A Statistical Model for the Distribution of Horizontal Edges in a Planar Scene

John Stowers
Department of Electrical and Computer Engineering
University of Canterbury
New Zealand
Email: john.stowers@ieee.org

Michael Hayes
Department of Electrical and Computer Engineering
University of Canterbury
New Zealand
Email: michael.hayes@canterbury.ac.nz

Peter Smith
Department of Electrical and Computer Engineering
University of Canterbury
New Zealand
Email: peter.smith@canterbury.ac.nz

***Abstract*—This paper presents a statistical model which describes the distribution of horizontal information, "edges", measured by a camera looking over a planar scene. The inspiration for the work is taken from experimental evidence that *Drosophila*, the common fruit fly uses detection of edges to control its flight. Our model is able to predict the distribution of edges for a given camera position, and also to estimate altitude from a given distribution of edges. The model is verified using a number of experiments in a virtual environment.**

***Index Terms*—planes, edges, edge detection, altitude estimation, honeybee, drosophila**

I. Introduction

Experimental work [1], [2], [3], [4] on insects (the common fruit fly *Drosophila* and the honeybee) has demonstrated these organisms recognise lines and edges in their environment (in addition to image motion / optical flow). These features are subsequently used to drive instantaneous flight responses, such as the attraction to long edges [1], behaviour and context specific strategies such as landing [3], [4], and also have recently been shown to provide an absolute global reference for altitude control [2]. The duality of this use of edge information, i.e, the simultaneous use of it for instantaneous response and longer term control, point towards some internal model of ones environment using these features[1].

Our previous work has recreated the edge-based instantaneous altitude control of these organisms on a quadrotor helicopter [6]. This paper addresses the other half of the duality, investigating if it is possible to calculate global information (such as altitude) from the edges in an image if some geometric properties of the camera and environment are known.

To enable estimation of altitude, this paper presents a statistical model that describes the distribution of edges on a plane.

This paper begins with Section II, an introduction to the statistical model describing the distribution of edges. Section III describes the algorithm for detecting edges in an image. Section IV includes the results of measured edges, and those predicted by the model in a virtual environment for a number of camera configurations. The paper concludes with Section V, a discussion of the accuracy of the model, future work and applications of the technique in biomimetic visual flight control.

[1] structural evidence [5] through imaging the *Drosophila* brain also indicate that an internal model may use unconventional parametrisations of the environment as the visual information in these organisms is reduced when it passes through the visual projection neurons before entering the central brain.

II. Image edge statistical model

In this section a statistical model for the images edges detected by a camera looking across a scene is developed (see Figure 1). To make the model tractable, a number of assumptions are made:

1) Edges are assumed to be distributed along the plane as a Poisson process with intensity λ (this is the mean number of edges per metre).
2) The scene is planar or at least has small height variations compared to the camera height h above the scene.
3) Because of limited resolution on the camera, edges close together will not be distinguished.
4) Edges are not missed independently of other edges; they are missed because they are close. Hence the process of detected edges is not Markov and so an exact analysis is difficult.

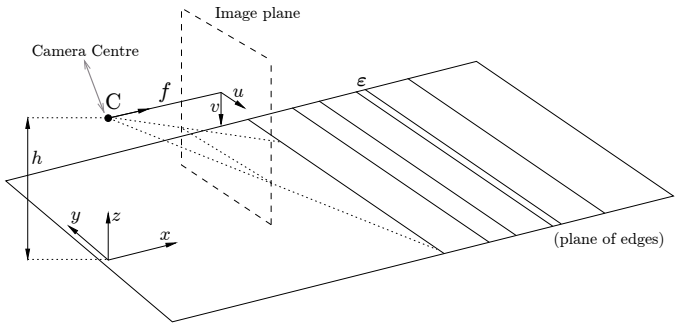


Fig. 1. The geometry of the problem. A camera, C, looks out over a plane containing many edges, ε. The camera has a focal length, f and is at an altitude, h above the plane. The camera image plane co-ordinates are measured in pixels, (u, v).

