most of the already transmitted colour parameters. In the final step of image analysis, image areas are detected which cannot be described by the already transmitted colour parameters and the estimated shape, motion and illumination parameters. These areas of model failures are detected according to a receiver model. They tend to be small in size. This allows to transmit colour parameters for the areas for model failures. Colour parameter transmission is expensive by means of bit rate. Motion, shape, colour and illumination parameters are coded, transmitted, decoded and stored in the parameter memory.

Fig. 1

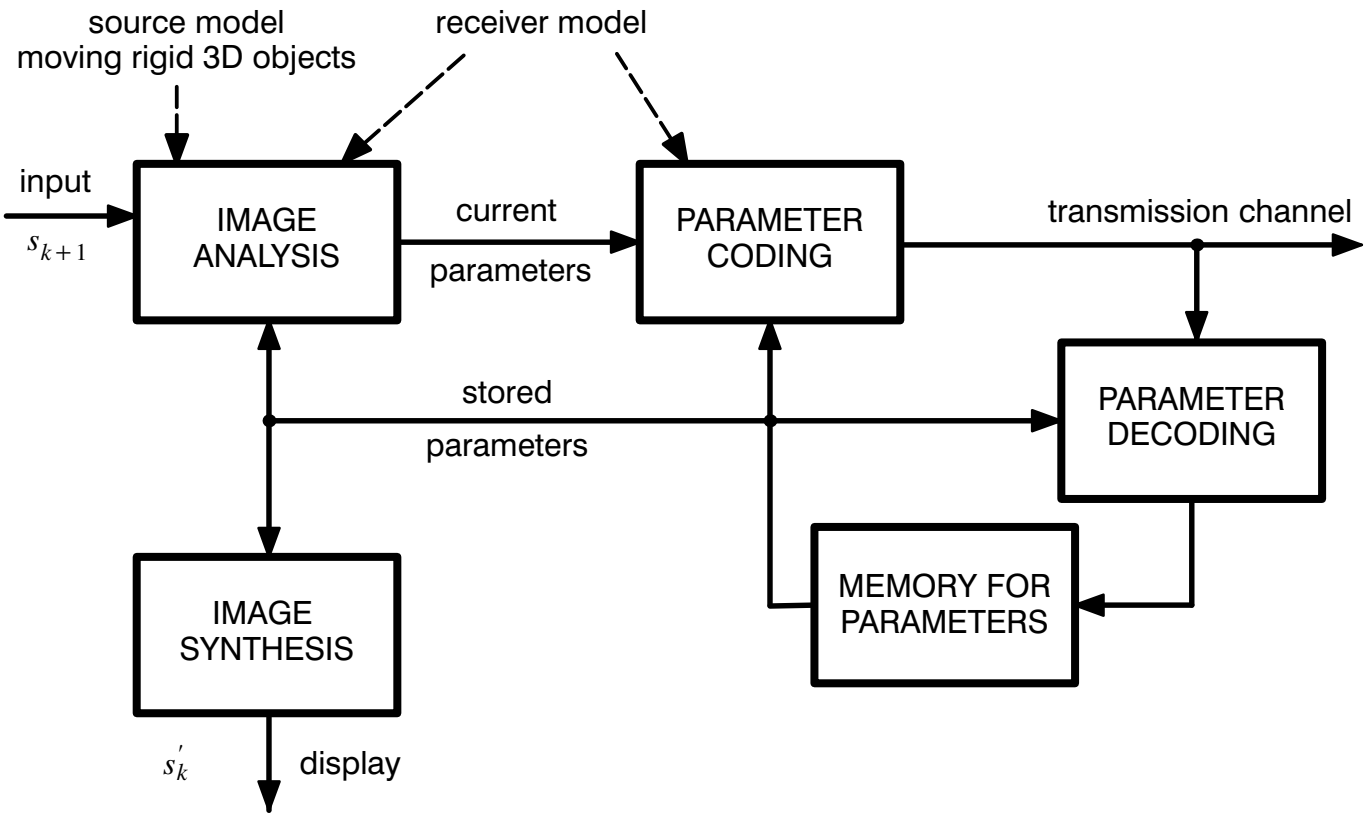


Fig. 1 Principle of Object–Based Analysis–Synthesis Coding based on the source model of moving rigid 3D objects

In OBASC, the suitability of source models can be judged by comparing the size of model failures analyzing the same image sequence at the same image quality, because the coding of colour parameters in the areas of model failures requires a huge part of available bit rate [14].

3. Source Model

Ostermann proposes the source model of moving rigid 3D objects [14] for OBASC. Implicitly he assumes *diffuse* illumination. This assumption is rejected in this paper. Therefore, a distant point light source and diffuse ambient light is introduced as an explicit illumination model. The new source model is that of moving rigid 3D objects with *non–diffuse* local illumination.

In this section the complete source model is described focussing on novelties caused by the introduction of the explicit illumination model. The source model is defined by a set of camera, illumination, and object parameters. They will be well defined in this section too. Finally, it is shown that for an irradiance compensated image synthesis in an OBASC scheme, a reduced number of illumination parameters is sufficient.

3.1 Illumination Model

The interaction of incident light from a light source with the surface of an object at point $\boldsymbol{P}$ is described by the distribution $L_{r,\lambda}$ of reflected radiance from an object surface depending on the distribution $E_{i,\lambda}$ of incident irradiance and the object surface reflectance function R at this point according to

$$L_{r,\lambda}(\boldsymbol{L},\boldsymbol{V},\boldsymbol{N},\boldsymbol{P},\lambda) \;=\; R(\boldsymbol{L},\boldsymbol{V},\boldsymbol{N},\boldsymbol{P},\lambda)\;\; E_{i,\lambda}(\boldsymbol{L},\boldsymbol{N},\lambda) \tag{1}$$

Here, $\boldsymbol{N}$ is the surface normal vector at point $\boldsymbol{P}$, $\boldsymbol{L}$ the illumination direction, $\boldsymbol{V}$ the viewing direction and λ the wave length of light, see Fig. 2. It is assumed that the surface of the objects is opaque, i.e. the incident irradiance is either reflected or absorbed. Further on, video telephone scenes with time–invariant viewing direction and scene illumination are assumed. Thus, Eq. (1) reduces to

Fig. 2

$$L_{r,\lambda}(\boldsymbol{N},\boldsymbol{P},\lambda) \;=\; R(\boldsymbol{N},\boldsymbol{P},\lambda)\;\; E_{i,\lambda}(\boldsymbol{N},\lambda) \tag{2}$$

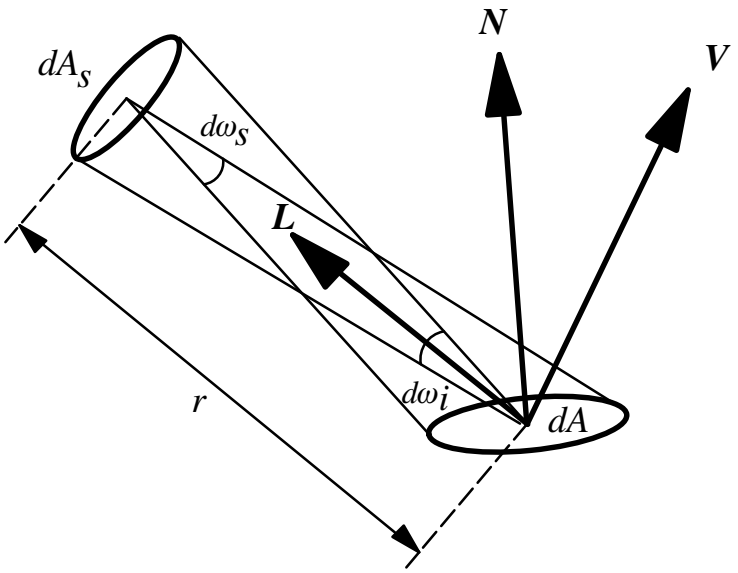


Fig. 2 Surface patch dA with normal vector $\boldsymbol{N}$ illuminated from direction $\boldsymbol{L}$ by a point light source with the infinitesimal small area dA_s and viewn from direction $\boldsymbol{V}$.

The distribution $E_{i,\lambda}(\boldsymbol{N},\lambda)$ of incident irradiance is factorized into a geometric and a photometric term:

$$E_{i,\lambda}(\boldsymbol{N},\lambda) = \beta_i(\lambda) \;\; E_i(\boldsymbol{N}) \tag{3}$$

The geometric term $E_i(\boldsymbol{N})$ denotes the dependency of the incident irradiance on the surface normal vector $\boldsymbol{N}$ and is modelled by the illumination model. Further on, the geometric term is called irradiance. The photometric term $\beta_i(\lambda)$ denotes the spectral distribution of incident light and will be covered by the object model, see Section 3.3.

In this paper, the objects are assumed to be illuminated by a point light source and ambient diffuse light. First, the incident irradiance due to the point light source will be evaluated. Therefore, an infinitesimal small object surface patch dA around point $\boldsymbol{P}$ is regarded, see Fig. 2. The patch is seen from the point light source under a solid angle $d\omega_s$. The light source has an infinitesimal small area dA_s and emits in direction to the object surface patch the radiance

$$L_s = \frac{d^2\Phi}{dA_s d\omega_s} \tag{4}$$

Radiance is defined to be the emitted power flux $d^2\Phi$ per solid angle and per emitting surface. The power flux $d^2\Phi$ emitted in direction to the surface patch dA causes there an irradiance

$$dE_i = \frac{d^2\Phi}{dA} \tag{5}$$

Irradiance is defined to be the incident power flux per receiving surface area. Inserting Eq. (4) into Eq. (5) gives

$$dE_i = \frac{L_s dA_s d\omega_s}{dA} \tag{6}$$

The surface patch is seen from the point light source under a solid angle

$$d\omega_s = \frac{dA \cos \measuredangle(\boldsymbol{L},\boldsymbol{N})}{r^2} \tag{7}$$

Here, $\boldsymbol{L}$ is a unit vector pointing from dA into direction of the point light source. $\boldsymbol{L}$ is defined by the spheric angles ϕ and θ according to

$$\boldsymbol{L} = \cos\phi \sin\theta \;\; \boldsymbol{e}_x \;\; + \;\; \sin\phi \sin\theta \;\; \boldsymbol{e}_y \;\; + \;\; \cos\theta \;\; \boldsymbol{e}_z \tag{8}$$

with *(x,y,z)* the world coordinate system. Inserting Eq. (7) into Eq. (6) leads to

$$dE_i = \frac{L_s \cos \measuredangle(\boldsymbol{L},\boldsymbol{N})}{r^2} dA_s \tag{9}$$

The total irradiance $E_i(\boldsymbol{N})$ follows from integration of dE_i over the area A_s of the light source according to

$$E_i(\boldsymbol{N}) = \int_{A_s} \frac{L_s \cos \measuredangle(\boldsymbol{L},\boldsymbol{N})}{r^2} dA_s \;\; = \;\; c_{LAMBERT}\, \boldsymbol{LN} \qquad c_{LAMBERT} = \frac{L_s\, A_s}{r^2} \tag{10}$$

During integration, $\boldsymbol{L}$ and r are assumed to be constant. Here, $c_{LAMBERT}$ is the irradiance caused by the point light source in case of perpendicular incidence. Eq. (10) is known as the cosine law [8][13] found by Lambert in 1760. Allowing additionally for ambient diffuse light and taking into account that the point light source will have influence only, if it illuminates the patch from above, the irradiance becomes

$$E_i(\boldsymbol{N}) = c_{AMBIENT} + c_{LAMBERT} \max\{0, \boldsymbol{LN}\} \tag{11}$$

Here, $\boldsymbol{L}$ and $\boldsymbol{N}$ are unit vectors, $c_{AMBIENT}$ is the ambient irradiance and $c_{LAMBERT}$ the point light source irradiance. Eq. (11) is part of the Phong reflection model [18] used in early computer graphics. Eq. (11) is a local illumination model due to the following assumptions:

- the distance r between light source and surface patch (and thus $c_{LAMBERT}$) is assumed to be constant all over the scene
- the illumination direction $\boldsymbol{L}$ is assumed to be constant all over the scene
- the illumination colour is assumed to be constant all over the scene
- shadowing is not considered
- mutual illumination is not considered

The illumination parameters are the spheric angles ϕ and θ, and the irradiances $c_{AMBIENT}$ and $c_{LAMBERT}$.

An alternative illumination model for OBASC is the discrete irradiance map [20][21]. It is a discrete map of irradiance values of the incident illumination on a half Gaussian sphere . This illumination model is not restricted to any kind or number of light sources. It was not chosen here, because there are a lot of unknown parameters (i.e. 9×9 irradiance values) to be estimated.

3.2 Camera Model

The camera model defines the geometric and photometric behaviour of the model camera. The geometric camera model describes where the reflected radiance from a point on the surface of an object hits the image plane. The photometric camera model describes, how much radiance hits the image plane.

3.2.1 Geometric Camera Model

The geometric camera model is a pinhole camera projecting a point $\boldsymbol{P} = (P_x, P_y, P_z)^T$ in the world coordinate system *(x,y,z)* onto the point $\boldsymbol{p} = (p_X, p_Y)^T$ on the camera's image plane according to

$$\boldsymbol{p}_X = F \frac{P_x}{P_z} \qquad \boldsymbol{p}_Y = F \frac{P_y}{P_z} \tag{12}$$

Fig. 3 where F is the focal length and *(X,Y)* the 2D image coordinate system, see Fig. 3. The viewing direction $\boldsymbol{V}$ is a unit vector in z–direction. The camera's origin is $(0,0,0)^T$.

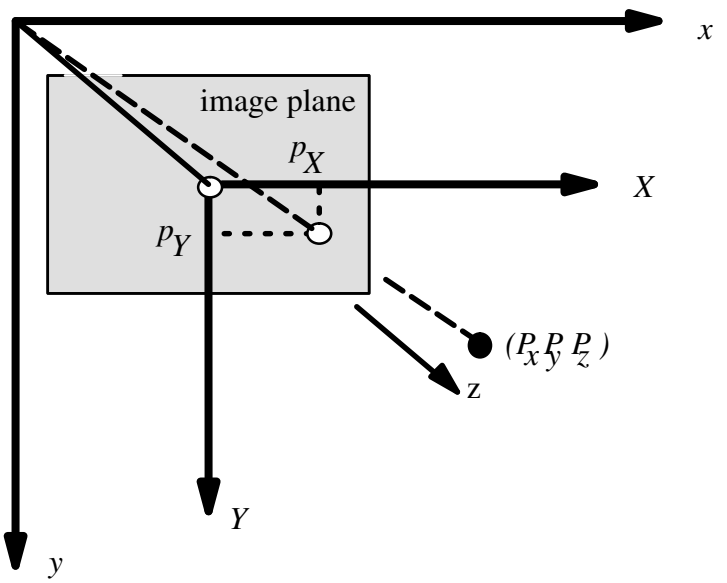


Fig. 3 Pin–hole camera model and coordinate systems: The origin of the camera coordinate system is at location *(0,0,F)* of the world coordinate system. *F* is the focal length of the camera.

