Global–Shape from Motion and Shading for 3D–Object–Based Analysis–Synthesis Coding

Jürgen Stauder
Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung
Universität Hannover, Appelstraße 9A, D–30167 Hannover 1, FRG
Phone: +49–511–762–5311
Email: stauder@tnt.uni–hannover.de

Abstract

In OBASC based on the source model "moving textured rigid 3D objects with non–diffuse illumination" the shape of model objects describing real objects is estimated from the silhouettes of the real objects assuming an ellipsoid–like shape. In this paper, the image analysis is extended by a global shape estimation method evaluating directly the image signal. By one global shape parameter, the depth of a model object is scaled in direction of the lines of sight. The extension of OBASC image analysis reduces the data rate necessary for a test sequence showing a textured, rapidly moving ball from 107 kbit/s to 101 kbit/s.

1 Introduction

For object–based analysis–synthesis coding (OBASC) [5] of image sequences at low data rates, an implementation has been proposed where the image sequence is described by a model world consisting of moving 3D model objects [7] and non–diffuse model illumination [9][10]. The model objects are defined by colour parameters describing the reflectance of a model object surface, motion parameters describing 3D rotation and translation and shape parameters describing a 3D wireframe, which is put up by I vertices $\boldsymbol{P}^{(i)} = (P_x^{(i)}, P_y^{(i)}, P_z^{(i)}), i{=}0..I-1$. The model illumination is defined by illumination parameters describing a distant point light source and ambient diffuse light.

In [7][10], for shape estimation two steps are applied: shape initialization and shape adaptation. By shape initialization, the shape of a model object is derived from the silhouette of the corresponding real object at time instant $k{=}0$ assuming an ellipsoidal shape with a fixed depth–to–width–ratio of 1.5 [7]. By shape adaptation, the shape of the model object is adapted to the silhouette of the corresponding real object at time instant k>0 for each image k>0 of the image sequence [7].

To enhance shape adaption, the image signal may be used directly instead of using object silhouettes. Koch [3] suggested a shape from motion method for shape adaption. First, a model object is moved according to estimated motion parameters [7]. Then, for shape adaption at time instant $k{>}0$, the vertices of the model object are moved along the line of sight using shape parameters $s^{(i)}$ according to

$$\boldsymbol{P}_{k+1}^{(i)} = \boldsymbol{P}_k^{(i)} + s^{(i)} \cdot \boldsymbol{V}_k^{(i)} \ . \tag{1.1}$$

The line of sight vectors $\boldsymbol{V}_k^{(i)}$ have unit length, are fixed to the model objects at time instant $k{=}0$ and are moved together with the model object. The shape parameters are estimated minimizing a cost function by Gauss–Newton method [8]. The cost function evaluates the frame difference between a model image generated from the model world and the corresponding real image. The frame difference is evaluated only at a set of image positions with high spatial image luminance gradients [7] inside an image region corresponding to the neighbourhood of $\boldsymbol{P}_k^{(i)}$. The frame difference is approximated by a linear image luminance model [10].

Falkenhagen [1] reduces the shape parameters to one global shape parameter

$$s = s^{(i)} \quad i{=}0..I{-}1 \tag{1.2}$$

in Eq. (1.1). Therefore, he presets the length of the line of sight vectors $\boldsymbol{V}_k^{(i)}$ at time instant $k{=}0$ to

$$|\boldsymbol{V}_0^{(i)}| = \left| P_{0,z}^{(i)} - \frac{1}{I}\sum_{j=0}^{I-1} P_{0,z}^{(j)} \right| \tag{1.3}$$

with z–direction as viewing direction. Thus, the depth of the object can be scaled by s. For estimation of s, the cost function of Koch [3] is evaluated at a set of image positions with high spatial image luminance gradients inside the image region corresponding to the model object. The cost function is minimized by Gauss–Newton method.

Koch and Falkenhagen assume diffuse illumination. Lee and Kuo [4] suggested a shape from shading approach for shape adaption assuming an illumination by a distant point light source. For shape from shading, the reflectance of the object surface has to be known [2]. Thus, in [4], the reflectance is assumed to be spatially constant all over the object surface.

In this contribution, the shape of a model object is adapted by one global shape parameter s as in [1]. To take account of shading clues, the cost function from [3] will be adapted to the illumination model of OBASC [10]. By the illumination model, the restriction of spatially constant reflectance of [4] can be overcome. Therefore, the reflectance of each point on the model object surface is derived from a real image eliminating the influence of illumination using the illumination model.

In Section 2, the applied source model defining the model world is summarized. Based on the source model, in Section 3 an optical flow equation is set up. Section 4 describes the global shape estimation method. In Section 5, results for OBASC coding efficiency are given.