Using a pin-hole camera model [7], a scene coordinate x is mapped to a image vertical coordinate v (referenced to the

centre of the camera) using

$$v = \frac{fh}{x} = \frac{k}{x}, \tag{1}$$

where f is the focal length of the camera and h is the height of the camera above the scene. These parameters can be combined into a single scale factor k. If the maximum image vertical coordinate is $v_{\max}$, the closest visible scene position is

$$x_{\min} = \frac{k}{v_{\max}}. \tag{2}$$

If the camera has a vertical resolution[2] δ, then multiple edges in the region $x_{\min} \le x < x_1$ are not resolvable, where

$$\begin{aligned} x_1 &= \frac{k}{v_{\max} - \delta}, \\ &= \frac{k}{\frac{k}{x_{\min}} - \delta}. \end{aligned} \tag{3}$$

Assuming a non-homogeneous Poisson process for detected edges where the intensity varies according to the probability an edge is visible, the probability of detecting an edge ε in the region $(x, x+\Delta x)$ where $x_{\min} \le x < x_1$ is the probability of an edge inside $(x, x+\Delta x)$ and no other edge, in the preceding resolution interval. Mathematically, this can be expressed as

$$\begin{aligned} &P\left(\varepsilon_D \in (x, x+\Delta x)\right) \\ &\quad = P\left(\varepsilon \in (x, x+\Delta x)\right) \times P\left(\varepsilon_D \notin (x_{\min}, x)\right), \\ &\quad = (\lambda \Delta x) \times P\left(\varepsilon_D \notin (x_{\min}, x)\right) \\ &\quad = \lambda \Delta x \exp\left(-\lambda (x - x_{\min})\right). \end{aligned} \tag{4}$$

Here ε denotes an actual edge and ε_D denotes a detected edge. Hence, assuming a non-homogeneous Poisson process for detected edges, the intensity at x is

$$\lambda(x) = \lambda \exp\left(-\lambda (x - x_{\min})\right). \tag{5}$$

Now consider a more distant edge at $x > x_1$. The probability of detecting it in the image is

$$\begin{aligned} &P\left(\varepsilon_D \in (x, x+\Delta x)\right) \\ &\quad = P\left(\varepsilon \in (x, x+\Delta x)\right) \times P\left(\varepsilon_D \notin (x_2, x)\right), \\ &\quad = \lambda \Delta x \exp\left(-\lambda (x - x_2)\right), \end{aligned} \tag{6}$$

where

$$\begin{aligned} x_2 &= \frac{k}{v + \delta}, \\ &= \frac{k}{\frac{k}{x} + \delta}. \end{aligned} \tag{7}$$

So the equivalent intensity for a non-homogeneous Poisson process is

$$\lambda(x) = \lambda \exp\left(-\lambda \left(\frac{k}{\frac{k}{x} + \delta}\right)\right). \tag{8}$$

[2]we use the official meaning of the word resolution; the smallest detectable change in the quantity being measured, and not the variation in general use; the number of pixels in an image

If many scenes are simulated and the positions of all detected edges recorded, the empirical cumulative distribution function (CDF) of the edge position converges to

$$\begin{aligned} P(X \le x) &= \lim_{N \to \infty} \left\{ \frac{\#\varepsilon_D < x}{\text{total } \#\varepsilon_D} \right\}, \\ &= \frac{E\left\{\#\epsilon < x\right\}}{E\left\{\text{total } \#\epsilon\right\}}, \\ &= \frac{m(x)}{m(\infty)}, \end{aligned} \tag{9}$$

where

$$m(x) = \int_{x_{\min}}^{x} \lambda(u)\, \mathrm{d}u, \tag{10}$$

and hence

$$\begin{aligned} f_X(x) &= \frac{m'(x)}{m(\infty)} = \frac{\lambda(x)}{m(\infty)}, \\ &= \begin{cases} \lambda \exp\left(-\lambda (x - x_{\min})\right) & x_{\min} \le x \le \frac{k}{\frac{k}{x_{\min}} - \delta} \\ \lambda \exp\left(-\lambda (x - \frac{k}{\frac{k}{x} + \delta})\right) & x_{\min} > \frac{k}{\frac{k}{x_{\min}} - \delta} \end{cases} \end{aligned} \tag{11}$$