3.2.2 Photometric Camera Model

The photometric model neglects optical blurr. The perspective projection is approximated by orthogonal projection [4]. Then, the distribution $E_{image,\lambda}$ of irradiance incident on the image plane at point $\boldsymbol{p}$ is proportional to the distribution $L_{r,\lambda}$ of the radiance emitted from an object surface point $\boldsymbol{P}$ according to

$$E_{image,\lambda}(\boldsymbol{N},\boldsymbol{p},\lambda) \;=\; k_{image}\;\; L_{r,\lambda}(\boldsymbol{N},\boldsymbol{P},\lambda) \qquad ; \qquad k_{image} = const. \tag{13}$$

The luminance $s_Y(\boldsymbol{p})$ at a point $\boldsymbol{p}$ of the model image generated by the model camera is

$$s_Y(\boldsymbol{p}) = \int_{\lambda=-\infty}^{\infty} \alpha(\lambda)\;\; E_{image,\lambda}(\boldsymbol{N},\boldsymbol{p},\lambda)\;\; d\lambda = k_{image} \int_{\lambda=-\infty}^{\infty} \alpha(\lambda)\;\; L_{r,\lambda}(\boldsymbol{N},\boldsymbol{P},\lambda)\;\; d\lambda \tag{14}$$

where $\alpha(\lambda)$ is the camera luminance responsivity. In this paper, the luminance image s_Y is refered to as *image*. The chrominances are not considered.

3.3 Object Model

The object model defines the shape of a model object, the change of its position by motion, and its surface properties, the colour. Colour includes luminance and chrominance. Chrominances are not considered in this paper.

3.3.1 Object Shape Model

The shape of a model object is a polygone approximation of the shape of the corresponding real object. It is defined by a set of control points $\boldsymbol{P}_C^{(i)}$ *i=0..I–1* spanning a mesh of triangles. Each control point is defined by its position $\boldsymbol{P}_C^{(i)} = (P_{C,x}^{(i)}, P_{C,y}^{(i)}, P_{C,z}^{(i)})^T$ in world coordinates. Each triangle is defined by the indices of its three edge points.

3.3.2 Object Motion Model

The 3D motion of any point $\boldsymbol{P}$ on the surface of a model object to its new position $\boldsymbol{P}'$ is described by 3D rotation and translation according to

$$\boldsymbol{P}' = [\boldsymbol{R}] \cdot (\boldsymbol{P} - \boldsymbol{C}) + \boldsymbol{C} + \boldsymbol{T} \tag{15}$$

with the translation vector $\boldsymbol{T} = (T_x, T_y, T_z)^T$, the object center $\boldsymbol{C} = \frac{1}{I} \sum_{i=0}^{I-1} \boldsymbol{P}_C^{(i)}$ and the rotation matrix according to

$$[\boldsymbol{R}] = \begin{bmatrix} \cos R_y \cos R_z & \sin R_x \sin R_y \cos R_z - \cos R_x \sin R_z & \cos R_x \sin R_y \cos R_z + \sin R_x \sin R_z \\ \cos R_y \sin R_z & \sin R_x \sin R_y \sin R_z + \cos R_x \cos R_z & \cos R_x \sin R_y \sin R_z - \sin R_x \cos R_z \\ -\sin R_y & \sin R_x \cos R_y & \cos R_x \cos R_y \end{bmatrix} \tag{16}$$

with the rotation angles $\boldsymbol{R} = (R_x, R_y, R_z)^T$ around x–, y–, and z–axis. Each model object has its own set of parameters $\boldsymbol{T}$,$\boldsymbol{R}$.

3.3.3 Object Colour Model

The colour of a model object surface point $\boldsymbol{P}$ is defined by the colour parameters. They are stored at time instant 0 after being taken from the real image $s_{Y,0}$ at position $\boldsymbol{p}_0$ as shown for the luminance parameter $s_{Y,0}(\boldsymbol{p}_0)$:

$$s_{Y,0}(\boldsymbol{p}_0) = k_{image} \int_{\lambda=-\infty}^{\infty} \alpha(\lambda) \; L_{r,\lambda}(\boldsymbol{N}_0, \boldsymbol{P}_0, \lambda) \; d\lambda \tag{17}$$

where $\boldsymbol{N}_0$ is the surface normal vector at the considered point $\boldsymbol{P}$ on the surface of an object at time instant 0. To reduce complexity of presentation in this paper, the time of colour parameter storage is fixed to time instant 0.

On the surface of the objects diffuse reflection is assumed. This inherits three assumptions:

- The wavelength of incident light λ is assumed to be much smaller than surface and material irregularities [13].

- Object surfaces are assumed to be matt, i.e. no reflection occurs at the interface between air and object, but only inside the body of the object [7], specular reflection [23] is not considered.
- Object surfaces are assumed to be isotropic [24].

To utilize the assumption of diffuse reflection, the reflectance of the object surface at a point $\boldsymbol{P}$ is factorized into a geometric and a photometric term according to

$$R(\boldsymbol{N},\boldsymbol{P},\lambda) = \varrho(\boldsymbol{N},\boldsymbol{P}) \;\; t(\lambda,\boldsymbol{P}) = \varrho \;\; t(\lambda,\boldsymbol{P}) \tag{18}$$

where $\varrho(\boldsymbol{N},\boldsymbol{P}) = \varrho$ is the geometric term depending no longer on $\boldsymbol{N}$ nor $\boldsymbol{P}$ due to diffuse reflection all over the object surface. Introducing the reflection from Eq. (2) and the reflectance from Eq. (18) into the luminance parameter in Eq. (17) leads to

$$s_{Y,0}(\boldsymbol{p}_0) = k_{image} \int_{\lambda=-\infty}^{\infty} \alpha(\lambda) \;\; \varrho \;\; t(\lambda,\boldsymbol{P}) \;\; E_i(\boldsymbol{N}_0) \;\; \beta_i(\lambda) \;\; d\lambda = t_Y(\boldsymbol{P}) \;\; E_i(\boldsymbol{N}_0) \tag{19}$$

where $t_Y(\boldsymbol{P})$ is defined to be the effective luminance reflectance of the point on the surface of an object according to

$$t_Y(\boldsymbol{P}) = k_{image} \;\; \varrho \int_{\lambda=-\infty}^{\infty} \alpha(\lambda) \;\; t(\lambda,\boldsymbol{P}) \;\; \beta_i(\lambda) \;\; d\lambda \tag{20}$$

Note that the reflectance $t_Y(\boldsymbol{P})$ includes camera and illumination features additionally to object features. This is no problem, because the reflectance, the camera, and the illumination are assumed to be time–invariant. To remember the orientation of the surface surrounding the considered point at colour parameter storage time, $\boldsymbol{N}_0$ is stored additionally.

The colour parameters are the luminance $s_{Y,0}(\boldsymbol{p}_0)$, the corresponding chrominances, and $\boldsymbol{N}_0$.

3.4 Image Synthesis

In OBASC, at the decoder a model image is generated by image synthesis. Therefore, the model camera projects the model scene onto its image plane. For image synthesis, the colour parameters of a model object are projected onto the image plane according to the object shape, the object position, and the camera model using an affine transform for speed up [14]. If an object has moved since the colour parameters were stored, the sampling rate may not match that of the image plane. Thus, bilinear interpolation is applied.

At time instant $k+1$ a model image $s'_{Y,k+1}$ is to be generated. If a model object has rotated since the colour parameters were stored, the irradiance has to be compensated. The colour parameters stored for a point $\boldsymbol{P}_{k+1}$ on the surface of a model object are irradiance compensated before projection at $\boldsymbol{p}_{k+1}$ on the image plane as shown for the luminance parameter:

$$s'_{Y,k+1}(\boldsymbol{p}_{k+1}) = s_{Y,0}(\boldsymbol{p}_0) \;\; \gamma(\boldsymbol{N}_0, \boldsymbol{N}_{k+1}) \tag{21}$$

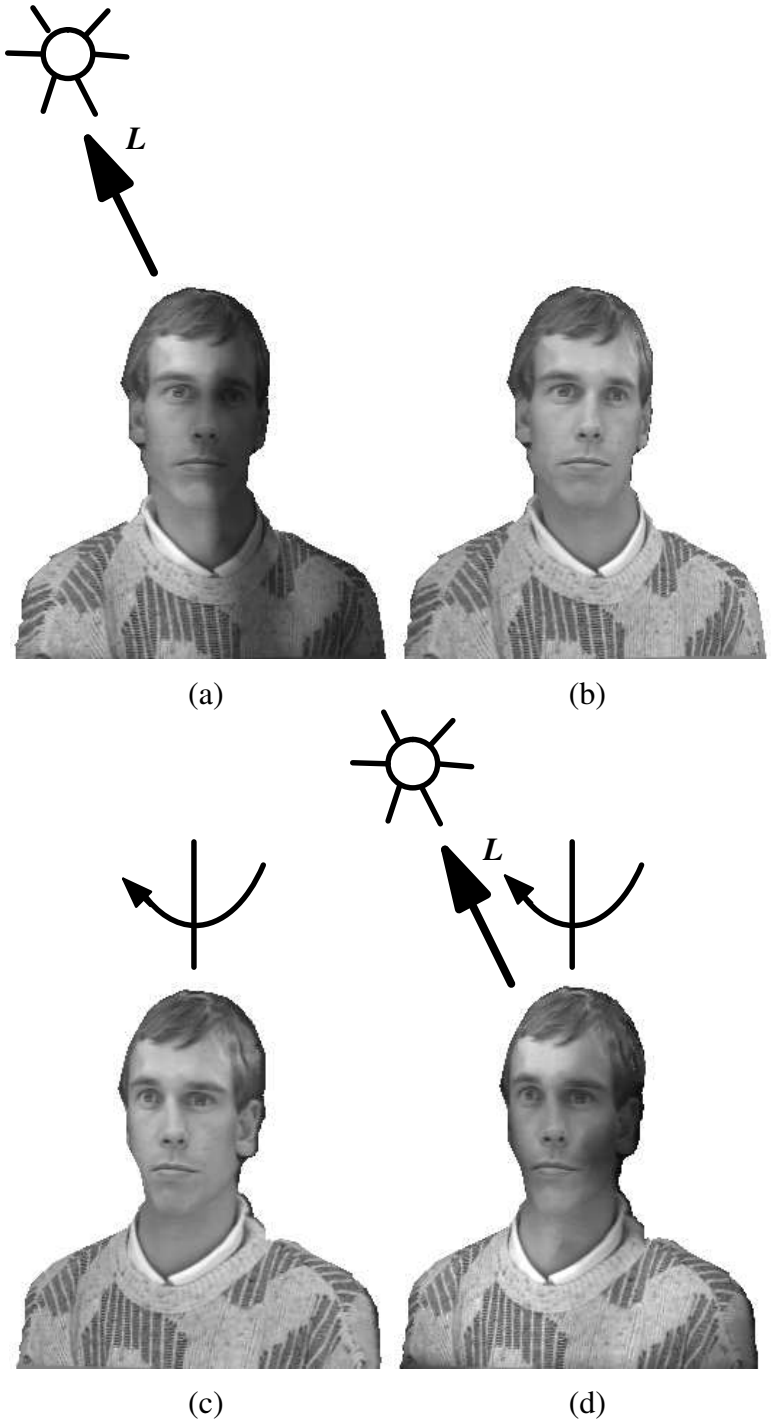


Fig. 4 Steps of irradiance compensation: **(a)** model object projected onto the image plane at time instant *0* using the stored colour parameters; **(b)** model object projected at time instant *0* with colour parameters free of inherent irradiance by application of the denominator of $\gamma(\boldsymbol{N}_{k+1}, \boldsymbol{N}_0)$ in Eq. (21); **(c)** same as (b) but at time instant k+1 after some rotation; **(d)** model object projected at time instant k+1 with irradiance compensated colour parameters according to Eq. (21).

Here, $s'_{Y,k+1}(\boldsymbol{p}_{k+1})$ is the luminance of the model image $s'_{Y,k+1}$ at point $\boldsymbol{p}_{k+1}$, $\boldsymbol{N}_{k+1}$ the surface normal vector at point $\boldsymbol{P}_{k+1}$ at time instant k, and $\gamma(\boldsymbol{N}_0, \boldsymbol{N}_{k+1})$ the irradiance compensation due to rotation of the surface patch surrounding $\boldsymbol{P}_{k+1}$ from $\boldsymbol{N}_0$ to $\boldsymbol{N}_{k+1}$ according to

$$\gamma(\boldsymbol{N}_0, \boldsymbol{N}_{k+1}) = \frac{E_i(\boldsymbol{N}_{k+1})}{E_i(\boldsymbol{N}_0)} \tag{22}$$

Inserting Eqs. (19) and (22) into Eq. (21) shows that the irradiance compensation in Eq. (21) gives the current image luminance $s'_{Y,k+1}(\boldsymbol{p}_{k+1}) = t_Y(\boldsymbol{P})\ E_i(\boldsymbol{N}_{k+1})$. Fig. 4 shows an example for irradiance compensation.

Fig. 4

As can be seen in Eq. (21), the expression for the irradiance is used only in a ratio to another irradiance. Inserting Eq. (11) in Eq. (21) shows that the illumination parameters $c_{AMBIENT}$ and $c_{LAMBERT}$ are used only relatively according to:

$$c = \frac{c_{LAMBERT}}{c_{AMBIENT}} \tag{23}$$

Thus, instead of those two parameters only c is necessary to be known. Hereby, c is introduced as a new illumination parameter substituting $c_{AMBIENT}$ and $c_{LAMBERT}$.

Due to the shape approximation by triangles, an irradiance compensation by Eq. (21) is discontinuous at the triangle edges. To smooth the compensation, $E_i(\boldsymbol{N}_{k+1})$ as well as $E_i(\boldsymbol{N}_0)$ are interpolated over the model object by Phong shading [18].

