Proceedings of the 2nd International Workshop on Biologically Motivated Computer Vision (BMCV 2002), Eds. H. H. Bülthoff, S.-W. Lee, T. Poggio and C. Wallraven,
LNCS 2525, pp. 360–367,

Modeling Insect Compound Eyes: Space-Variant Spherical Vision

Titus R. Neumann

Max Planck Institute for Biological Cybernetics
Spemannstraße 38, 72076 Tübingen, Germany
titus.neumann@tuebingen.mpg.de

Abstract. Insect compound eyes are highly optimized for the visual acquisition of behaviorally relevant information from the environment. Typical sampling and filtering properties include a spherical field of view, a singular viewpoint, low image resolution, overlapping Gaussian-shaped receptive fields, and a space-variant receptor distribution. I present an accurate and efficient compound eye simulation model capable of reconstructing an insect's view of highly realistic virtual environments. The algorithm generates low resolution spherical images from multiple perspective views which can be produced at high frame rates by current computer graphics technology. The sensitivity distribution of each receptor unit is projected on the planar views to compensate for perspective distortions. Applications of this approach can be envisioned for both modeling visual information processing in insects and for the development of novel, biomimetic vision systems.

1 Introduction

The compound eyes of flying insects are the primary source of visual information for a variety of highly efficient orientation strategies and can be regarded as the initial processing stage in the insect visual system. Thus, an accurate and efficient eye model is a prerequisite for comprehensive computer simulations of insect vision and behavior. Furthermore, it facilitates the development of novel, insect-inspired computer vision systems specialized on tasks like visual flight control and navigation.

The fundamental sampling and filtering properties of insect compound eyes [1,2] include discrete receptor units (ommatidia) separated by an interommatidial angle $\Delta\varphi$ (Fig. 1) and distributed over a spherical field of view with a singular viewpoint (valid in the context of flight behavior). The image resolution is low and ranges from approximately 700 ommatidia in the fruitfly *Drosophila* to a maximum of >28000 in the dragonfly *Anax junius*. Spatial wavelengths shorter than $2\Delta\varphi$ cannot be resolved without artifacts, thus high spatial frequencies are suppressed by overlapping Gaussian-shaped receptive fields with acceptance angle $\Delta\rho$ (Fig. 1). Many compound eyes exhibit a space-variant receptor distribution resulting from ecological and behavioral specialization. Insects living in flat environments, such as desert ants or water striders, have an increased resolution around the horizon, whereas some dragonflies have dual 'acute zones' in the frontal and frontodorsal regions for prey detection and tracking.

A number of insect eye simulation models have been developed during the past decade. Early studies include a ray tracing-based simulation of a hoverfly compound

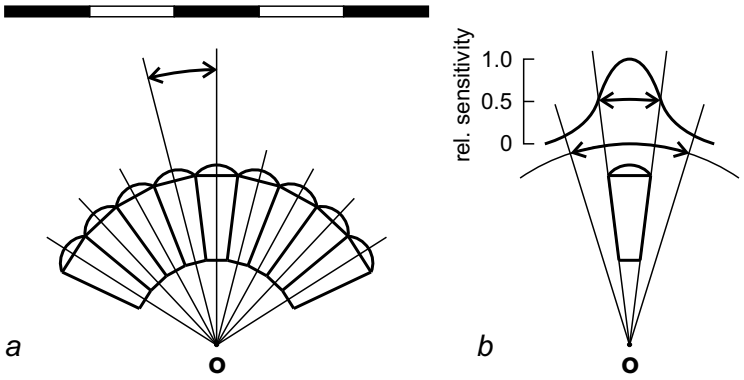

Fig. 1. Visual acuity of compound eyes (Götz, 1965; Land, 1997). *a*. The interommatidial angle $\Delta\varphi$ defines the spatial sampling frequency. *b*. For each receptor unit, the half-width of the Gaussian-shaped sensitivity distribution is defined by the acceptance angle $\Delta\rho$. In the simulation model presented here the receptive field size is restricted by an aperture angle $\Delta\alpha$.

eye looking at a simple black and white stripe pattern [3], and a demonstration of what static planar images might look like for a honeybee [4]. More recently, neural images present in an array of simulated fly visual interneurons have been reconstructed from an image sequence recorded in a simple virtual environment [5], and the visual input of *Drosophila* has been approximated by applying a regular sampling grid to a Mercator projection of the surrounding environment [6]. The common goal of these studies is the reconstruction of particular stimulus properties as seen through a specific insect eye.