From (1) a random scene variable X maps to a random image variable V using

$$V = \frac{k}{X}, \tag{12}$$

and so the probability distribution function for an image edge, $f_V(v)$, is related to the probability distribution function for a scene edge, $f_X(x)$, by

$$f_V(v) = f_X\left(\frac{k}{v}\right) \left| \frac{-k}{v^2} \right|, \tag{13}$$

and thus,

$$\begin{aligned} &f_V(v) \\ &= \begin{cases} \frac{\lambda k}{v^2} \exp\left(-\lambda \left(\frac{k}{v} - x_{\min}\right)\right) & x_{\min} \le \frac{k}{v} \le \frac{k}{\frac{k}{x_{\min}} - \delta} \\ \frac{\lambda k}{v^2} \exp\left(-\lambda \left(\frac{k}{v} - \frac{k}{v+\delta}\right)\right) & \frac{k}{v} > \frac{k}{\frac{k}{x_{\min}} - \delta} \end{cases} \\ &= \begin{cases} \frac{\lambda k}{v^2} \exp\left(-\lambda k \left(\frac{1}{v} - \frac{1}{v_{\max}}\right)\right) & v_{\max} - \delta \le v \le v_{\max} \\ \frac{\lambda k}{v^2} \exp\left(-\lambda k \left(\frac{1}{v} - \frac{1}{v+\delta}\right)\right) & v < v_{\max} - \delta \end{cases} \end{aligned} \tag{14}$$

III. Edge detection

In this section we describe the algorithm used to detect edges in the image. We first calculate the vertical image gradient by convolving the image with the Sobel kernel,

$$\mathbf{G}_v = \mathbf{I} \odot \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix}. \tag{15}$$

Here $\mathbf{I} = \mathbf{I}[m, n]$ is the input image, $\mathbf{G}_v = \mathbf{G}_v[m, n]$ is the vertical gradient image (in the v-axis), and $\odot$ denotes convolution. The Sobel kernels have a smoothing effect, so they are less sensitive to noise than some techniques, and they avoid needing to introduce an additional Gaussian blurring step. The

Sobel operator is also less computationally demanding than the Canny algorithm, which is also commonly used for edge detection,

The next step quantizes the gradient image using a threshold, k_s, based on the RMS value of the gradient image[3],

$$\mathbf{B}_v[m,n] = \begin{cases} 1 & \text{if } \mathbf{G}_v[m,n] \geq k_s \\ 0 & \text{otherwise} \end{cases}. \tag{16}$$

This results in an image with only horizontal images remaining (Figure 2b).

The quantized gradient is then summed in the horizontal direction,

$$\mathbf{E}_v[n] = \sum_{m=1}^{M} \mathbf{B}_v[m,n]. \tag{17}$$

This represents the number of horizontal edges in the image (Figure 2b). Finally, a running sum of $\mathbf{E}_v$ is performed to estimate the empirical cumulative distribution of horizontal scene edges,

$$\mathbf{C}_v[n] = \frac{\sum_{j=1}^{n} \mathbf{E}_v[n]}{\sum_{j=1}^{N} \mathbf{E}_v[n]}. \tag{18}$$

IV. Results and Discussion

The model was verified in a virtual environment (Figure 2a), created using the VRML modelling language. The experiments involved moving the camera in the vertical axis and recording the image, I, at each instant.