4. Illumination estimation

In case of non–diffuse illumination in the real scene, both, real object motion and real illumination have an impact on the change between two successive images of an image sequence. Thus, object motion estimation and illumination estimation are joint problems [17]. Therefore, in this chapter, an illumination estimation method is developed under close consideration of motion estimation. Finally, it is shown, how the illumination estimation method is integrated in the image analysis of an OBASC scheme.

4.1 Estimation of 3D Motion and Illumination

In this section, the basics of motion and illumination estimation are developed together. For estimation of 3D motion parameters $\boldsymbol{R}, \boldsymbol{T}$ and illumination parameters ϕ, θ, c, an analysis–synthesis technique is applied. This technique aims to motion and illumination parameters minimizing the frame difference between the real image $s_{Y,k+1}$ from an image sequence and a model image generated from the updated model world W_{k+1}. For estimation, it is assumed that the frame difference is caused by object motion under non–diffuse illumination only. Without motion, there will be no difference. With diffuse illumination or sole object translation, there will be a difference due to geometric effects only. With object rotation under non–diffuse illumination there will be a frame difference due to geometric and photometric effects.

4.1.1 Observation Points

For 3D motion and illumination estimation, the frame difference is evaluated only at a set of points in the image plane. The corresponding points on the surface of the model objects are called observation points. J observation points $\boldsymbol{O}^{(j)} = (\boldsymbol{P}_k^{(j)}, \boldsymbol{N}_k^{(j)}, \boldsymbol{N}_0^{(j)}, s_{Y,0}(\boldsymbol{p}_0^{(j)}), \boldsymbol{g}_{Y,0}(\boldsymbol{p}_0^{(j)}))$, $j \in [0,J-1]$ are selected, where $\boldsymbol{P}_k^{(j)}$ is the current 3D position, $\boldsymbol{N}_k^{(j)}$ the current surface normal vector at point $\boldsymbol{P}$, and $\boldsymbol{N}_0$ the surface normal vector at colour parameter storage time. At colour parameter storage time, the luminance $s_{Y,0}(\boldsymbol{p}_0^{(j)})$ and the spatial gradient $\boldsymbol{g}_{Y,0}(\boldsymbol{p}_0^{(j)})$ are measured at the corresponding 2D position $\boldsymbol{p}_0^{(j)}$ in the image $s_{Y,0}$ that was used for colour parameter storage. For each observation point, an optical flow equation is set up.

4.1.2 Optical Flow Equation

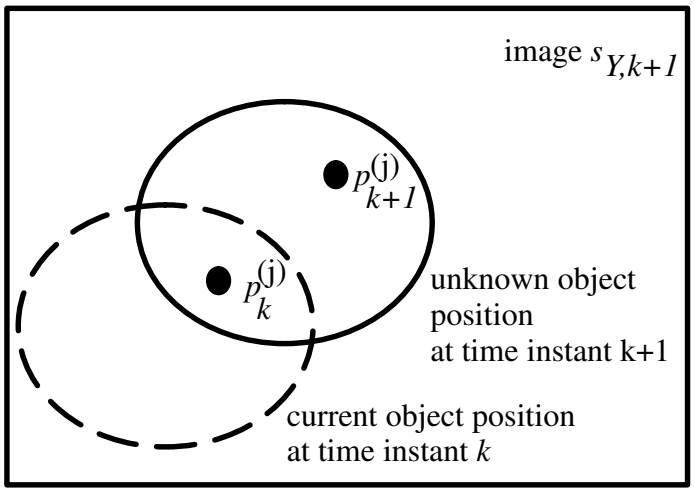


Fig. 5 Observation point $O^{(j)}$ in the image plane moving from $\boldsymbol{p}^{(j)}_{k}$ to $\boldsymbol{p}^{(j)}_{k+1}$.

The optical flow equation describes the geometric and photometric effects of motion for an observation point $\boldsymbol{O}^{(j)}$, i.e. a displacement in the image plane and a temporal luminance variation. In Fig. 5, the current position $\boldsymbol{p}^{(j)}_{k}$ and the unknown new position $\boldsymbol{p}^{(j)}_{k+1}$ of an observation point $\boldsymbol{O}^{(j)}$ in the current image $s_{Y,k+1}$ are shown. The luminance $s_{Y,k+1}(\boldsymbol{p}^{(j)}_{k})$ is known, whereas the luminance $s_{Y,k+1}(\boldsymbol{p}^{(j)}_{k+1})$ is unknown but can required to be

$$s_{Y,k+1}(\boldsymbol{p}^{(j)}_{k+1}) \stackrel{!}{=} s'_{Y,k+1}(\boldsymbol{p}^{(j)}_{k+1}) = s_{Y,0}(\boldsymbol{p}^{(j)}_{0}) \;\; \gamma \;\; (\boldsymbol{N}^{(j)}_{0}, \boldsymbol{N}^{(j)}_{k+1}) \tag{24}$$

with $\boldsymbol{N}_{k+1}$ the surface normal $\boldsymbol{N}_{k}$ rotated by the unknown rotation angles $\boldsymbol{R}$. In Eq. (24), irradiance compensation for the observed luminance of an observation point $\boldsymbol{O}^{(j)}$ due to the unknown rotation is shown. The difference between these luminances at $\boldsymbol{p}^{(j)}_{k}$ and $\boldsymbol{p}^{(j)}_{k+1}$ can be explained applying a linear signal model to $s_{Y,k+1}$ using the unknown displacement $\boldsymbol{p}^{(j)}_{k+1} - \boldsymbol{p}^{(j)}_{k}$ and the gradient $\boldsymbol{g}^{(j)}_{Y} = \textit{constant}$ of the image $s_{Y,k+1}$:

$$s_{Y,0}(\boldsymbol{p}^{(j)}_{0}) \;\; \gamma \;\; (\boldsymbol{N}^{(j)}_{0}, \boldsymbol{N}^{(j)}_{k+1}) \;\; - \;\; s_{Y,k+1}(\boldsymbol{p}^{(j)}_{k}) \;\; = \;\; \boldsymbol{g}^{(j)}_{Y} \;\; \left(\boldsymbol{p}^{(j)}_{k+1} - \boldsymbol{p}^{(j)}_{k} \right) \tag{25}$$

Eq. (25) can be improved by applying a quadratic signal model to $s_{Y,k+1}$ for evaluation of $\boldsymbol{g}^{(j)}_{Y}$ by averaging the gradients at the positions $\boldsymbol{p}^{(j)}_{k}$ and $\boldsymbol{p}^{(j)}_{k+1}$ according to

$$\boldsymbol{g}^{(j)}_{Y} \;\; = \;\; \frac{1}{2} \left(\boldsymbol{g}_{Y,k+1}(\boldsymbol{p}^{(j)}_{k}) \;\; + \;\; \boldsymbol{g}_{Y,0}(\boldsymbol{p}^{(j)}_{0}) \;\; \gamma(\boldsymbol{N}^{(j)}_{0}, \boldsymbol{N}^{(j)}_{k+1}) \right) \tag{26}$$

Eq. (26) has to be solved for the six motion parameters $\boldsymbol{R}, \boldsymbol{T}$ (see Eq. (15)) contained in $\boldsymbol{N}^{(j)}_{k+1}$ and $\boldsymbol{p}^{(j)}_{k+1}$ and for three illumination parameters ϕ, θ, c (see Eqs. (8) and (23)) contained in γ. The parameters are combined to a vector $\boldsymbol{U} = (R_x, R_y, R_z, T_x, T_y, T_z, \theta, \phi, c)^T$.

4.1.3 Optimization

Due to image noise and simplifications in the source model and image model, Eq. (25) will not match exactly for every observation point. For each observation point $\boldsymbol{O}^{(j)}$ there will be the residual error

$$f^{(j)}(\boldsymbol{U}) \;=\; s_{Y,0}(\boldsymbol{p}_0^{(j)})\;\gamma(\boldsymbol{N}_0^{(j)},\boldsymbol{N}_{k+1}^{(j)}) \;-\; s_{Y,k+1}(\boldsymbol{p}_k^{(j)}) \;-\; g_Y^{(j)}\left(\boldsymbol{p}_{k+1}^{(j)} - \boldsymbol{p}_k^{(j)}\right) \tag{27}$$

For a robust estimation of the 9 unknown parameters, $J \gg 9$ observation points have to be considered. If the residual error is modeled as Gaussian uncorrelated zero–mean noise, a Maximum–Likelihood estimator [2] has to minimize the MSE:

$$F\;(\boldsymbol{U}) \;=\; \sum_{j=0}^{J-1} f^{(j)\,2}\;(\boldsymbol{U}) \rightarrow \underset{\boldsymbol{U}}{MIN} \tag{28}$$

Eq. (28) is a nonlinear least square optimization problem. It is solved here by a Gauß–Newton iteration method [19]. Generally, the method works as follows. First, a start value $\boldsymbol{U}^0$ for the unknown parameters is evaluated. For each iteration m, $0<m\leq M-1$, starting at point $\boldsymbol{U}^{m-1}$ in the solution space, a new solution $\boldsymbol{U}^m = \boldsymbol{U}^{m-1} + \Delta\boldsymbol{U}^m$ is evaluated. Therefore, the residual error in Eq. (27) is linearized at point $\boldsymbol{U}^{m-1}$. Then, $\boldsymbol{U}$ is substituted by $\boldsymbol{U}^{m-1} + \Delta\boldsymbol{U}^m$. Eq. (28) is thus reduced to a *linear* least square problem and the Gauß algorithm of least squares can be applied. Therefore, a set of J linear equations has to be solved for $\Delta\boldsymbol{U}^m$. To exclude outlyers, only observation points $\boldsymbol{O}^{(j)}$ with $|f^{(j)}(\boldsymbol{U})| < \sigma_f$ are used in Eq. (28) with σ_f the standard deviation of the residuals in Eq. (27) of all observation points.

For estimation of illumination and motion parameters, it was found that a combined estimation is not convergent for the illumination parameters. Reason is that $F(\boldsymbol{U})$ is not a convex function and the start point $\boldsymbol{U}^0$ may not be close enough to the absolute minimum. The Gauß–Newton method performs well for convex functions and a start point $\boldsymbol{U}^0$ near to the absolute minimum [19]. Assuming little object motion, $\boldsymbol{R}^0 = \boldsymbol{0}, \boldsymbol{T}^0 = \boldsymbol{0}$ as start point works well for motion estimation [14]. The derivation of a start point ϕ^0, θ^0, c^0 for the illumination parameters is presented in Section 4.3.2.

But even having found a "good" start point $\boldsymbol{U}^0$, a combined estimation was found to be non convergent. This is, because Eq. (25) is only a rough approximation based on a linear signal model during the first iterations. Looking at $F(\boldsymbol{U})$ to be minimized (Fig. 6), it can be seen that the motion clues are much stronger than the illumination clues. During the first iterations, the Gauß–Newton method converges to the absolute minimum by means of the motion parameters. Because there is no markable concavity in $F(\boldsymbol{U})$ by means of the illumination parameters (see Fig. 6), the illumination parameters often get worse during the first iterations. Going on from such a bad point in solution space, the Gauß–Newton method is not able to converge to the absolute minimum, although Eq. (25) is now valid. This is due to saddle points and local minima. Thus, motion and illumination parameters are estimated separately.

Fig. 6

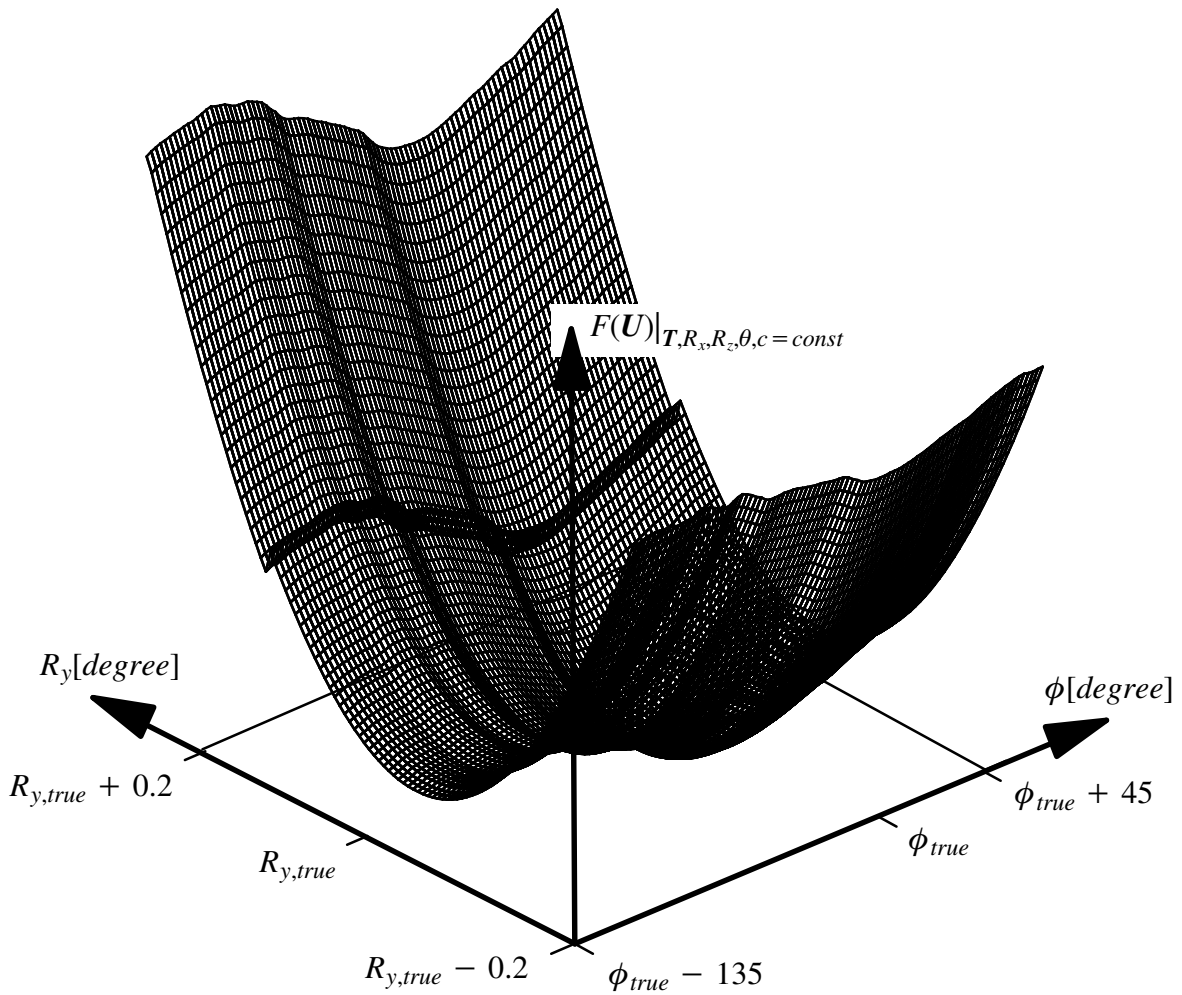


Fig. 6 Cost function to be minimized for motion and illumination parameter estimation measured using a synthetic ball and 3000 observation points with gradients $| \boldsymbol{g}_{Y,0}(\boldsymbol{p}_0^{(j)}) | > 50$. The coordinate system $(e_x, e_z, -e_y)$ is used here for the spheric illumination parameters: The parameters $\boldsymbol{T}, R_x, R_z, \theta, c$ are held constant at the optimal values, and R_y, ϕ are varied around their optimal values $R_{y,true}, \phi_{true}$. By variation of R_y, ϕ, the object or the point light source can be rotated around the y–axis, respectively. For minimization in ϕ–direction, much more saddle points and local minima can be seen than in R_y–direction.