Aspects of compound eye optics are also implemented in various insect-inspired computer vision systems. Examples are a one-dimensional, space-variant distribution of optical axes used for altitude control [7], a one-dimensional, panoramic receptor arrangement for view-based navigation [8], and a spherical arrangement of receptor clusters for 3D flight control [9]. Since most of these applications are based on ray tracing, both the eye model and the visual stimulus are strongly simplified to speed up processing.

All of these eye models suffer from one or more of the following restrictions: The field of view is limited and does not comprise the entire sphere [4,5,7,8,9], point sampling or insufficient filtering may lead to spatial aliasing [3,4,7], and regular sampling grids in the Mercator plane result in systematic errors due to polar singularities [6]. Furthermore, ray tracing is not supported by current computer graphics hardware and is therefore inefficient and slow for complex scenes and large numbers of samples.

In this work I present a novel compound eye simulation model that is both more accurate and more efficient than previous approaches.

2 Compound Eye Model

The compound eyes of flying insects map the surrounding environment onto a spherical retinal image. Thus, the involved filtering operations have to be defined on the sphere.

2.1 Space Variant Image Processing on the Sphere

Each viewing direction corresponds to a point on the sphere and can be described by a three-dimensional unit vector

$$\mathbf{d} \in U = \{\mathbf{u} \in \Re^3 \mid \mathbf{u} \cdot \mathbf{u} = 1\}. \tag{1}$$

Extending a notation for space-variant neural mapping [10] to the spherical domain, the set of all possible viewing directions originating from the current eye position is denoted as source area S, the set of all directions covered by the retinal image as target area T, with $S, T \subseteq U$. The current source and target images are defined by the functions $I_S(\mathbf{d}) : S \rightarrow \Re$ and $I_T(\mathbf{d}) : T \rightarrow \Re$, respectively, assigning a gray value to each local viewing direction on the sphere. An extension for color images is straightforward.

The compound eye model determines the retinal image from the surrounding environment using a space-variant linear operator

$$I_S(\mathbf{d}_S) \longmapsto I_T(\mathbf{d}_T) := \int_S I_S(\mathbf{d}_S)\, K(\mathbf{d}_S; \mathbf{d}_T)\; d\mathbf{d}_S \tag{2}$$

integrating over the two-dimensional, spherical source domain S. For each local viewing direction $\mathbf{d}_S \in S$ in the source area and each direction $\mathbf{d}_T \in T$ in the retinal target area the space-variant kernel

$$K(\mathbf{d}_S; \mathbf{d}_T) : S \times T \rightarrow \Re \tag{3}$$

specifies the weight by which the input stimulus influences the retinal image. A kernel closely approximating the spatial low pass filtering properties of compound eye optics is the Gaussian function [1].

2.2 Omnidirectional World Projections

The spherical input images required for the compound eye model described above cannot be generated directly since current raster graphics technology is optimized to produce planar, perspective images. However, the entire surrounding environment can be represented as a cube environment map composed of six square perspective views, one for each face of a cube centered around the viewpoint (Fig. 2 and 3*a*). A color value is determined by intersecting a viewing direction with the corresponding face of the environment map.

To avoid aliasing due to point sampling, Greene and Heckbert [11] proposed the elliptical weighted average filter, a Gaussian-shaped, concentric weight distribution around the intersection point in the image plane. As shown in Fig. 4, this approximation deviates from the correct spherical weight distribution which is defined as a function of the angular distance from the viewing direction (Fig. 3*b*). Since the error increases with the angular width of the filter mask, the elliptical weighted average filter is not suitable for the large ommatidial acceptance angles of compound eyes. In the following I describe the correct spherical filtering transformation.