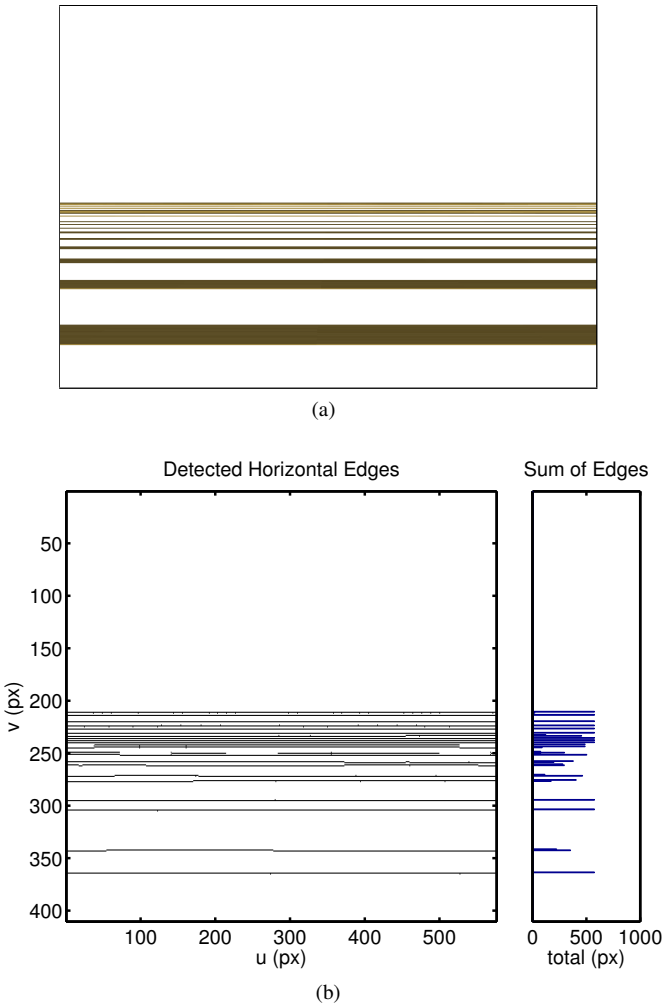
Figure 3 shows the model-predicted CDF, from (14), plotted against the real cumulative sum of edges in the image, from (18). The model fits reasonably well with a larger field of view but gets poorer with a decreasing field of view. Figure 7 shows the affect of the unknown parameter λ on the height estimation when matching the empirical data to the statistical model.

Figure 4 compares the accuracy of the model-predicted height with the actual camera height. The height was estimated by fitting the model CDF to the empirical CDF. There is a slight bias and an increasing variance with camera height and reduced field of view.

An interesting observation is that the peak of the empirical histogram appears to be a linear function of camera height as shown in Figure 5. There is a small bias that needs to be subtracted. It is expected that the proportionality factor is a function of λ; this is yet to be confirmed.

Differentiating the modelled probability distribution function against the image coordinate v predicts a square-root dependence of the position of the mode versus camera height (see Figure 6). Thus while the statistical model is a simple continuous first approximation, it only captures some of the behaviour of the system. An improvement to the model would incorporate the effects of image sampling.

[3]this is the recommended approach for thresholding Sobel generated gradient images



(a)

(b)

Fig. 2. The virtual environment used for validation of the probabilistic model of edges. (a) An image from the virtual scene (note, this has be recoloured, replacing the black background with white for readability). (b) The result of the edge detection algorithm (left), $\mathbf{B}_v$, and the row-wise (v) sum of edges (right), $\mathbf{E}_v$.

V. Conclusion

Our results show that the derived model has skill; in Figure 3 the model is able to approximate the distribution of edge information of a planar scene, and in Figure 4 the model is able to estimate the height given a distribution of edges.

These two results demonstrate that it is possible to extract sufficient information from edges in an image in order to estimate the camera height.

In the biological context from which this work was inspired, it is interesting to consider what accuracy the algorithm needs to be in order to have utility to an organism. An analogous situation is the avoidance response in *Drosophila*[4], despite the insect continually measuring image motion, the avoidance response seems to be active only when optical flow exceeds a predetermined threshold. Comparatively, the model-

[4]the avoidance response is where the insect turns away from regions in its environment when the image motion / optical flow exceeds a certain value of dilation.

predicted height looks suitable for such decisions, qualitatively Figure 3 shows the model can distinguish large from small heights. Quantitatively, Figure 4 shows the model can provide a reasonable height estimate at low altitudes.

Further work is required to improve the model to predict the observed linear height dependence of the empirical histogram mode. Additional work remains to assess the model performance when the scene is non-planar, in in non-ideal lighting conditions, or when λ is unknown. Even in these scenarios the results of Figure 4 are encouraging. If λ can be assumed constant (such as over a small number of frames) the agreement between the model-predicted and actual height at low values is again reminiscent of time-to-collision approaches [8] which emulate the biological expansion avoidance response.

REFERENCES

[1] G. Maimon, A. D. Straw, and M. H. Dickinson, "A Simple Vision-Based Algorithm for Decision Making in Flying Drosophila," *Current Biology*, 2008.

[2] A. D. Straw, S. Lee, and M. H. Dickinson, "Visual Control of Altitude in Flying Drosophila," *Current Biology*, vol. 20, no. 17, pp. 1550–1556, 2010.