4.2 Motion Estimation

The 3D motion parameters are estimated first. Therefore, the illumination parameters ϕ, θ, c are held constant during motion estimation. Either the estimates of the analysis of the last image are used or, at time instant 1, $c=0$, i.e. diffuse illumination is applied. In [14], the motion estimation method is described in detail. After decision for a seperate estimation of motion and illumination parameters, different sets of observation points can be used. For motion estimation a set of observation points with high spatial gradients $\boldsymbol{g}_{Y,0}(\boldsymbol{p}_0^{(j)})$ is chosen [14].

There are two extensions to the motion estimation method described in [14]. First, Eq. (25) contains an irradiance compensation now. Secondly, a hierarchical motion estimation by means of spatial image resolution is carried out, because the convergence of illumination parameter estimation will heavily depend on the convergence of the motion estimation. Therefore, two sets of observation points are generated. They differ in their luminance and their gradients. The first set is taken from a lowpass filtered version of the image $s_{Y,0}$, the second from the original $s_{Y,0}$. For hierarchical motion estimation, first, motion is estimated using the first set of observation points and a lowpass filtered version of $s_{Y,k+1}$. This step is to overcome local minima in $F(\boldsymbol{U})$ that can be seen in Fig. 6. The model objects are motion compensated. Then, the residual motion is estimated using the second set of observation points and the original image $s_{Y,k+1}$. Again, the model objects are motion compensated.

The described enhancements of the 3D motion estimation method from [14] were found to be effectual in order to get convergent iterations even in case of large motion under non–diffuse illumination.

4.3 Illumination Estimation

The illumination parameters are estimated after motion estimation. Therefore, the motion parameters $\boldsymbol{R}, \boldsymbol{T}$ are held constant during illumination estimation. It is assumed that the motion estimation was successful, i.e. the remaining differences between the observation points and the current image $s_{Y,k+1}$ are assumed to be caused by photometric effects of object motion.

4.3.1 Optimization of Illumination Parameters

For illumination estimation, Eq. (27) reduces to

$$f^{(j)}(\boldsymbol{U}) \;=\; s_{Y,0}(\boldsymbol{p}_0^{(j)}) \;\gamma\; (\boldsymbol{N}_0^{(j)}, \boldsymbol{N}_{k+1}^{(j)}) \;-\; s_{Y,k+1}(\boldsymbol{p}_k^{(j)}) \tag{29}$$

For illumination estimation, a set of observation points with low spatial gradients $\boldsymbol{g}_{Y,0}(\boldsymbol{p}_0^{(j)})$ is chosen. This reduces the sensitivity of Eq. (29) to slight motion estimation errors. The observation points are taken from a lowpass filtered image and for illumination estimation a lowpass filtered version of image $s_{Y,k+1}$ is used. The new optimization problem is

$$F\;(\boldsymbol{U}) \;=\; \sum_{j=0}^{J-1} f^{(j)\,2}\;(\boldsymbol{U}) \rightarrow \underset{\boldsymbol{U}}{MIN} \quad ; \quad \boldsymbol{U} \;=\; (\phi, \theta, c)^T \tag{30}$$

$F(\boldsymbol{U})$ is minimized applying the iterative Gauß–Newton method. Thus, $\boldsymbol{U}$ is substituted by $\boldsymbol{U}^{m-1} + \Delta\boldsymbol{U}^m$ and Eq. (29) is linearized at $\boldsymbol{U}^{m-1}$ leading to

$$
\begin{aligned}
f^{(j)}_{linear}(\boldsymbol{U}^m) \;&=\; f^{(j)}(\boldsymbol{U}^{m-1}) \\
&+ \frac{s_{Y,0}(\boldsymbol{p}_0)}{c^{m-1}} \frac{a_{k+1}\,\boldsymbol{L}^{m-1}\boldsymbol{N}_{k+1} \;-\; a_0\,\boldsymbol{L}^{m-1}\boldsymbol{N}_0}{\left(a_0\,\boldsymbol{L}^{m-1}\boldsymbol{N}_0 + 1\right)^2} \quad \Delta c^m \\
&+ s_{Y,0}(\boldsymbol{p}_0)\; c^{m-1} \frac{a_{k+1}\boldsymbol{L}^{m-1}_{\theta}\boldsymbol{N}_{k+1} \;-\; a_0\boldsymbol{L}^{m-1}_{\theta}\boldsymbol{N}_0}{\left(a_0\boldsymbol{L}^{m-1}\boldsymbol{N}_0 + 1\right)^2} \quad \Delta\theta^m \\
&+ s_{Y,0}(\boldsymbol{p}_0)\; c^{m-1} \frac{a_{k+1}\boldsymbol{L}^{m-1}_{\phi}\boldsymbol{N}_{k+1} \;-\; a_0\boldsymbol{L}^{m-1}_{\phi}\boldsymbol{N}_0}{\left(a_0\boldsymbol{L}^{m-1}\boldsymbol{N}_0 + 1\right)^2} \quad \Delta\phi^m
\end{aligned}
\tag{31}
$$

$$
\text{with } a_k \;=\; \begin{cases} 1 & \boldsymbol{L}^{m-1}\boldsymbol{N}_k \;>\; 0 \\ 0 & else \end{cases}
$$

$$
\text{and } \boldsymbol{L}^{m-1}_{\phi} \;=\; \frac{d}{d\phi}\boldsymbol{L}\Big|_{\phi=\phi^{m-1},\theta=\theta^{m-1}} \;=\; \begin{pmatrix} -\sin\phi^{m-1}\sin\theta^{m-1} \\ \cos\phi^{m-1}\sin\theta^{m-1} \\ 0 \end{pmatrix}
$$

$$
\text{and } \boldsymbol{L}^{m-1}_{\theta} \;=\; \frac{d}{d\theta}\boldsymbol{L}\Big|_{\phi=\phi^{m-1},\theta=\theta^{m-1}} \;=\; \begin{pmatrix} \cos\phi^{m-1}\cos\theta^{m-1} \\ \sin\phi^{m-1}\cos\theta^{m-1} \\ -\sin\theta^{m-1} \end{pmatrix}
$$

The linear least square optimization problem is

$$
F_{linear}\,(\Delta\boldsymbol{U}^m) \;=\; \sum_{j=0}^{J-1} f^{(j)\,2}_{linear}\,(\boldsymbol{U}^m) \;\rightarrow\; \underset{\Delta\boldsymbol{U}^m}{MIN} \quad ; \quad \Delta\boldsymbol{U}^m \;=\; (\Delta\phi^m, \Delta\theta^m, \Delta c^m)^T \tag{32}
$$

and is solved by the Gauß algorithm of least squares. Therefore, Eq. (31) for all *j=0..J–1* is set to 0 and written in matrix form according to

$$
\boldsymbol{f}^m \;+\; [\boldsymbol{A}]^m \cdot \Delta\boldsymbol{U}^m \;=\; \boldsymbol{0} \tag{33}
$$

with $\boldsymbol{f}^m$ a J–dimensional vector and $[\boldsymbol{A}]^m$ an $3 \times J$ matrix. Eq. (33) is overdetermined and solved for $\Delta\boldsymbol{U}^m$ by the Gauß algorithm of least squares according to

$$
\Delta\boldsymbol{U}^m \;=\; -\,[\boldsymbol{C}]^m\,[\boldsymbol{A}]^{m,T}\,\boldsymbol{f}^m \quad with \quad [\boldsymbol{C}]^m \;=\; \left[[\boldsymbol{A}]^{m,T}[\boldsymbol{A}]^m\right]^{-1} \tag{34}
$$

with $[\boldsymbol{C}]^m$ the covariance matrix of the estimation error.

Fig. 7

The optimization problem Eq. (32) is sometimes ill–posed. In Fig. 7, a rotating object under different illumination situations is shown. The more parallel the illumination direction is to the rotation axis the less photometric effects can be observed in the current image and the more unstable is an estimation of the illumination parameters. In those cases a dependency between

θ, ϕ and c leads to unstable results of Eq. (34) and large estimation error variances, the diagonal elements of $[\boldsymbol{C}]^m$.

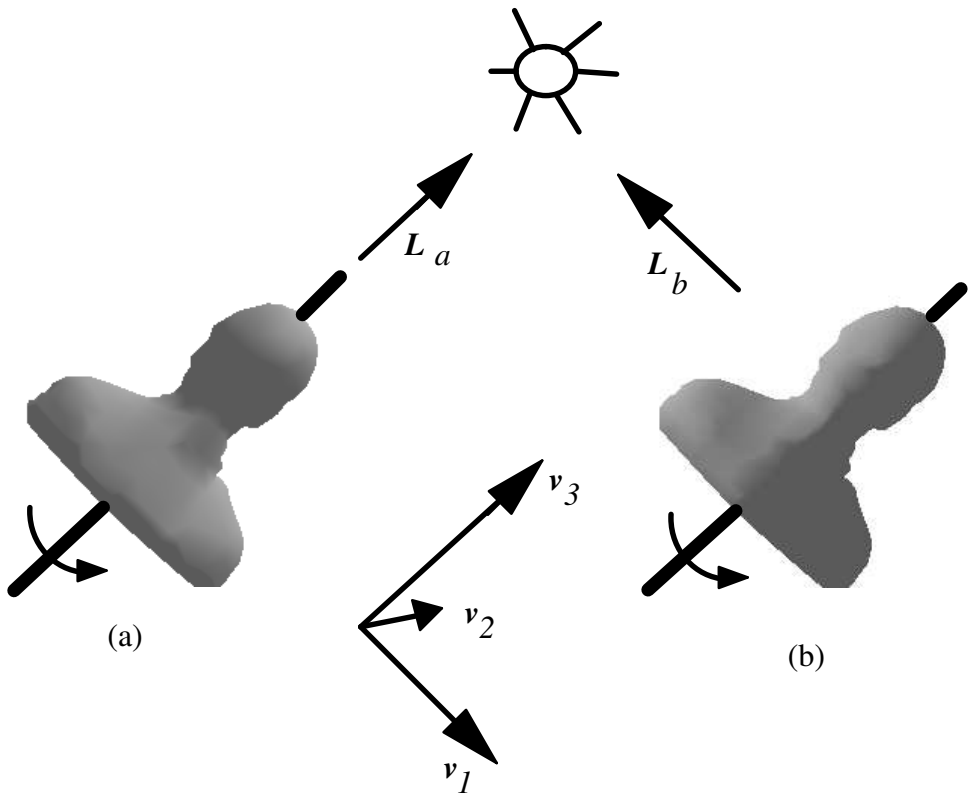


Fig. 7 An illuminated rotating object under two illumination situations: The illumination direction is (a) parallel to the rotation axis or (b) perpendicular. A rotation axis oriented coordinate system $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ is shown.

The dependency between θ, ϕ and c will be explained now. Therefore, an observation point $\boldsymbol{O}^{(j)}$ is considered visible for the point light source at time instant *0* (colour parameter storage time) and the current time instant *k+1*, i.e. $\boldsymbol{L}\ \boldsymbol{N}_0 > 0$ and $\boldsymbol{L}\ \boldsymbol{N}_{k+1} > 0$. The index j is suppressed here. According to Eq. (22), for such an observation point the following irradiance compensation will occur in Eq. (29):

$$\begin{aligned}\gamma\ (\boldsymbol{N}_0, \boldsymbol{N}_{k+1}) &= \frac{1 + c\ \cos\phi \sin\theta\ N_{k+1,1} + c\ \sin\phi \sin\theta\ N_{k+1,2} + c\ \cos\theta\ N_3}{1 + c\ \cos\phi \sin\theta\ N_{0,1} + c\ \sin\phi \sin\theta\ N_{0,2} + c\ \cos\theta\ N_3} \\ &= \frac{1 + c\ \sin\theta \left(\cos\phi\ N_{k+1,1} + \sin\phi\ N_{k+1,2} + \operatorname{atan}\theta\ N_3\right)}{1 + c\ \sin\theta \left(\cos\phi\ N_{0,1} + \sin\phi\ N_{0,2} + \operatorname{atan}\theta\ N_3\right)}\end{aligned} \tag{35}$$

In Eq. (35), the vectors $\boldsymbol{L}, \boldsymbol{N}_0, \boldsymbol{N}_{k+1}$ are represented in a rotation axis oriented coordinate system $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ shown in Fig. 7 instead of $(\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$ for explanation purpose. It is $N_{0,3} = N_{k+1,3} = N_3$, because $\boldsymbol{v}_3$ is the rotation axis. In Eq. (35), it can be seen that the

parameters c and θ are dependent if $N_3 \ll N_{0,1}, N_{0,2}, N_{k+1,1}, N_{k+1,2}$ or $N_3 \gg N_{0,1}, N_{0,2}, N_{k+1,1}, N_{k+1,2}$. This is for the equatorial and polar regions of the object in Fig. 7. For observation points not close to polar or equatorial regions or observation points with $\boldsymbol{L}\ \boldsymbol{N}_0 < 0$ or $\boldsymbol{L}\ \boldsymbol{N}_{k+1} < 0$, the dependency is not present. For estimation, vectors are represented in $(\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$ coordinate system instead of $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$. Thus, ϕ as well as θ may be dependent to c.

To overcome unstable estimation results, the illumination parameters are estimated iteratively alternately, i.e. Eq. (30) is iteratively alternately solved for ϕ, θ and for c. The covariances of the last iterations for ϕ, θ or c are denoted as $[\boldsymbol{C}]_{\phi,\theta}$ and $[\boldsymbol{C}]_c$, respectively.