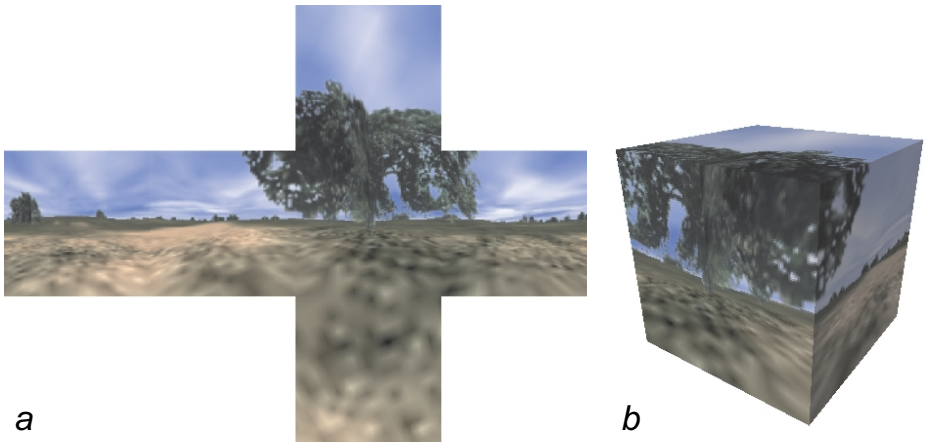

Fig. 2. Cube environment map composed of six square perspective images. *a*. Unfolded. *b*. On a cube surface.

2.3 Filtering Transformation for Discrete Pixels

Both the environment map and the compound eye retinal image are composed of discrete pixels, referred to as source pixels $p \in P_S = \{1, 2, \ldots, |P_S|\}$ and target pixels $q \in P_T = \{1, 2, \ldots, |P_T|\}$, respectively. The source image is represented by the one-dimensional vector $I_S \in \Re^{|P_S|}$ containing the entire environment map. Each source pixel p has a gray value $I_S[p]$. The target image $I_T \in \Re^{|P_T|}$ contains one gray value $I_T[q]$ for each receptor unit q.

The receptive field of receptor unit $q \in P_T$ is described by the weight vector

$$\mathbf{w}_q^{\mathrm{Rcpt}} = \left(w_{q,1}, \ldots, w_{q,p}, \ldots, w_{q,|P_S|}\right)^{\mathrm{T}} \tag{4}$$

indicating the contribution of each input pixel value $I_S[p]$ to the output value $I_T[q]$. The weight matrix $\mathbf{W}^{\mathrm{Rcpt}} \in \Re^{|P_T| \times |P_S|}$ is defined as

$$\mathbf{W}^{\mathrm{Rcpt}} = \left(\mathbf{w}_1^{\mathrm{Rcpt}}, \ldots, \mathbf{w}_q^{\mathrm{Rcpt}}, \ldots, \mathbf{w}_{|P_T|}^{\mathrm{Rcpt}}\right)^{\mathrm{T}} \tag{5}$$

and contains the receptive fields of all $|P_T|$ receptor units. The complete filtering transformation from the input image I_S to the vector I_T of receptor responses is

$$I_T = \mathbf{W}^{\mathrm{Rcpt}} \cdot I_S \,. \tag{6}$$

2.4 Receptive Field Projection

Each receptor unit q has a specific local viewing direction $\mathbf{d}_q \in U$. To prevent spatial aliasing due to point-sampling, the incoming light intensity is integrated over a conical receptive field around $\mathbf{d}_q$. Here, a Gaussian-shaped sensitivity distribution