[3] M. Lehrer, M. V. Srinivasan, and S. W. Zhang, "Visual Edge Detection in the Honeybee and its Chromatic Properties," *Proceedings of the Royal Society of London. B. Biological Sciences*, vol. 238, no. 1293, pp. 321–330, 1990.

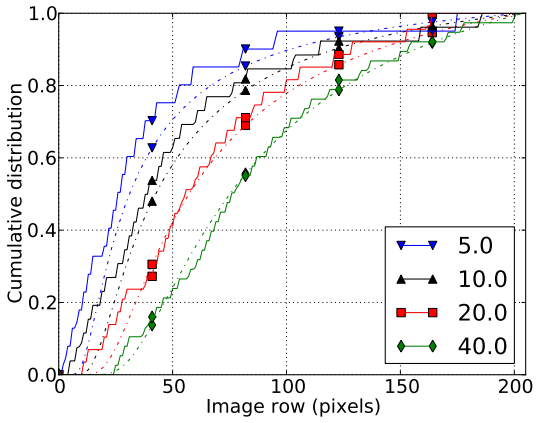
[4] M. Lehrer and M. V. Srinivasan, "Object detection by honeybees: Why do they land on edges?" *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 173, no. 1, pp. 23–32, 1993, 10.1007/BF00209615. [Online]. Available: {http://dx.doi.org/10.1007/BF00209615}

[5] H. Otsuna and K. Ito, "Systematic analysis of the visual projection neurons of Drosophila melanogaster. I. Lobula-specific pathways," *The Journal of Comparative Neurology*, vol. 497, no. 6, pp. 928–958, 2006.

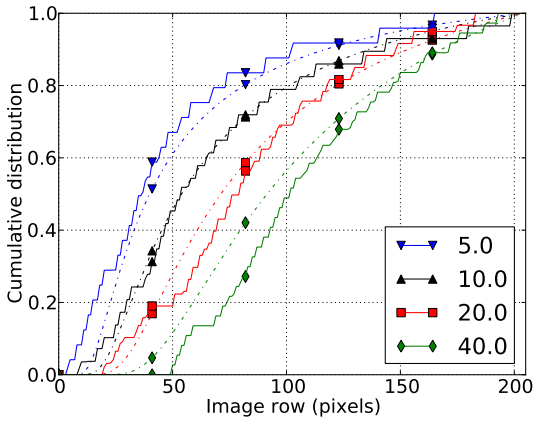
[6] J. Stowers, M. Hayes, and A. Bainbridge-Smith, "Beyond Optical Flow - Biomimetic UAV Altitude Control using Horizontal Edge Information," in *The 5th International Conference on Automation, Robotics and Applications (ICARA 2011)*, Wellington, New Zealand, dec 2011.

[7] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Second ed. Cambridge University Press, ISBN: 0521540518, 2004.

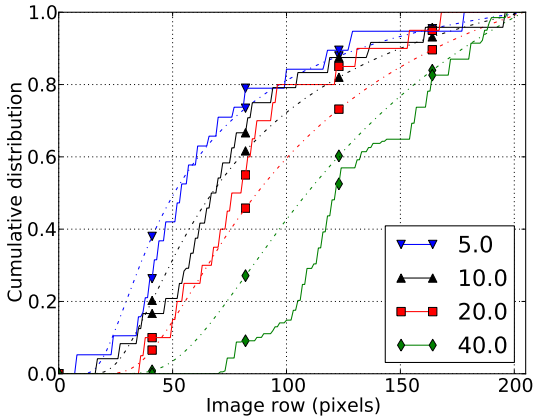
[8] B. K. P. Horn, Y. Fang, and I. Masaki, "Time to Contact Relative to a Planar Surface," in *Intelligent Vehicles Symposium, 2007 IEEE*, june 2007, pp. 68–74.