4.3.2 Start Point Evaluation

At time instant k+1, the ML Gauß–Newton estimation technique requires a start point $\boldsymbol{U}^0_{k+1} = (\phi^0, \theta^0, c^0)^T$ in solution space close to the absolute minimum of $F(\boldsymbol{U})$. There are two cases to be examined: Finding a start point at time instant 1 is more difficult than at time instants k+1>1.

In [22], at time instant 1, an empirical start point was used, i.e. a weak relative irradiance $c^0 = 0.1$ in viewing direction $\boldsymbol{L}^0 = \boldsymbol{e}_z$. At the following time instants k+1>1, the estimate $\boldsymbol{U}_k$ of time instant k is used as start point. This leads to suboptimal results [22]. In this paper, a start point is evaluated by an empiric search (at time instant 1) or by Kalman filtering (at time instants k+*1*>1).

The first start point is found by an empiric search. The MSE in Eq. (30) is evaluated for 20 points in the solution space: According to Eqs. (8) and (23), first, the MSE for $c = \frac{1}{9}$ and θ=135^0, and for 16 values for ϕ in the range of $0..360^0$ in steps of 22.5^0 is evaluated. Secondly, the MSE for $c = \frac{1}{9}$, θ=180^0, and ϕ=0 is evaluated. A first start point candidate $\boldsymbol{U}^0 = (\phi^0, \theta^0, c^0)^T$ is chosen by means of minimum MSE. Additionally, at constant ϕ^0, θ^0, the MSE at $c = \frac{2}{8}$, $c = \frac{4}{6}$, and $c = \frac{8}{2}$ is evaluated. If the MSE can be reduced, the corresponding point is chosen as start point.

For k+*1*>1, the start point is found by Kalman filtering [1]. Therefore, the former estimates, refered to as observations $\boldsymbol{U}_1, ..., \boldsymbol{U}_k$, are Kalman filtered. For a time–invariant real illumination, a simplified Kalman filter [6] is shown in Fig. 8**.** The real world is modelled by time–invariant but unknown set of illumination parameters $\boldsymbol{X}_k$ to be estimated. They are observed at time instant k by the ML Gauß–Newton estimation, described in section 4.3. The observation error is modelled as Gaussian uncorrelated zero–mean noise with the covariance matrix $[\boldsymbol{C}]_k$. The covariance matrix is assumed to be diagonal. The diagonal elements are the variances of the observation errors for the three illumination parameters. They are taken from $[\boldsymbol{C}]_{\phi,\theta}$ and $[\boldsymbol{C}]_c$. The Kalman filter gives a minimum error variance estimate $\hat{\boldsymbol{X}}_k$ used as start point $\boldsymbol{U}^0_{k+1}$ for the ML Gauß–Newton estimation at time instant k+1. Additionally, the minimum error

Fig. 8

covariance $[\boldsymbol{P}]_k$ of the Kalman estimate is calculated. The iterative Kalman filter weights the new observation $\boldsymbol{U}_k$ by the correction matrix $[\boldsymbol{K}]_k$ according to:

$$\begin{aligned} [\boldsymbol{K}]_k &= [\boldsymbol{P}]_{k-1} \cdot \left([\boldsymbol{P}]_{k-1} + [\boldsymbol{C}]_k\right)^{-1} \\ \boldsymbol{X}_k &= \boldsymbol{X}_{k-1} + [\boldsymbol{K}]_k \cdot \left(\boldsymbol{U}_k - \boldsymbol{X}_{k-1}\right) \qquad k \geq 2 \\ [\boldsymbol{P}]_k &= [\boldsymbol{P}]_{k-1} - [\boldsymbol{K}]_k \cdot [\boldsymbol{P}]_{k-1} \end{aligned} \tag{36}$$

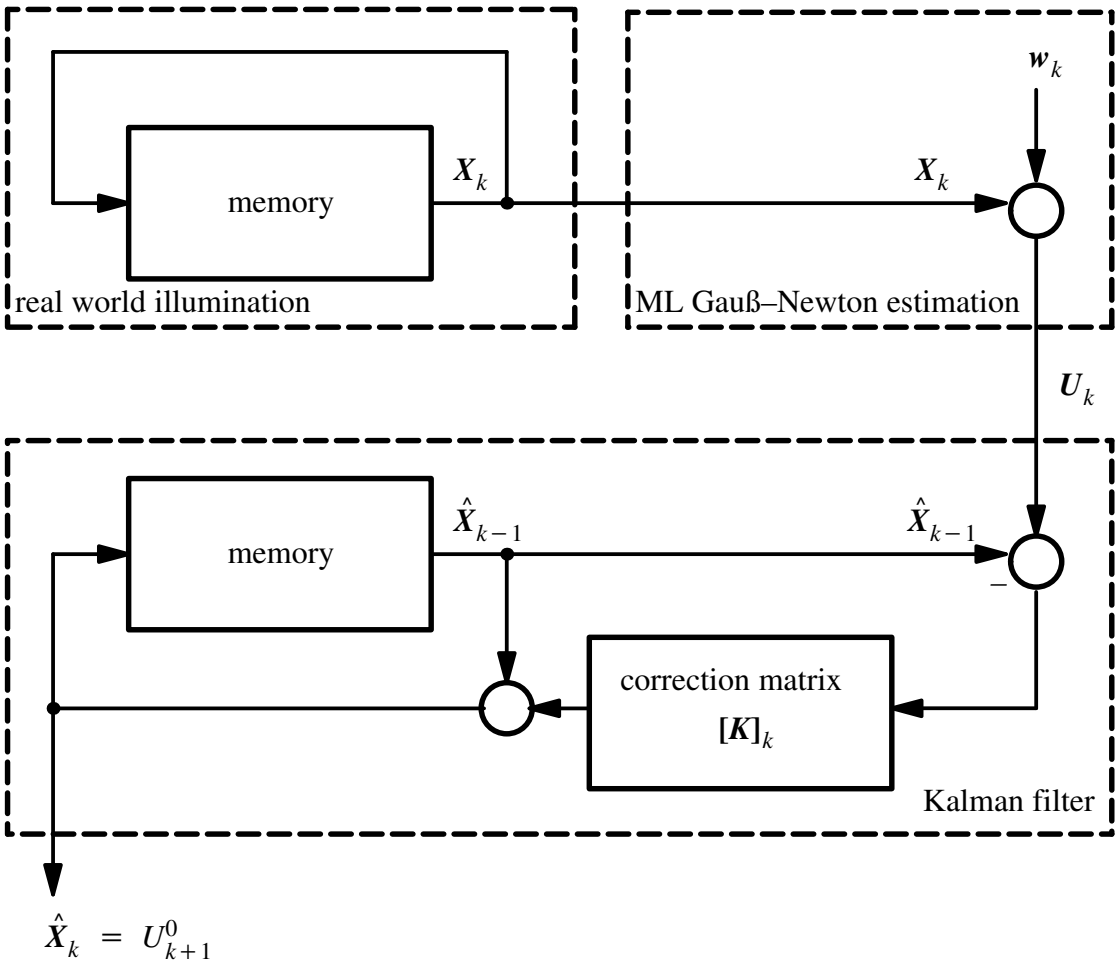


Fig. 8 Kalman estimation at time instant k: The true time–invariant illumination parameters $\boldsymbol{X}_k = \boldsymbol{X} = (\phi_{true}, \theta_{true}, c_{true})^T$ of the real world are observed by the described ML Gauß–Newton estimator resulting in the observed $\boldsymbol{U}_k$. The observation error is modelled as Gaussian uncorrelated zero–mean noise $\boldsymbol{w}_k$. The Kalman filter weights the new observation $\boldsymbol{U}_k$ by the correction matrix $[\boldsymbol{K}]_k$ and gives an estimation $\hat{\boldsymbol{X}}_k$ that is taken as start point $\boldsymbol{U}^0_{k+1}$ for the ML Gauß–Newton estimation at time instant k+1.

The Kalman filter is initialized by $[\boldsymbol{P}]_1 = [\boldsymbol{C}]_1$ and $\boldsymbol{X}_1 = \boldsymbol{U}_1$. There are some comments on taking Kalman estimates as start points for ML Gauß–Newton least square estimation. First,

in case of little motion or ill–conditioned illumination direction (parallel to rotation axis) the ML Gauß–Newton estimation fails, while the covariance matrix show high amplitudes. Looking at Eq. (36) makes clear, that in case of $\| [\boldsymbol{P}]_{k-1} \| \ll \| [\boldsymbol{C}]_k \|$ the bad ML Gauß–Newton estimate has little influence on the next start point. Thus, the illumination estimation is stabilized over the image sequence.

Secondly, due to erroneous shape and motion estimates, the true illumination parameters are often not those ones for minimum least square according to Eq. (30). Erroneous model object shape and motion can be compensated partly by a moving model illumination. Thus, the illumination parameters used in OBASC for model image synthesis are the ML Gauß–Newton estimates. The Kalman estimates inherit minimum error variance and thus may be closer to the true illumination parameters.

Finally, a Kalman filter can be used together with a dynamic model. This may help in future for more complicate scenes with moving light sources or a moving camera.

4.4 Overview on image analysis in an OBASC

The developed illumination estimation technique as well as the adapted motion estimation technique are embedded in the image analysis of Ostermann's OBASC scheme [14]. For image analysis, first, the model world is initialized using the first two images $s_{Y,0}$ and $s_{Y,1}$ of a real image sequence. Then, for each new image $s_{Y,k+1}$ $k+1 \geq 1$, the model world is updated by image analysis. In the real scene, objects are assumed to move in front of static background without occluding each other.

4.4.1 Model World Update

Fig. 9

Fig. 9 gives an overview on image analysis. Input to image analysis are the current model world W_k and the current image $s_{Y,k+1}$. First, a **change detection** mask is evaluated marking the corresponding image area of the current model objects *and* the current real objects that may have moved. This mask is required for shape estimation.

Secondly, the **3D motion** parameters of the objects and the **illumination** parameters of the scene are estimated as explained in Sections 4.2 and 4.3, respectively.

Then, the **3D shape** of the objects ist estimated [14]. Therefore, the **silhouettes** of the real objects are detected using the change detection mask, the estimated 3D motion parameters and the current shape of the model objects. Then, the control points of the model objects are moved such that the model object silhouettes get close to the silhouettes of the real objects.

Finally, **model failures** are detected. Model failures are areas of the model image, where the image quality is not sufficient enough according to a receiver model. This is done by change detection and some post processing [14]. For those model object surfaces that correspond to the model failures, new colour parameters are taken from the current image.

4.4.2 Model World Initialization

At the beginning, the model world W_0 is initialized. First, from image $s_{Y,0}$ and $s_{Y,1}$ a change detection mask is evaluated and interpreted as real object silhouettes. From these silhouettes,

the **shape** of the model objects is initialized using a distance transform assuming an ellipsoid–like shape [14]. The **colour** of the model objects is taken from the first image $s_{Y,0}$. The **illumination** is initialized to be diffuse, i.e. c=0. The **motion** of the objects do not have to be initialized.

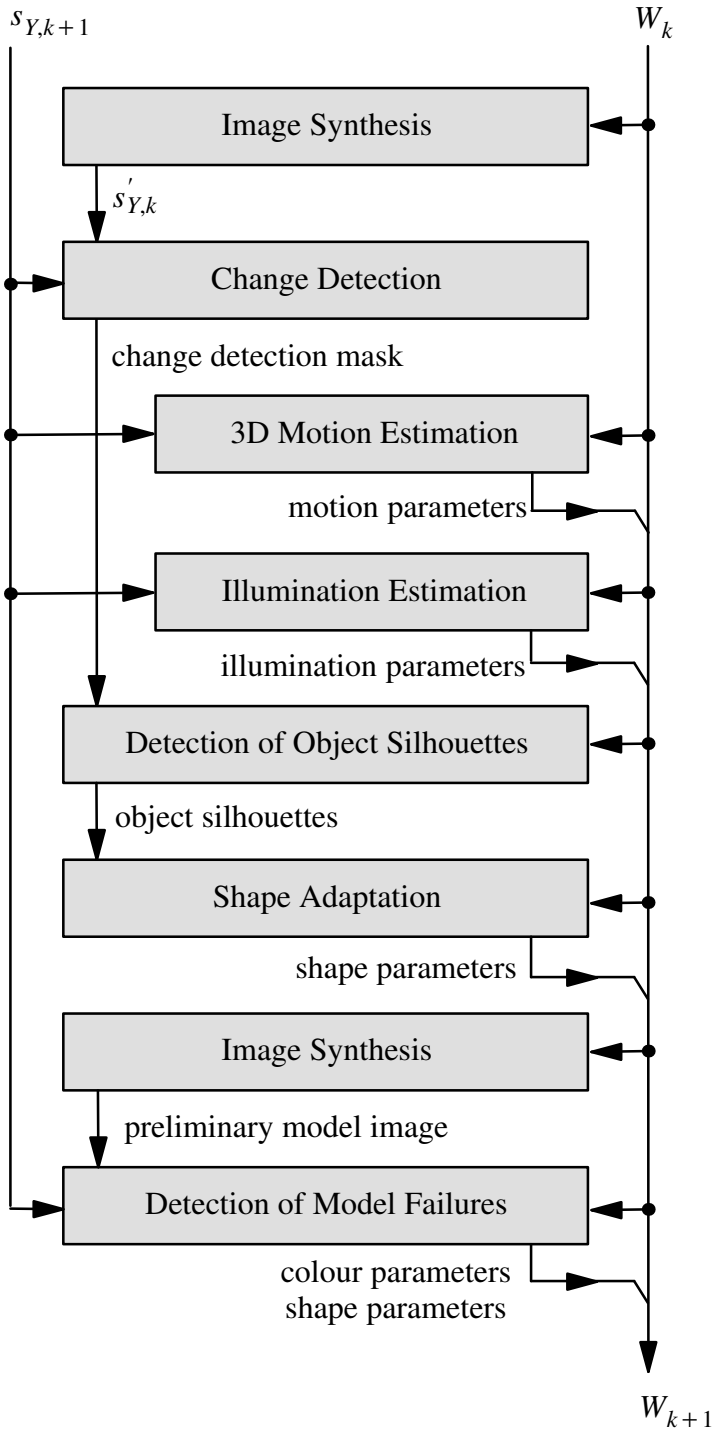


Fig. 9 Overview on image analysis: The current image $s_{Y,k+1}$ and the current model world W_k are used to estimate the motion, illumination, shape and colour parameters resulting to the updated model world W_{k+1}.