2 Source model

The source model describes the real image sequence and consists of an illumination model, an object model and a camera model. The illumination model describes the irradiance E (power per area) incident on the surface of an object by a distant point light source and ambient diffuse light:

$$E(\boldsymbol{N}) = c_{AMBIENT} + c_{LAMBERT}\ \max\{0, \boldsymbol{LN}\} \tag{2.1}$$

Here, $\boldsymbol{N}$ is the object surface normal at a specific point on the object surface, $\boldsymbol{L}$ the illumination direction, $c_{AMBIENT}$ the ambient irradiance and $c_{LAMBERT}$ the point light source irradiance [9][10].

The object model describes the real objects by coloured wireframes. The shape can be adapted according to Eqs. (1.1) and (1.2), the position can be changed according to

$$\boldsymbol{P}_{k+1}^{(i)} = [\boldsymbol{R}] \cdot (\boldsymbol{P}_k^{(i)} - \boldsymbol{C}) + \boldsymbol{C} + \boldsymbol{T} \tag{2.2}$$

with $[\boldsymbol{R}]$ a 3D rotation matrix, $\boldsymbol{T}$ a 3D translation vector and $\boldsymbol{C}$ the mean of $\boldsymbol{P}^{(i)}$ [7]. The colour parameters of a point $\boldsymbol{O}$ on the surface of a model object are taken from image $k=0$ at the image position $\boldsymbol{p}_0$ as shown for the luminance parameter $s_{Y,0}(\boldsymbol{p}_0)$:

$$s_{Y,0}(\boldsymbol{p}_0) = t_Y(\boldsymbol{O})\ \ E(\boldsymbol{N}_0) \tag{2.3}$$

with $s_{Y,0}$ the luminance of image $k=0$, $t_Y(\boldsymbol{O})$ the effective luminance reflectance of $\boldsymbol{O}$ [10] and $\boldsymbol{p}_0$ the corresponding 2D image position of the object surface point $\boldsymbol{O}$ at time instant $k=0$.

The camera model describes the geometric camera properties by perspective projection, and the photometric properties are assumed to be linear [10].

3 Optical flow equation

For estimation of s, the frame difference between model image and real image is evaluated for a set of J points in the image plane. The corresponding points on the surface of the model objects are called observation points. Each observation point $\boldsymbol{O}^{(j)} = (\boldsymbol{P}_k^{(j)}, \boldsymbol{N}_k^{(j)}, \boldsymbol{N}_0^{(j)}, s_{Y,0}(\boldsymbol{p}_0^{(j)}))$, $j=0..J-1$ is defined by its current 3D position $\boldsymbol{P}_k^{(j)}$, the current surface normal $\boldsymbol{N}_k^{(j)}$, the surface normal $\boldsymbol{N}_0$ at time instant $k=0$ and the luminance parameter $s_{Y,0}(\boldsymbol{p}_0^{(j)})$. For each observation point, an optical flow equation is set up:

$$f^{(j)} = s_{Y,0}(\boldsymbol{p}_0^{(j)}) \cdot \frac{E(\boldsymbol{N}_{k+1}^{(j)})}{E(\boldsymbol{N}_0^{(j)})} - s_{Y,k+1}(\boldsymbol{p}_k^{(j)}) - \boldsymbol{g}_Y^{(j)} \left(\boldsymbol{p}_{k+1}^{(j)} - \boldsymbol{p}_k^{(j)}\right) = 0 \qquad (3.1)$$

with $\boldsymbol{g}_Y^{(j)}$ the spatial image luminance gradient of a linear image luminance model [10]. In Eq. (3.1), the unknown effective luminance reflectance $t_Y(\boldsymbol{O}^{(j)})$ of the observation point is calculated individually by inverting Eq. (2.3). The Eqs. (3.1) are combined to a cost function

$$F = \sum_{j=0}^{J-1} f^{(j)\,2} \rightarrow MIN \quad . \qquad (3.2)$$

The cost function has to be minimized with respect to the unknowns. The unknowns are $([\boldsymbol{R}],\boldsymbol{T})$ for motion estimation [7], $(\boldsymbol{L}, c_{AMBIENT}, c_{LAMBERT})$ for illumination estimation [10], or s for global shape estimation.

4 Global shape estimation method

The global shape estimation consists of two steps: global shape initialization and global shape optimization. For global shape initialization, an initial value for s is found by a systematic search in the interval $S=[-0.4,0.4]$, i.e. the depth of a real object is assumed to be in between 60% and 140% of the depth of the corresponding model object. Therefore, for each model object, M equally spaced values $s_m \in S, m=0..M-1$, are applied temporally to the model object. For each of the M scaled model objects, motion is estimated and temporally compensated minimizing Eq. (3.2) by Gauss–Newton method [8] using observation points with high spatial image luminance gradients [10]. For each of the M scaled and motion compensated model objects, the luminance frame difference between model image and real image is evaluated, denoted as D_m. s_m is taken as initial value for s if $D_m = \min\{\ D_n\ ,\ n=0..M-1\ \}$. M is set to 8.