$$G_q(\zeta) = \exp(-2\zeta^2/\Delta\rho_q^2) \tag{7}$$

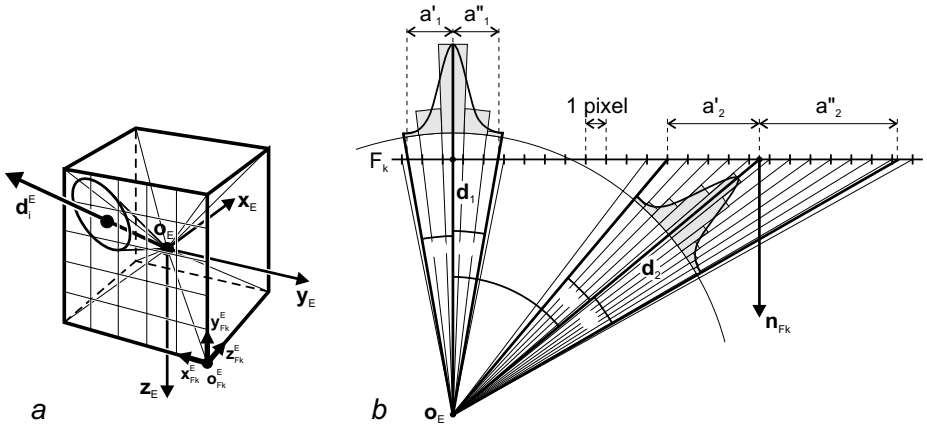

Fig. 3. Receptive field projection. *a*. Eye (E) and face (F_k) coordinate systems of a cube environment map. *b*. Projection and distortion of a Gaussian-shaped sensitivity distribution on a face of the environment map. Although the aperture angle $\Delta\alpha = 2\alpha$ remains constant, the receptive field diameter $\Delta \mathrm{a}_i = \mathrm{a}'_i + \mathrm{a}''_i$ varies for different viewing directions $\mathbf{d}_i$.

with space variant half-width angle $\Delta\rho_q$ (Fig. 1*b*) is used. It is defined as a function of the angular distance ζ from the optical axis. Thus, the relative sensitivity of the receptive field q in the direction $\mathbf{p}_p \in U$ of pixel p is

$$\widehat{w}_{q,p}^{\mathrm{Rcpt}} = \begin{cases} G_q\left(\arccos\left(\mathbf{d}_q \cdot \mathbf{p}_p\right)\right) \cdot A_p, & \arccos\left(\mathbf{d}_q \cdot \mathbf{p}_p\right) \leq \frac{1}{2}\Delta\alpha_q \\ 0, & \text{else} \end{cases} . \tag{8}$$

The solid angle A_p covered by pixel p depends on the position on the image plane (Fig. 3*b*) and is therefore required as a correction factor. The space-variant aperture angle $\Delta\alpha_q$ (Fig. 1*b*) truncates the Gaussian sensitivity distribution to a conical region around the optical axis. Thus, each receptive field (Eq. 4) needs to be normalized to yield a unit gain factor.

2.5 Lookup Table Implementation

Each weight vector $\mathbf{w}_q^{\mathrm{Rcpt}}$ has one entry for each pixel of the input image. This leads to a large but sparsely occupied weight matrix $\mathbf{W}^{\mathrm{Rcpt}}$, since each receptive field covers only a small solid angle containing only a small portion of the environment map. An efficient implementation of the weight matrix is achieved by storing only non-zero weights together with the corresponding pixel indices in a pre-computed lookup table. For each receptive field q the set $\widetilde{P}_q$ of source pixel indices with non-zero weights is determined by

$$\forall q \in P_T : \widetilde{P}_q = \{p \in P_S | w_{q,p}^{\mathrm{Rcpt}} \neq 0\} . \tag{9}$$

With an arbitrary bijective function $\widetilde{o}_q : \{1, \ldots, |\widetilde{P}_q|\} \rightarrow \widetilde{P}_q$ ordering the source pixel indices, and a weighting function $\widetilde{w}_q : \{1, \ldots, |\widetilde{P}_q|\} \rightarrow \Re$ defined as

$$\widetilde{w}_q(p) := w_{q,\widetilde{o}_q(p)}^{\mathrm{Rcpt}} \tag{10}$$

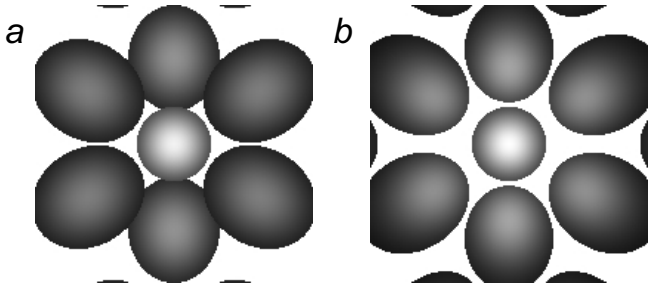

Fig. 4. Distorted receptive fields on a cube map face. *a*. The elliptical weighted average filter (Greene & Heckbert, 1986) shifts the receptive fields towards the image center. *b*. Correct distortion using the receptive field projection method presented here.

the filtering transformation can be efficiently computed by

$$\forall q \in P_T : I_T[q] = \sum_{p=1}^{|\widetilde{P}_q|} \widetilde{w}_q(p) \cdot I_S[\widetilde{o}_q(p)] \,. \tag{11}$$