5. Results

In this chapter, experimental results using synthetically generated and real image sequences are presented. The experiments on synthetically generated image sequences focus on the robustness and performance of the proposed illumination estimation method depending on varying scene contents. Then, a first experiment on a real image sequence shows the performance of the image analysis of an OBASC scheme based on the new source model of moving rigid 3D objects with *non–diffuse* illumination (3Dn). The performance is compared to that of an image analysis based on Ostermann's source model of moving rigid 3D objects assuming *diffuse* illumination (3Dd).

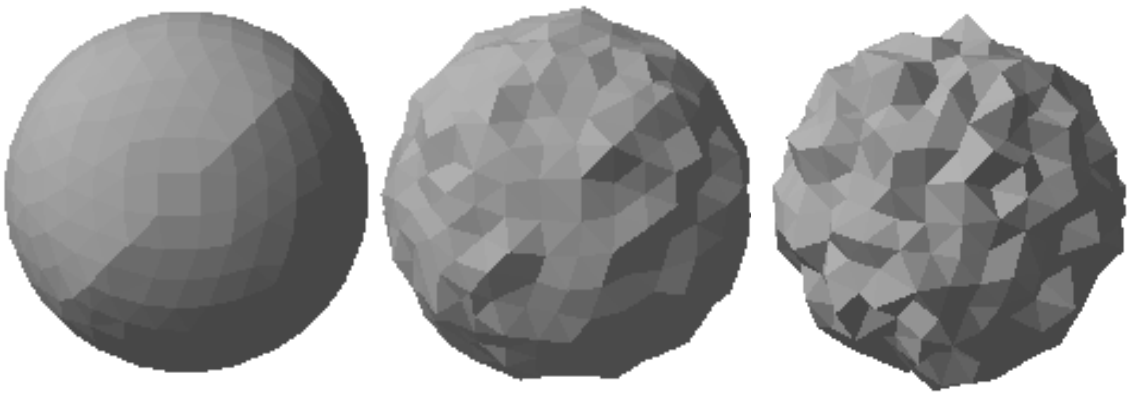
For hierarchical motion estimation, two sets of observation points, each about 10000, with gradients $\left|\boldsymbol{g}_{Y,0}(\boldsymbol{p}_0)\right| > 33$ are chosen. For the first set, a simple 3×3 averaging window is used as lowpass. For illumination estimation, about 3000 observation points with gradients $\left|\boldsymbol{g}_{Y,0}(\boldsymbol{p}_0)\right| < 50$ are chosen. A simple 5×5 averaging window is used as lowpass.

5.1 Experiments on synthetically generated image sequences

To examine the effect of varying scene contents (object shape, object motion and illumination) on the robustness and performance of the proposed illumination estimation method, several synthetic **image sequences** were generated. Each image sequence shows in 31 images a textured rigid ball under non–diffuse illumination. The illumination is defined by the direction $\boldsymbol{L}_{true} = (-\frac{2}{3}, -\frac{1}{3}, -\frac{2}{3})^T$ and the irradiances $c_{LAMBERT}$ and $c_{AMBIENT} = 1 - c_{LAMBERT}$. The ball rotates around all three axes up to a maximum angle $\alpha_{\max}$ away from its initial position. On average, the ball is $0.76 \cdot \alpha_{\max}$ degrees away from its initial position and rotates $0.33 \cdot \alpha_{\max}$ degrees per frame. For image sequence generation, the shape of the ball was derived from a reference shape shown in Fig. 10 (a). Therefore, the reference shape was first compressed in z–direction by the factor s_{global} (variation of global shape). Then, a stochastic distortion was applied (variation of local shape). The distortion is realized by a shift of each control point of the model object on the line through its position and the center of gravity $\boldsymbol{C}$. The shift amplitude is equally distributed between $\pm\, s_{local}$ percent of the ball radius. Fig. 10 shows some examples.

Fig. 10

To focus on the new illumination estimation method, a simplified **image analysis** is used here. The model object shape is initialized by the reference shape and left unchanged over the image sequence. The colour parameters are taken from the first image and are left unchanged as well. In case of Ostermann's source model of moving rigid 3D objects assuming *diffuse* illumination, for each image, motion parameters are estimated and the model object is motion compensated. In case of the new source model of moving rigid 3D objects with *non–diffuse* illumination, for each image, first motion, then illumination parameters are estimated. Further on, the model object is motion compensated.



(a) (b) (c)

Fig. 10 Examples for the shape of the model object "ball" used to generate the synthetical image sequences: (a) smooth reference shape (as assumed and used by analysis); (b) shape with local shape differences up to 5% of radius (as used for generation of nearly all synthetic image sequences); (c) shape with local shape differences up to 10% of radius.

To assess robustness and performance of the proposed illumination estimation technique and the new source model, two **criteria** are used. First, the accurateness of the estimated parameters c and $\boldsymbol{L}$ is assessed by the mean m_e of the relative parameter estimation error

$$e = \frac{\mid c\boldsymbol{L} - c_{true}\boldsymbol{L}_{true} \mid}{c_{true}} \tag{37}$$

Here, $c_{true} = \frac{c_{LAMBERT}}{c_{AMBIENT}}$ and $\boldsymbol{L}_{true}$ are the "true" parameters used for sequence generation. The mean error m_e is measured over the image sequence. This criterion is chosen, because it evaluates all illumination parameters by only one measure. Secondly, the MSE of the frame difference between the current image $s_{Y,k+1}$ and a model image $s'_{Y,k+1}$ generated from W_{k+1} is evaluated in the image area corresponding to the moving ball. The MSE is averaged over the image sequence. The MSE is measured for Ostermann's source model of moving rigid 3D objects (3Dd) assuming *diffuse* illumination and for the new source model of moving rigid 3D objects with *non–diffuse* illumination (3Dn). The source models are compared by a prediction gain according to

$$G = 10 \log \frac{MSE_{3Dd}}{MSE_{3Dn}} \tag{38}$$

Two more detailed criteria are used in a special case. These are the standard deviation σ_c of the estimated relative point light source irradiance and the mean of the angle $\measuredangle(\boldsymbol{L}, \boldsymbol{L}_{true})$ between true and estimated illumination direction.

5.1.1 Dependency on object shape

For illumination parameter estimation, it is assumed that the shape of the model objects is correct. Thus, the performance depends on the accurateness of the model object shape. In an

OBASC scheme, the model object shape is generated from an object silhouette by a distance transform assuming an ellipsoid–like shape. It is as smooth as the reference shape in Fig. 10 (a) that is used for image analysis here. The shape of the corresponding real object may differ in global and local shape.

To examinate the effect of differences in global shape between real and model object, a set of image sequences with $c_{LAMBERT}$=0.4, $\alpha_{\max}$=15^0, s_{local}=5% and varying s_{global} was generated and analyzed. As can be seen in Fig. 11, the mean parameter error is not much effected by global shape differences. The prediction gain g is prevailingly less if there are differences in global shape. For some cases the prediction gain is even higher in case of of global shape differences. In those cases, the global shape differences compensate the local shape differences of s_{local}=5% partly. These results show that the assumption of an ellipsoid–like shape as used by Ostermann's OBASC scheme for the initialization of the model object shape does not much disturb illumination estimation in case of convex smooth–shaped objects like a ball.

Fig. 11

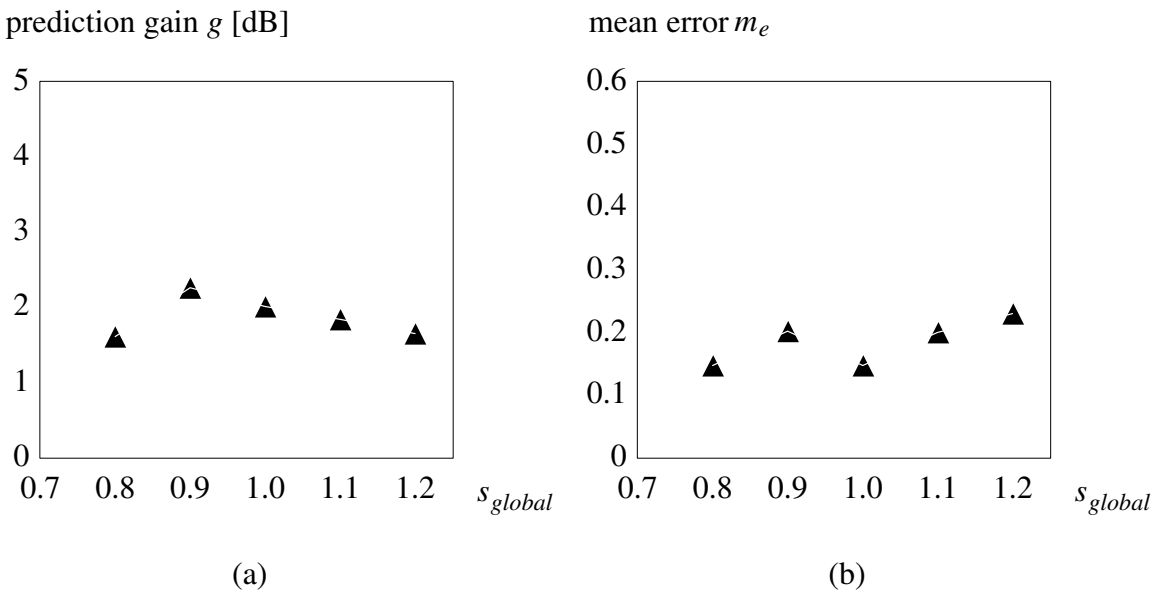


Fig. 11 Influence of global shape differences: Prediction gain g achieved by the illumination model of *non–diffuse* illumination and mean parameter error m_e depending on the global shape differences between real and model object. The analyzed image sequences were generated with $c_{LAMBERT}$=0.4, $\alpha_{\max}$=15^0, and s_{local}=5%.

To examinate the effect of differences in local shape between real and model object, a set of sequences with $c_{LAMBERT}$=0.4, $\alpha_{\max}$=15^0, s_{global}=1.0 and varying s_{local} was generated and analyzed. In an OBASC scheme, local shape differences may be due to a nose or chin of a person which is not contained in the corresponding model object. The results are shown in Fig. 12. With increasing differences, the prediction gain decreases and the mean parameter error of the illumination parameters gets larger. For local shape differences up to 10% of the object radius, the source model 3Dn has better performance than 3Dd. In case of more differences, the

Fig. 12

estimated illumination parameters get worse. The usage of the new source model (3Dn) in an OBASC scheme for coding of image sequences containing real objects with a lot details in shape (e.g. persons in a video telephone sequence) may be less effective.

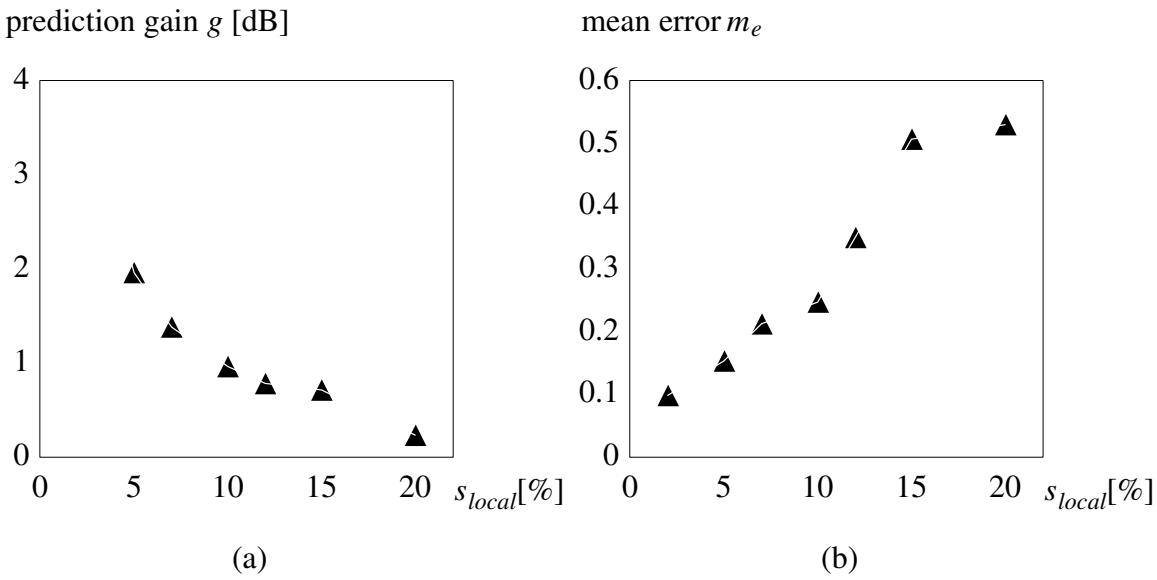


Fig. 12 Influence of local shape differences: Prediction gain g achieved by the illumination model of *non–diffuse* illumination and mean parameter error m_e depending on the local shape differences s_{local} between real and model object. The analyzed image sequences were generated with $c_{LAMBERT}$=0.4, $\alpha_{\max}$=15^0, and s_{global}=1.0.

5.1.2 Dependency on object motion

The photometric effects of non–diffuse illumination depend on both, object rotation and amount of relative point light source irradiance. To examine the effect of varying object motion, a set of image sequences with $c_{LAMBERT}$=0.4, s_{local}=5%, s_{global}=1.0 and varying $\alpha_{\max}$ was generated and analyzed. The results in Fig. 13 show clearly a higher prediction gain and a smaller mean parameter error the more the real object rotates. This behaviour is expectable and, for an OBASC, it is desirable, because the stronger photometric effects due to real object rotation the more effective is their successful estimation and compensation.

Fig. 13

5.1.3 Dependency on illumination

To examine the effect of varying point light source irradiance, a set of image sequences with $\alpha_{\max}$=15^0, s_{local}=5%, s_{global}=1.0 and varying $c_{LAMBERT}$ was generated and analyzed. The results in Fig. 14 show that the mean parameter error is not much affected by the amount of relative point light source irradiance $c_{LAMBERT}$. For very little point light source irradiance only, the estimation of the point light source direction gets worse, because the observable photome-

Fig. 14

tric effects due to real object rotation under non–diffuse illumination are too weak. Again, this behaviour is expectable and, for an OBASC, desirable. However, the prediction gain increases with increasing point light source irradiance.