(a)

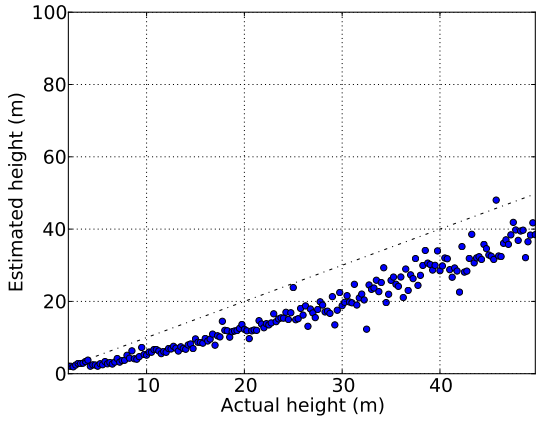


(b)

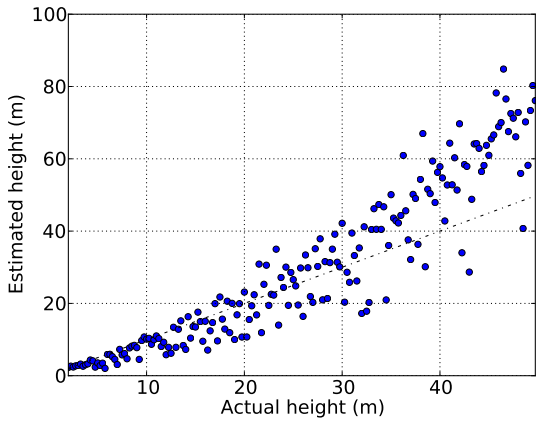


(c)

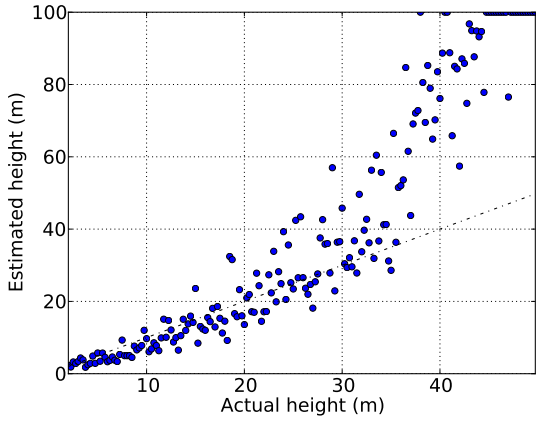
Fig. 3. Comparison of estimated cumulative probability distribution with the model (dashed) as a function of height h for $\lambda = 0.35$: (a) camera field of view 120 degrees, (b) camera field of view 90 degrees, (c) camera field of view 60 degrees.



(a)

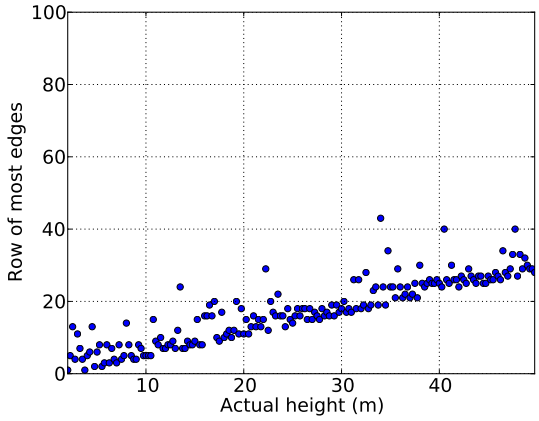


(b)

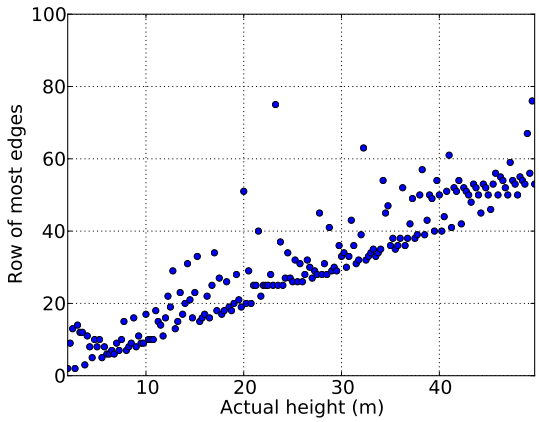


(c)

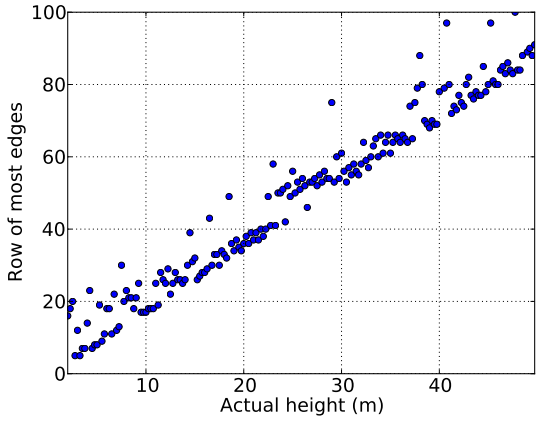
Fig. 4. Comparison of estimated camera height versus actual camera height with $\lambda = 0.35$. (a) camera field of view 120 degrees, (b) camera field of view 90 degrees, (c) camera field of view 60 degrees.



