Binocular Projection of a Random Scene

Miles Hansard
www.eecs.qmul.ac.uk/~milesh/

School of Elec. Eng. & Comp. Sci.
Queen Mary, University of London,
Mile End Road, London E1 4NS.

Current approaches to large-scale visual reconstruction would benefit from a statistical model of multiple-view projection. In particular, global constraints based on scene-clutter and occlusion are required. This work presents a new statistical model, starting from the simplest example of a random scene, viewed by two cameras. The most interesting aspects are revealed by working in a single epipolar plane, and supposing that the scene consists of identical discs of radius ε, as in fig. 1.
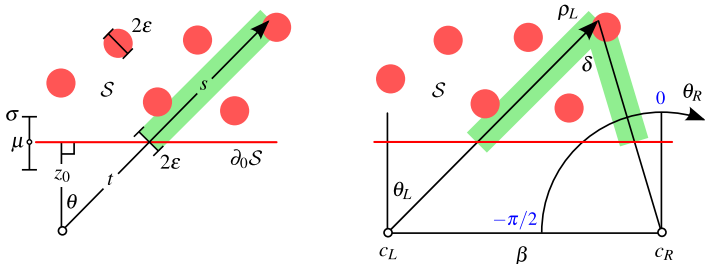

Figure 1: **Viewing geometry**. *Monocular case, left:* The scene $\mathcal{S}$ comprises discs of radius ε, with near-boundary $\partial_0 \mathcal{S}$. The depth z_0 of the latter is normally distributed around μ. A ray in direction θ extends to distance t through empty space, followed by distance s through the scene, before striking an object. Equivalently, there are no disc-centres in the $2\varepsilon \times s$ green rectangle, and so s is exponentially distributed. *Binocular case, right:* both rectangles must be empty for the disc to be binocularly visible. Note that equal increments of θ_R would span increasingly long segments of $\rho_L = s + t$ as the difference-angle δ decreases.

It has previously been shown [3, 4] that, if the discs are distributed according to a Poisson process [2] of intensity λ, then the distance to a visible object follows an exponential distribution F of intensity $2\varepsilon\lambda$;

$$\mathrm{pr}(s \,|\, \lambda, \varepsilon) = F(s, 2\varepsilon\lambda). \tag{1}$$

This is because if an object at distance s is visible, then the $s \times 2\varepsilon$ rectangle around the corresponding ray must be empty (cf. fig. 1, left). But the exponential model is not qualitatively realistic, because the mode of the distribution is at zero, implying that the optical centre is fully amid the clutter. Real range-data, in contrast, follows a two-tailed distribution along each ray, as seen in fig. 2.
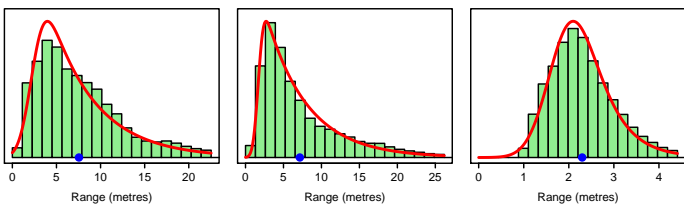

Figure 2: **Range densities**. *Left:* maximum likelihood ex-Gaussian fit to the upper 'canopy' regions of 14 range-scans of a forest scene. *Middle:* fit to the central $120° \times 40°$ 'trunks' regions, which is most representative of heterogeneous clutter. *Right:* fit to the lower 'ground' regions, which is more Gaussian. The blue dot is the mean of each data-set.