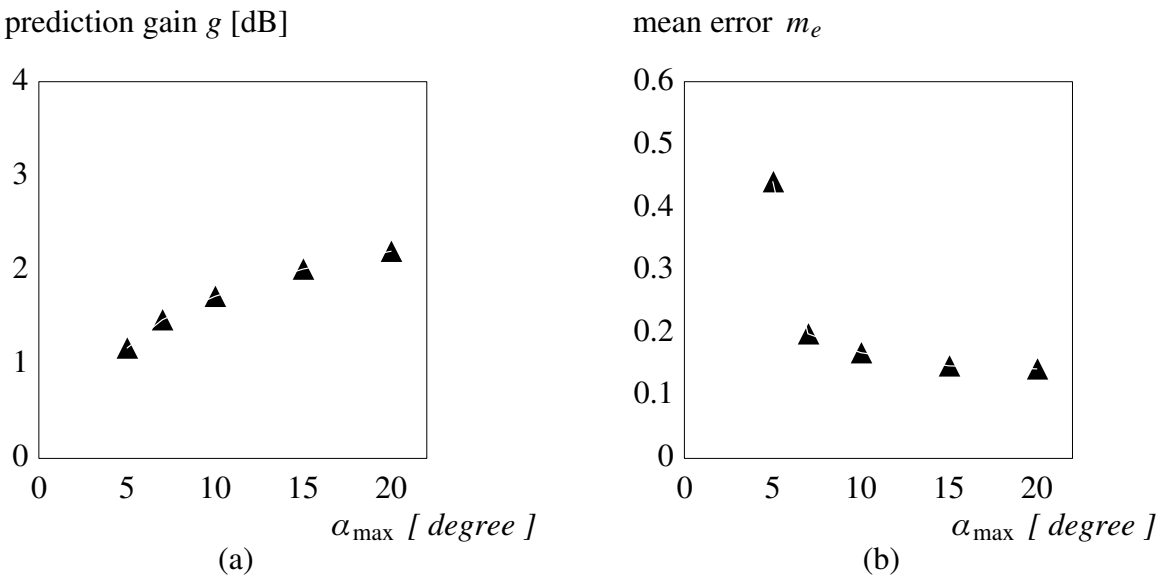


Fig. 13 Influence of amount of object rotation: Prediction gain g achieved by the illumination model of *non–diffuse* illumination and mean parameter error m_e depending on the amount of real object rotation defined by the angle $\alpha_{\max}$ of maximum rotation away from the initial position. The analyzed image sequences were generated with $c_{LAMBERT}$=0.4, s_{local}=5%, and s_{global}=1.0.

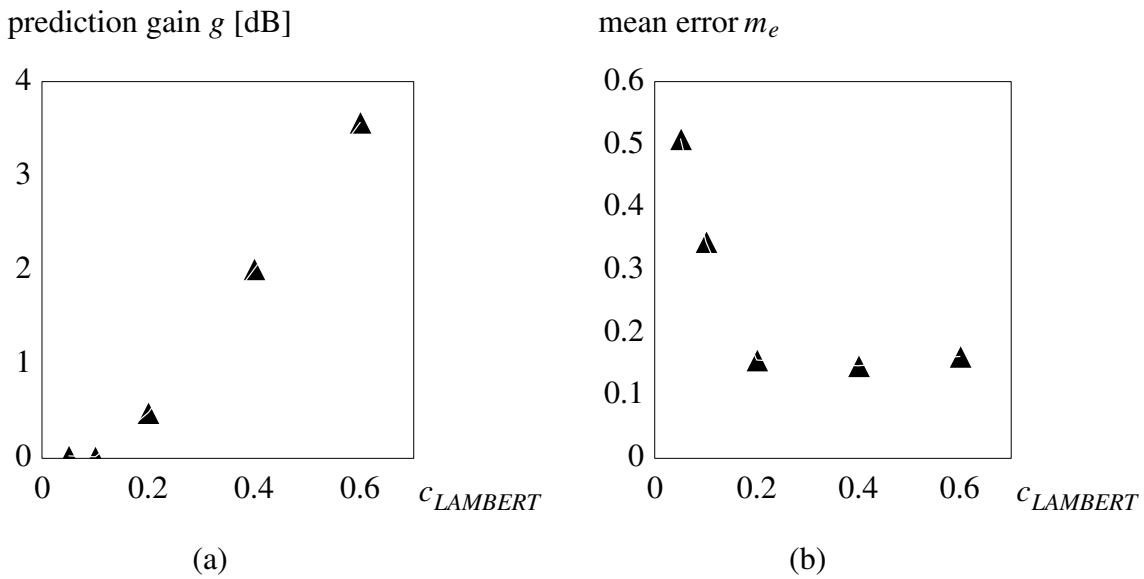


Fig. 14 Influence of relative point light source irradiance: Prediction gain g achieved by the illumination model of *non–diffuse* illumination and mean parameter error m_e depending on the relative point light source irradiance $c_{LAMBERT}$ of the real point light source. The analyzed image sequences were generated with $\alpha_{\max}$=15^0, s_{local}=5%, and s_{global}=1.0.

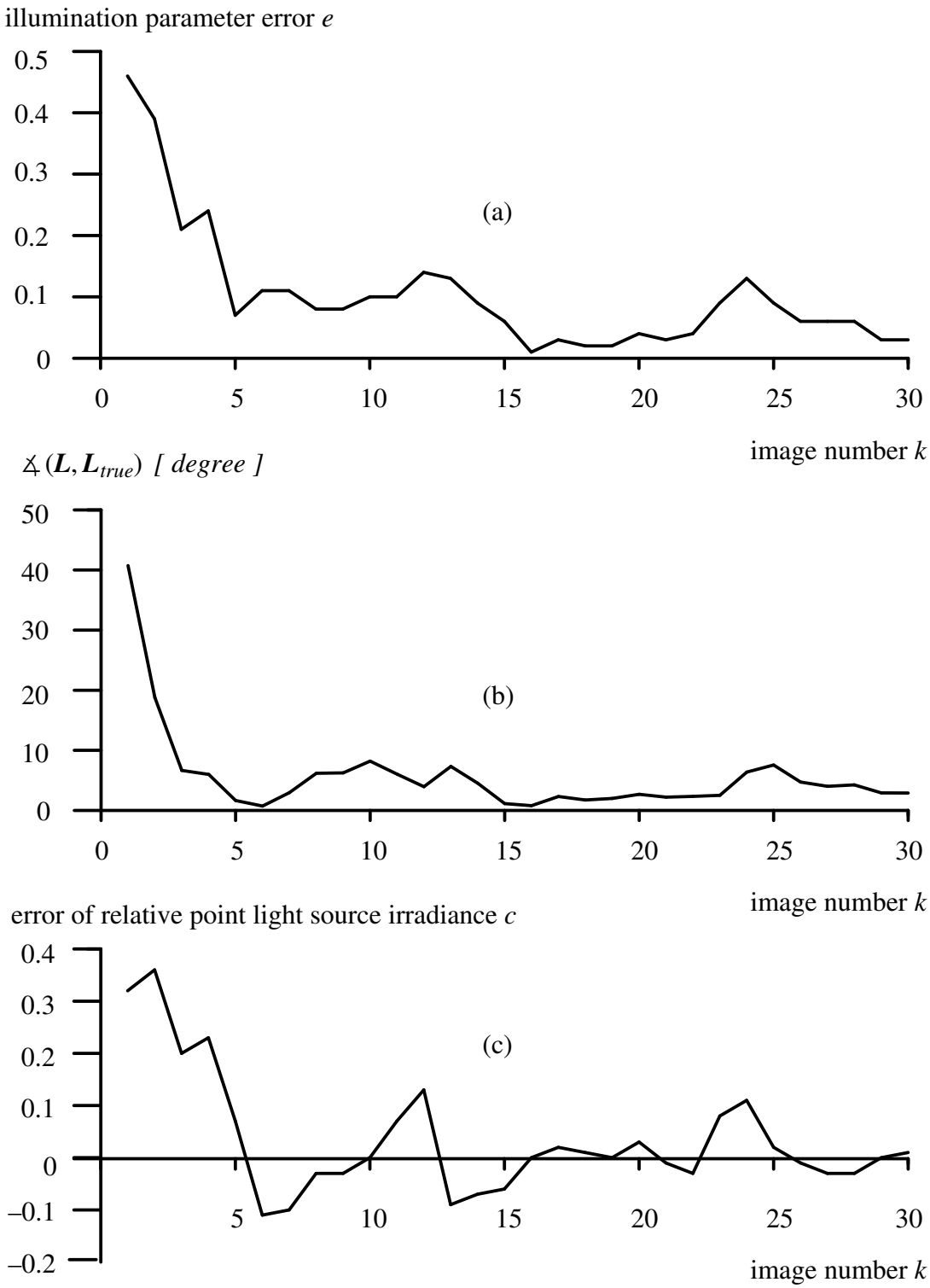


Fig. 15 Convergence of illumination parameters: (a) illumination parameter error e and (b) angle $\measuredangle(\boldsymbol{L}, \boldsymbol{L}_{true})$ between estimated and true illumination direction and (c) the error of the relative point light source irradiance over the time instant k. A convergence can be seen after the fivth image ($k=4$) of the image sequence. The analyzed image sequence was synthetically generated with $c_{LAMBERT}$=0.4, $\alpha_{\max}$=15^0, s_{local}=5%, and s_{global}=1.0.

5.1.4 Convergence speed of the illumination parameters

All of the above analyzed image sequences are variations of the case $c_{LAMBERT}$=0.4, $\alpha_{\max}$=15^0, s_{local}=5%, and s_{global}=1.0. In this case, the relative error e of the illumination parameters
Fig. 15 according to Eq. (37) is shown in Fig. 15 (a) versus time. As can be seen, there is a convergence of the illumination parameters after five images. In Fig. 15 (b) is shown that the mean of the angle $\measuredangle(\boldsymbol{L}, \boldsymbol{L}_{true})$ between true and estimated illumination direction is 3.8^0 averaged beginning with the sixth image of the sequence. In Fig. 15 (c) is shown, that the standard deviation σ_c of the estimated relative point light source irradiance is 5% of the total irradiance averaged from the sixth image on. The number of images necessary for convergence of the illumination
Table 1 parameters depend on the kind of real object motion. In Table 1 the position of the ball for each image of one of the synthesized image sequences is shown. It can be seen that during the first four images the ball rotates around the y–axis mainly. Not before the ball has rotated around more than two of three orthogonal axes the illumination parameter will converge.

	image 0	image 1	image 2	image 4	image 5	image 6
R_x	0	0	0	0	2	2
R_y	0	3	6	10	12	15
R_z	0	0	2	3	5	2

Table 1 Position of the ball for each image of one of the synthesized image sequences compared to the position for the first image (rotation angles in degree).

Although the estimated illumination parameters do not reach the true ones, in most cases, the prediction gain g is higher if the estimated parameters are applied instead of the true ones. This is due to erroneous object shape and motion parameters which can be compensated partly by erroneous illumination parameters.

The computational load for illumination parameter estimation is less than for motion estimation. 12 iterations are sufficient only, whereas motion estimation needs about 20 iterations. Only one or two illumination parameters per iteration in contrast to 6 motion parameters have to be estimated.

5.2 Experiments on real image sequences

To examine the performance of the image analysis of an OBASC scheme based on the new source model of moving rigid 3D objects with *non–diffuse* illumination (3Dn), see Fig. 9, the real **image sequence** ”Ball” was analyzed, see Fig. 17 (a) and (b). The sequence has a length of 31 images, CIF format, and 10Hz frame rate. It shows a textured ball rapidly moving (about 10^0 rotation per frame) under ambient and spot light. The ball covers about 20% of the image area. Due to the rotation and translation of the ball, a lot of new scene contents gets visible

during the image sequence. This is uncovered background and parts of the ball's surface that is visible first time. To assess the performance of an OBASC scheme based on the new source model 3Dn, the size of model failures is measured. For comparison, the size of model failures for Ostermann's source model [14] of moving rigid 3D objects (3Dd) assuming *diffuse* illumination is measured. In these experiments, the colour parameters are PCM coded, all other parameters are not coded.

Fig. 16 The results in Fig. 16 show the size of model failures in percent of the image area over the image sequence. In case of the new source model, 6.7% of the image area were found to be model failures, whereas in case of Ostermann's source model, the size is 9.9%. The image quality of the model images was held fixed at 33 dB PSNR.

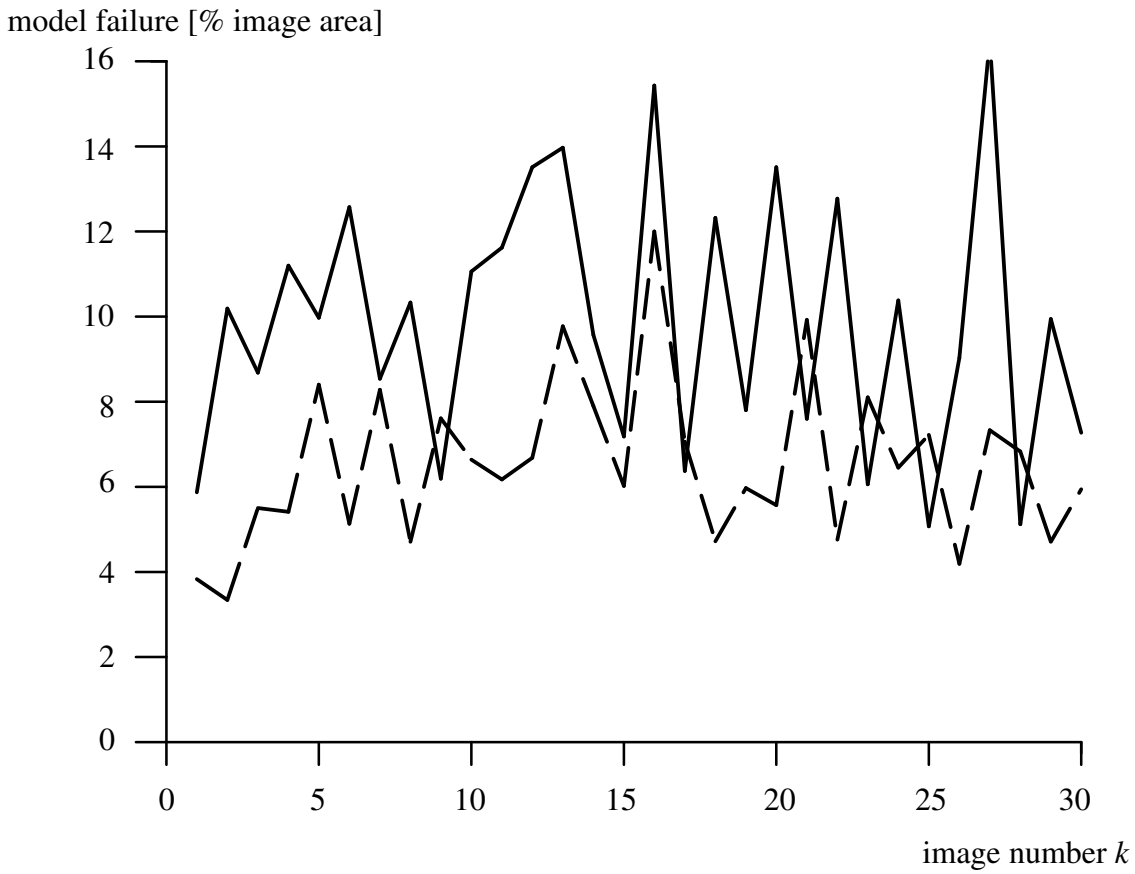


Fig. 16 Model failures in percent of the image area analyzing the real image sequence "ball" based on the source model 3Dd assuming *diffuse* illumination (solid line) or based on the source model 3Dn considering *non–diffuse* illumination (dashed line).