(a)

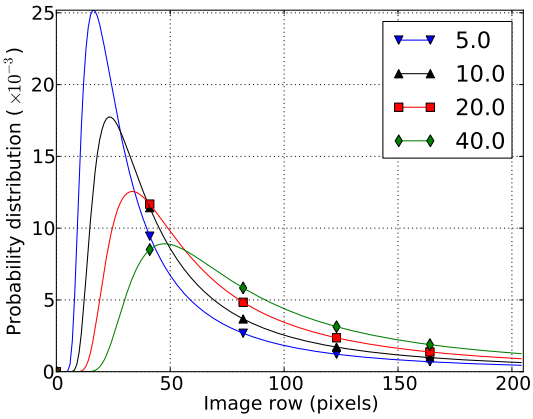


(b)

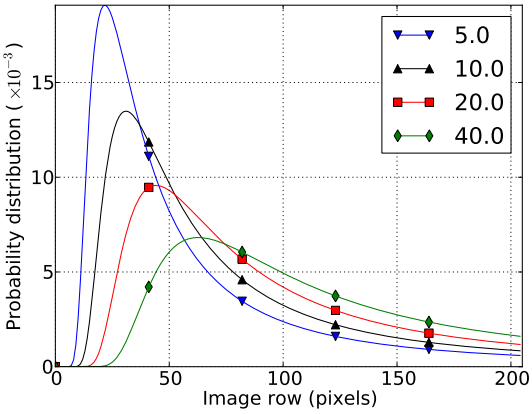


(c)

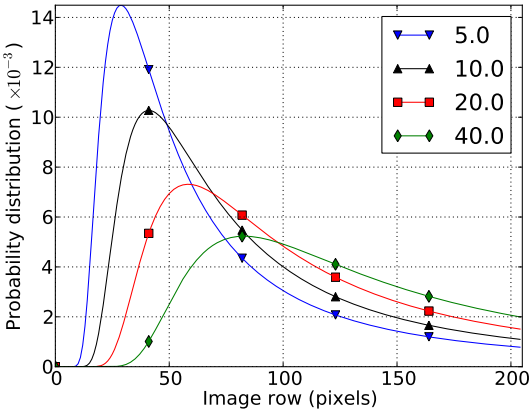
Fig. 5. Comparison of the image row with the most edges versus the camera height with $\lambda = 0.35$. (a) camera field of view 120 degrees, (b) camera field of view 90 degrees, (c) camera field of view 60 degrees.



(a)



(b)



(c)

Fig. 6. Modelled probability distribution functions as a function of camera height h for $\lambda = 0.35$: (a) camera field of view 120 degrees, (b) camera field of view 90 degrees, (c) camera field of view 60 degrees.

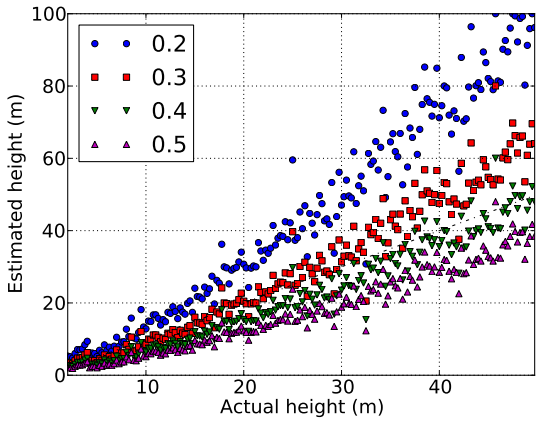


Fig. 7. Comparison of estimated camera height versus actual camera height as a function of λ with a camera field of view 120 degrees.