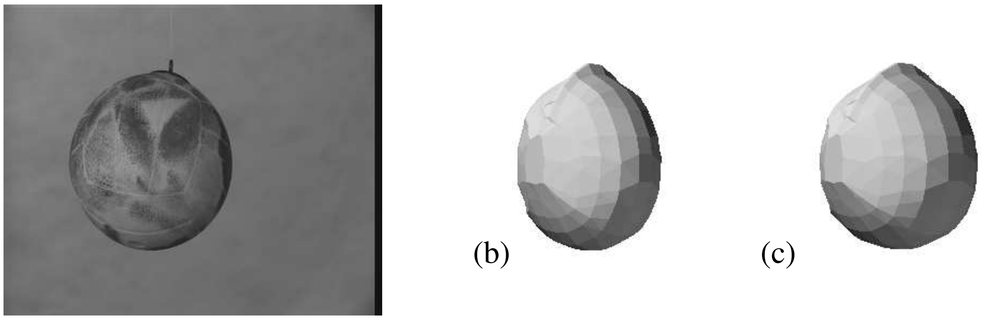
For global shape optimization, Eq. (3.2) is minimized by Downhill–Simplex method [6] using observation points with low spatial image luminance gradients. After global shape optimization, for each model object, again motion is estimated and compensated.

The global shape estimation is embedded in the image analysis of the considered OBASC scheme as follows: For each image of the image sequence, first motion is estimated and compensated, secondly illumination is estimated, then global shape is estimated and compensated, and finally shape is adapted to the object silhouettes. The order results from the fact that motion estimation was found to be not much sensible to erroneous model object shape neither to model illumination, and illumination estimation to be not much sensible to erroneous global–shape of the model object [10].

5 Results

OBASC simulations with and without global shape estimation has been carried out. In OBASC, after image analysis, motion and shape parameters parameters are quantized uniformly, entropy–coded and transmitted [7]. Colour parameters are transform–coded [7] and transmitted only in regions of model failures. In regions of model failures the real image cannot be described accurately enough by the model world. The other image regions are synthesized using formerly transmitted colour parameters of the model objects. Since the transmission of colour parameters is expensive by means of data rate, the OBASC aims at small sized model failure regions. The first frame is PCM coded.

OBASC coding efficiency was measured for the test sequence "Ball" (CIF, 10 Hz, 30 frames) showing a textured, rapidly moving ball (about 10 degrees rotation per frame) covering about 20% of the image area, see Fig. 5.1. Without global shape estimation, the size of model failure regions and regions of first visible texture (corresponding to uncovered background and rotated object surfaces) turned out to be 6.7% of the image area as measured in [10] resulting to a data rate of 107 kbit/s. By application of the global shape estimation method, the size of model failure and first visible texture regions can be reduced to 6.3% by applying the global shape estimation method. This corresponds to a reduction of the data rate to 101 kbit/s.



(a) (b) (c)

Figure 5.1: First frame of the test sequence "Ball" and shaded shapes of the model object "Ball" (45 degrees rotated for visualization) after analysis of 2 images (b) without and (c) with global shape estimation method.

6 Conclusions

The image analysis of an OBASC scheme based on the source model "moving textured rigid 3D objects with non–diffuse illumination" was extended by a global shape estimation method. Starting from an initial model object shape, one global shape parameter is estimated and applied to the model object. The global shape parameter scales the depth of the model object in direction of the lines of sight of the model camera. For estimation of the global shape parameter, first an initial value is found by a systematic search using the luminance frame difference between model and real image as criterion. Then, the global shape parameter is optimized by the Downhill–Simplex method using the frame difference at image points of low spatial image luminance gradient as cost function. The extension of OBASC image analysis reduces the data rate necessary for coding the test sequence "Ball" from 107 kbit/s to 101 kbit/s.

References

[1] L. Falkenhagen: "Bestimmung der dreidimensionalen Bewegungsparameter flexibler Objekte", diploma thesis, University of Hannover, February 1992.
[2] B.K.P. Horn: "Robot vision", MIT Press, Cambridge, 1986.
[3] R. Koch, "Dynamic 3D scene analysis through synthesis feedback control", IEEE transactions on pattern analysis and machine intelligence, vol. 15, no. 6, June 1993, pp. 556–568.
[4] K.M. Lee, C.–C.J. Kuo: "Shape from shading with perspective projection", CVGIP: Image Understanding, vol. 59, no. 2, March 1994, pp. 202–212.
[5] H.G. Musmann, M. Hötter, J. Ostermann: "Object–oriented analysis–synthesis coding of moving images", Signal Processing: Image Communication, vol. 1, no. 2, Nov. 1989, pp. 117–137.
[6] J.A. Nelder, R. Mead: "A simplex method for function minimization", Computer Journal, no. 7, July 1964, pp. 308–313.
[7] 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.
[8] L.E. Scales, "Introduction to non–linear optimization", Macmillan, London, 1985.
[9] 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.
[10] Stauder, J.: "Estimation of Point Light Source Parameters for Object–Based Coding", accepted for publication in Signal Processing: Image Communication.