Fig. 17 In Fig. 17 a typical example for model failures is shown. Three types of model failures can be seen: Model failures due to object rotation under non–diffuse illumination occur in areas, where a maximum of luminance variations can be observed. This area is about 3.2% of the image area

and can be identified in Fig. 17. These model failures are drastically reduced if the new source model is applied. The second type of model failures is due to new image contents. It occurs on object surfaces that can be seen the first time. In this image sequence, 2.3% of the image area are new scene contents due to uncovered background. About 0.5% of the image area are new scene contents due to object rotation. The model failures caused by object rotation can expected to be higher. This is because colour parameters on formerly tilted model object surfaces are used for image synthesis after model object rotation. This effect is examined by Pearson [15]. The third type of model failures are due to the wrong model object shape. They occur at the border of an object and near colour edges on the object surface.

In about 4% of the image area, Ostermann's OBASC scheme exhibits model failures coding the real image sequence "Claire", CIF format, 10Hz frame rate, at a bit rate of approximately 64kbit/s [14]. Coding of colour and shape parameters of these model failures takes more than 70% of the bit rate. Thus, the reduction of size of model failures from 9.9% to 6.7% of the image area measured for the real image sequence "Ball" will markedly reduce the bit rate for coding this image sequence.

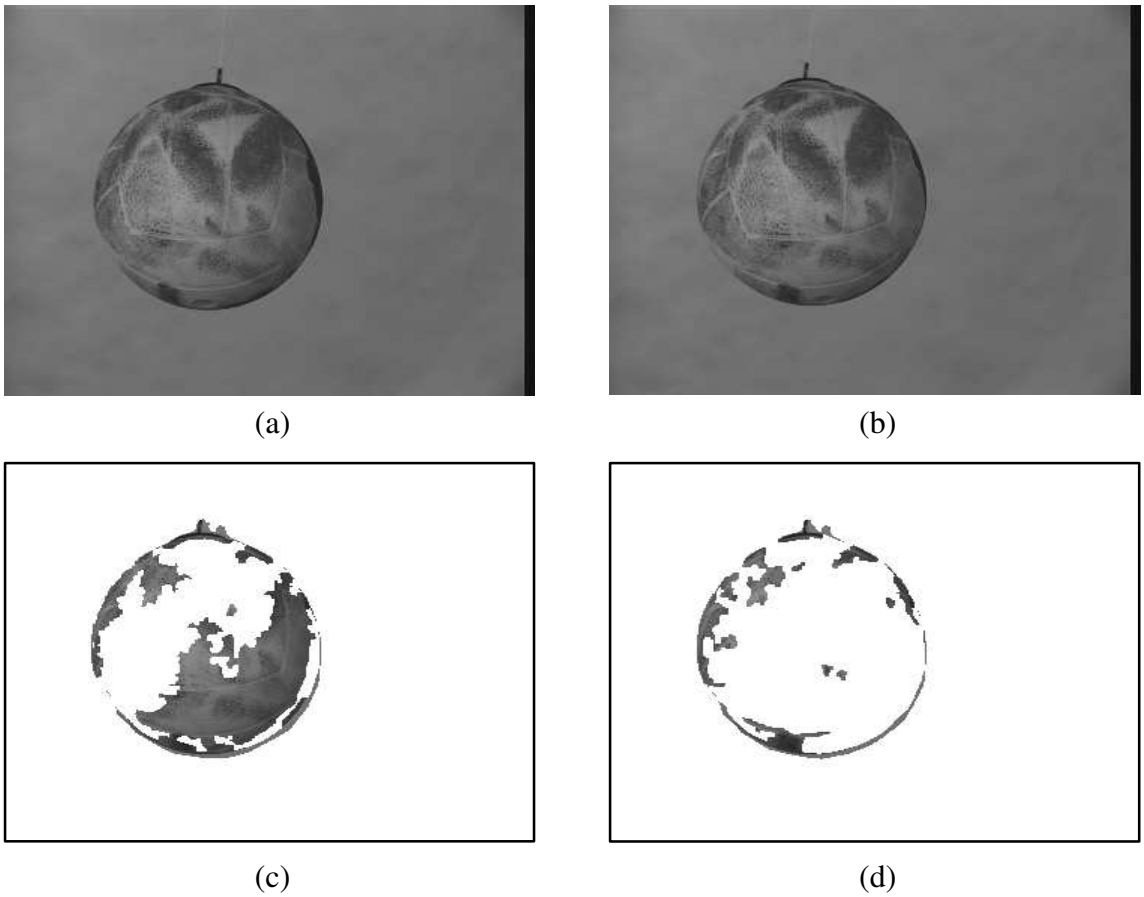


Fig. 17 Real image sequence "ball": (a) sample image $k=1$ and (b) sample image $k=2$; model failures found by image analysis at time instant $k=2$ (c) based on source model 3Dd assuming *diffuse* illumination or (d) based on source model 3Dn considering *non–diffuse* illumination. It can be seen that model failures due to photometric effects are reduced. The remaining model failures are mainly due to new image contents on the moving ball (left border of the ball), uncovered background (bottom right border of the ball), or wrong object shape (other regions).

6. Conclusion

For transmission of moving images at very low bit rates, Object–Based Analysis–Synthesis Coding (OBASC) based on the source model of moving rigid 3D objects [14] has been proposed. The inherent assumption of *diffuse* illumination leads to model failures due to photometric effects of object rotation under *non–diffuse* illumination.

In this paper, the source model of moving rigid 3D objects of an Object–Based Analysis–Synthesis Coder (OBASC) is extended from *diffuse* to *non–diffuse* illumination by an explicit illumination model. The applied illumination model consists of a distant point light source and ambient diffuse light. The illumination is described by two parameters defining the direction of the point light source and one for its relative irradiance compared to the total irradiance of point source and ambient light. It is shown that the definition of the illumination model by only three parameters is sufficient for establishing an irradiance compensated image synthesis in an OBASC scheme.

An algorithm for estimation of the illumination parameters is presented. The method assumes that shape and motion of the real objects have already been estimated. Object motion and illumination have a common impact on a real image sequence. Inspite of that, in the real image sequence, it is shown that photometric effects are weaker than geometric effects of object motion under *non–diffuse* illumination. Thus, illumination is estimated after motion estimation and compensation. For illumination estimation, a fast iterative Maximum–Likelihood Gauß–Newton estimation is carried out assuming Gaussian noise in the image. For each image of an image sequence, the estimation method minimizes the frame difference between the real image and a model image generated from the model world and thus ensures small model failures. As a start value for the first iteration, the Kalman filtered illumination parameters of the predecessing images are used.

Illumination estimation experiments on synthetically generated image sequences were carried out. The sequences show a moving textured rigid ball illuminated by a point light source and ambient light. The experiments examine the influence of the amount of real object motion, the relative irradiance of the real point light source, as well as differences between real and model object shape on the accuracy of the estimated illumination parameters. The accuracy of the illumination parameters is evaluated by the standard deviation of the estimation error of the relative point light source irradiance as well as by the mean of the inherently positive angle between estimated and true illumination direction. The experiments show that the accuracy increases with an increasing amount of real object rotation. The accuracy is almost independent of the relative irradiance of the real point light source. In case of a mean of 10^0 real object rotation per frame and 40% relative point light source irradiance, the illumination parameters converge after about five images. From the sixth image on, the mean error of the estimated illumination direction is about 4^o and the standard deviation of the estimated relative point light source irradiance is 5%.

The illumination estimation method was embedded in the image analysis of an OBASC scheme. Shape and motion is estimated as in [14], but motion estimation was adapted to the illumination model. The image synthesis was extended by an irradiance compensation.

The OBASC scheme was applied to a real image sequence showing a rigid textured ball translating and rapidly rotating about 10^0 per frame under spot and ambient light. The sequence has CIF format and 10Hz frame rate. The ball covers 20% of the image area. The colour parameters are PCM coded, all other parameters are not coded. As a quality criteria, the size of model failures was used. Comparing the source models of moving rigid 3D objects with *diffuse* and with *non–diffuse* illumination, a reduction of model failures from 9.9% to 6.7% of the image area, respectively, can be achieved at an SNR of 33dB. Thus, 3.2% model failures due to temporal luminance variations on the rotating ball under non–diffuse illumination are eliminated. The remaining 6.7% are due to uncovered background (2.3%) and errorneous model object shape or new scene contents on the moving model object (4.4%). It can be expected that a reduction of the size of model failures from 9.9% to 6.7% of the image area will markedly reduce the bit rate necessary for coding this image sequence.

Acknowledgements

The author thanks Prof. H.–G. Musmann for his support to this project and discussions about object–based coding. Further on, the author thanks Dr.–Ing. J. Ostermann for a lot of discussions about image analysis and, in particular, for his complete software implementation of an object–based analysis–synthesis coder. This research was supported by the German Federal Government.

7. References

[1] K. Brammer, G. Siffling, *Kalman–Bucy–Filter – Deterministische Beobachtung und stochastische Filterung*, Oldenbourg, München, 1985, Chapter 2, pp. 60ff.

[2] T.J. Broida, R. Chellappa, "Estimating the kinematics and structure of a rigid object from a sequence of monocular images", *IEEE transactions on pattern analysis and machine intelligence*, Vol. 13, No. 6, June 1991, pp. 497–512.

[3] CCITT, *Draft revision of recommendation H.261: video codec for audio visual services at $p \times 64$ kbit/s*, Study Group XV, WP /1/ Q4, Specialist group on coding for visual telephony, Doc. No. 584, November 1989.

[4] B. K. P. Horn, *Robot vision*, MIT Press, Cambridge, 1986, Chapter 10.3, pp. 206ff.

[5] M. Hötter: "Object–oriented analysis–synthesis coding based on moving twodimensional objects", *Signal Processing: Image Communication*, Vol. 2, No. 4, December 1990, pp. 409–428.

[6] A. Kattner: "*Verfahren zur Schätzung von Beleuchtung aus Bewegtbildsequenzen mit Hilfe eines Kalman–Filters*", diploma thesis, Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, University of Hannover, Hannover, 1994.

[7] G. J. Klinker, S. A. Shafer, T. Kanade, "The measurement of highlights in colour images", *Int'l Journal of Computer Vision*, Vol. 8, No. 2, June 1988, pp. 7–32.

[8] J. H. Lambert, *Lambert's Photometrie (Photometria sive de mensura de gratibus luminis, colorum et umbrae)*, Wilhelm Engelmann, Leipzig, Germany, 1892.

[9] K. Y. Lee, D. E. Pearson, "Luminanace compensation in model–based–prediction", *Picture Coding Symposium*, Tokyo, Japan, 2–4 September 1991, pp. 275–276.

[10] C. R. Moloney and E. Dubois, "Estimation of motion fields from image sequences with illumination variation", *ICASSP*, Toronto, Canada, 14–17 May 1991, Vol. 4, pp. 2425–2428.

[11] N. Mukawa, "Estimation of light source information from image sequence", *Systems and Computers in Japan*, Vol. 23, No. 10, October 1992, pp. 92–99.

[12] H. G. Musmann and M. Hötter and J. Ostermann, "Object–oriented analysis–synthesis coding of moving images", *Signal Processing: Image Communication*, Vol. 1, No. 2, November 1989, pp 117–137.

[13] S. K. Nayar, K. Ikeuchi, T. Kanade, "Surface Reflection: Physical and geometrical Perspectives", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 7, July 1991, pp. 611–634.

[14] J. Ostermann, "Object–based Analysis–synthesis Coding (OBASC) based on the source model of moving rigid 3D objects", *Signal Processing: Image Communication*, Vol. 6, No. 2, May 1994, pp. 143–161.

[15] D. E. Pearson, "Texture mapping in model–based image coding", *Signal Processing: Image Communication*, Vol. 2, No. 4, December 1990, pp. 377–395.

[16] A. P. Pentland, "Finding the illuminant direction", *Journal of the Optical Society of America*, Vol. 72, No. 4, April 1982, pp. 448–455.

[17] A. P. Pentland, "Photometric motion", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 9, September 1991, pp. 879–890.

[18] B. Phong, "Illumination for computer generated pictures", *Communications of the ACM*, Vol. 18, No. 6, June 1975, pp. 311–317.

[19] L.E. Scales, *Introduction to non–linear optimization*, Macmillan, London, 1985, Kap. 4, pp. 114–115.

[20] J. Stauder, "Estimation of nonuniform irradiance for object–oriented coding based on 3D objects", *Picture Coding Symposium*, Lausanne, Switzerland, 17–19 March 1993, pp. 18.2.1–18.2.2.

[21] J. Stauder, "An illumination estimation method for 3D–Object–based Analysis–synthesis Coding", *COST 211 ter European Workshop on New Techniques for*

Coding of Video Signals at Very Low Bitrates, Hannover, Germany, 1–2 December 1993, pp. 4.5.1–4.5.6.

[22] J. Stauder, "Joint estimation of object motion and point light illumination for 3D–object–based Analysis–synthesis Coding", *Int'l Workshop on Coding Techniques for Very Low Bit–rate Video*, Colchester, United Kingdom, 7–8 April 1994, pp. 4.1.1–4.1.4.

[23] J. Stauder, "Detection of Highlights for Object–Based Analysis–Synthesis Coding", *Picture Coding Symposium*, Sacramento, USA, 21–23 September 1994.

[24] G.J. Ward, "Measuring and modelling anisotropic reflection", *Computer Graphics*, Vol. 26, No. 2, July 1992, pp. 265–272.

[25] Q. Zheng and R. Chellappa, "Estimation of illuminant direction, albedo, and shape from shading", *IEEE Transactions on Pattern Recognition and Machine Intelligence*, Vol. 13, No. 7, July 1991, pp. 680–702.