3 Results

The algorithm was implemented on a standard PC with Intel Pentium II 450 MHz CPU and Nvidia GeForce2 GTS graphics accelerator, and tested for an eye model with 2562 receptor units, an interommatidial angle of $\Delta\varphi = 4.3°$, an acceptance angle of $\Delta\rho = 5.0°$, and an aperture angle of $\Delta\alpha = 10.0°$. On a cube environment map with a face width of 72 pixels and a total number of 31104 pixels, this receptor configuration resulted in distorted receptive fields with a minimum, mean, and maximum size of 30, 58.9, and 126 pixels, respectively. The complete simulation included rendering the example scene (Fig. 2), copying the resulting images to CPU memory, as well as the filtering transformation, and was updated with 9.7 Hz. The isolated filtering transformation achieved 107.9 Hz.

For large aperture and acceptance angles the elliptical weighted average filter [11] leads to considerable deformations of the eye geometry since it shifts the receptive fields towards the image center (Fig. 4*a*). In contrast, the correct projection method presented here preserves the centroid directions of the receptive fields coincident with their optical axes (Fig. 4*b*).

Fig. 5 shows different spherical receptor distributions and the resulting omnidirectional images. The relative local receptor densities are modeled after biological examples [2].

4 Conclusion

The compound eye simulation model presented here is both accurate and efficient. It allows to specify arbitrary, space-variant receptor distributions in a spherical field of view,

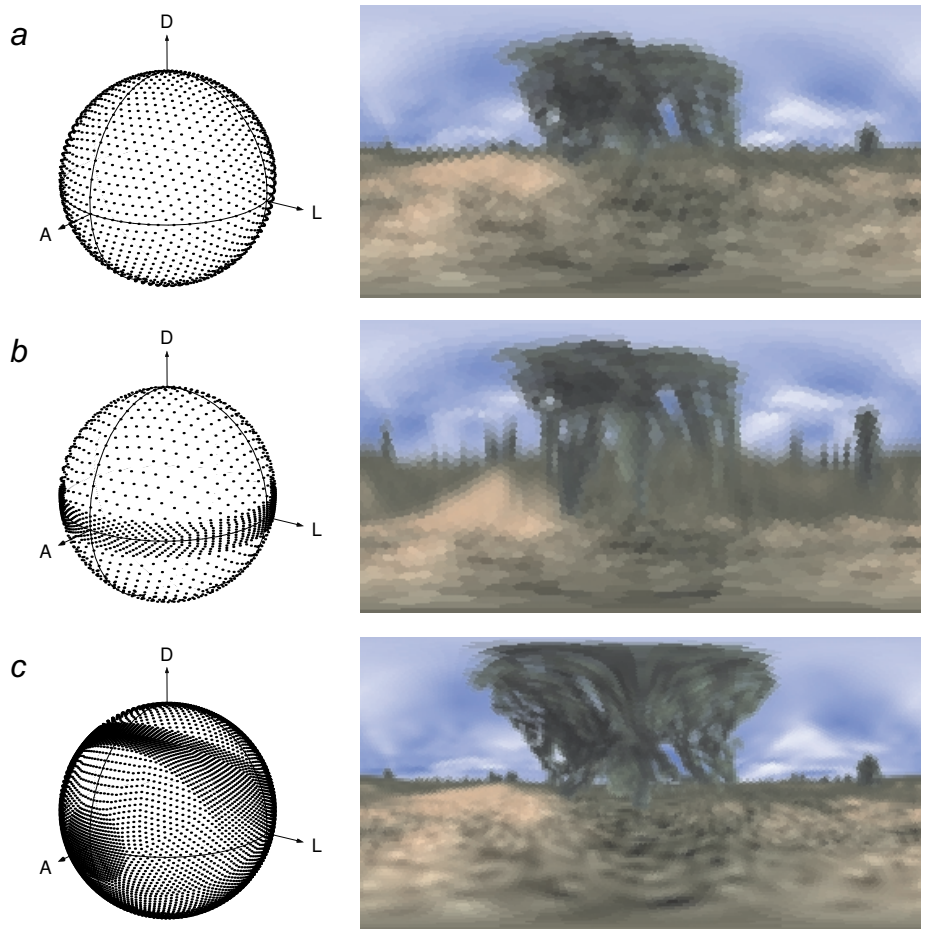

Fig. 5. Space-variant vision with different spherical eye models. Left column: receptor distributions on the sphere (A:anterior, L:lateral, D:dorsal). Right column: example scene (Fig. 2) as seen through each eye model, shown as distorted Mercator projections with full $360^\circ \times 180^\circ$ field of view. *a*. Homogeneous receptor distribution. *b*. Increased resolution along the horizon as in desert ants, water striders or empid flies. *c*. Overall increased resolution and double acute zones in the frontal and the frontodorsal regions as in dragonflies.

and is capable of reconstructing an insect's view in highly realistic virtual environments at high frame rates.

Applications of this approach can be envisioned for both modeling visual information processing in insects and for the development of biomimetic computer vision systems. In biological studies the compound eye simulation can be used to reconstruct

the exact visual stimulus as it is perceived by an insect during an experiment. This allows to find correlations of the visual input with the recorded physiological or behavioral responses. If the resulting processing and control models are implemented as part of the simulation, they can be tested and evaluated in the same open- and closed-loop experiments as the animals, and the observed responses can be compared immediately.

A further application of the proposed eye model is the development of novel, insect-inspired computer vision systems for tasks like visual self-motion control and navigation [12]. These systems are expected to be simpler, more robust and more efficient than existing technical solutions which are easily outperformed by flying insects in spite of their low visual acuity and small brain size.

Acknowledgments. The author thanks Heinrich Bülthoff, Roland Hengstenberg and Karl Götz for support and discussions, and Björn Kreher for technical assistance.

References

1. Götz, K.G.: Die optischen Übertragungseigenschaften der Komplexaugen von *Drosophila*. Kybernetik **2** (1965) 215–221
2. Land, M.F.: Visual acuity in insects. Annual Review of Entomology **42** (1997) 147–177
3. Cliff, D.: The computational hoverfly: A study in computational neuroethology. In Meyer, J.A., Wilson, S.W., eds.: From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior (SAB'90), Cambridge, MA, MIT Press Bradford Books (1991) 87–96
4. Giger, A.D.: B-EYE: The world through the eyes of a bee (http://cvs.anu.edu.au/andy/beye/beyehome.html). Centre for Visual Sciences, Australian National University (1995)
5. van Hateren, J.H.: Simulations of responses in the first neural layers during a flight (http://hlab.phys.rug.nl/demos/fly_eye_sim/index.html). Department of Neurobiophysics, University of Groningen (2001)
6. Tammero, L.F., Dickinson, M.H.: The influence of visual landscape on the free flight behavior of the fruit fly *Drosophila melanogaster*. Journal of Experimental Biology **205** (2002) 327–343
7. Mura, F., Franceschini, N.: Visual control of altitude and speed in a flying agent. In Cliff, D., Husbands, P., Meyer, J.A., Wilson, S.W., eds.: From Animals to Animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior (SAB'94), Cambridge, MA, MIT Press Bradford Books (1994) 91–99
8. Franz, M.O., Schölkopf, B., Mallot, H.A., Bülthoff, H.H.: Learning view graphs for robot navigation. Autonomous Robots **5** (1998) 111–125
9. Neumann, T.R., Bülthoff, H.H.: Insect inspired visual control of translatory flight. In Kelemen, J., Sosik, P., eds.: Advances in Artificial Life, Proceedings of ECAL 2001. Volume 2159 of LNCS/LNAI., Springer-Verlag, Berlin (2001) 627–636
10. Mallot, H.A., von Seelen, W., Giannakopoulos, F.: Neural mapping and space-variant image processing. Neural Networks **3** (1990) 245–263
11. Greene, N., Heckbert, P.S.: Creating raster Omnimax images from multiple perspective views using the elliptical weighted average filter. IEEE Computer Graphics and Applications **6** (1986) 21–27
12. Neumann, T.R., Bülthoff, H.H.: Behavior-oriented vision for biomimetic flight control. In: Proceedings of the International Workshop on Biologically-Inspired Robotics: The Legacy of W. Grey Walter (WGW 2002), 14–17 August, HP Labs Bristol, UK. (2002) 196–203