A more realistic visibility density is proposed here, in which the optical centre is displaced from the scene $\mathcal{S}$ by a shift t. If the near-boundary $\partial_0 \mathcal{S}$ of the scene is locally perpendicular to the straight-ahead direction, at a normally distributed distance $G(z_0, \mu, \sigma)$, then a ray at angle θ passes a distance t through empty space, where

$$\mathrm{pr}(t \,|\, \theta, \mu, \sigma) = G\left(t, \frac{\mu}{\cos\theta}, \frac{\sigma}{\cos\theta}\right). \tag{2}$$

The total distance ρ to a visible object is then $\rho = s + t$ where s has an exponential density (1) and t has a Gaussian density (2). It follows that

the density H of ρ is obtained by convolution, $H = F \star G$. This can be expressed as a standard *ex-Gaussian* distribution [1] where

$$\mathrm{pr}(\rho \,|\, \theta) = H\left(\rho, \, 2\varepsilon\lambda, \, \frac{\mu}{\cos\theta}, \, \frac{\sigma}{\cos\theta}\right). \tag{3}$$

Figure 2 shows that this model is a good fit to real range data, acquired with an outdoor laser scanner.

It is straightforward to extend (3) to the binocular case, in which an object is jointly visible if both left and right rays are unobstructed, hence

$$\mathrm{pr}(\rho_L, \rho_R) = \mathrm{pr}(\rho_L \,|\, \theta_L) \times \mathrm{pr}(\rho_R \,|\, \theta_R). \tag{4}$$

This model is expressed in terms of scene-distances ρ_L and ρ_R, which cannot be directly observed in practice. However, given that the two rays must intersect, ρ_L and ρ_R are functions of the left and right visual directions θ_L and θ_R from optical centres from $\boldsymbol{c}_L$ and $\boldsymbol{c}_R$ respectively (cf. fig. 1, right). Hence it is possible to reparameterize (4), in order to obtain the *conditional probability of observing a point in direction θ_R from $\boldsymbol{c}_R$, given that it is observed in direction θ_L from $\boldsymbol{c}_L$*. This density, which is supported on an epipolar line, involves the Jacobian

$$J_R(\theta) = \rho_R \sqrt{\csc^2 \delta + \cot^2 \delta}, \quad \text{where} \quad \delta = \theta - \theta_L \tag{5}$$

such that $J_R(\theta_R)$ accounts for the variation of the combined distances ρ_L and ρ_R, with respect to θ_R. The complete conditional density is then

$$\mathrm{pr}(\theta_R \,|\, \theta_L) = \mathrm{pr}(\rho_L, \rho_R) \times J_R(\theta_R) / S_R(\theta_L) \tag{6}$$

where $S_R(\theta_L)$ is the normalizing constant, which can be obtained (if required) by numerical integration. The density (6) is interesting, because it balances two opposite tendencies. On one hand, the tails of the ex-Gaussian parts (3) are exponentially decreasing, which means that distant objects are *less* likely to be seen. On the other hand, the Jacobian term (5) expresses the fact that, as the two visual directions become parallel, small changes of θ_R cause big changes in ρ_L and ρ_R, making it *more* likely that an object will be seen as θ_R approaches the vanishing point. Some examples and simulations of $\mathrm{pr}(\theta_R \,|\, \theta_L)$ are shown in fig. 3.
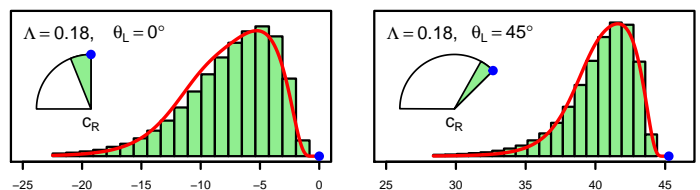

Figure 3: **Correspondence densities**. Histograms show the angular distributions $\mathrm{pr}(\theta_R | \theta_L)$ obtained by Monte Carlo simulation, as functions of θ_R, for two different directions θ_L. Red lines are *predicted* (not fitted) distributions defined by (6). The angular sectors extend from the epipole ($-90°$, cf. fig. 1, right) to θ_L, which is the vanishing-point (blue dot) of the ray through $\boldsymbol{c}_L$, as seen from $\boldsymbol{c}_R$.

In summary, the conditional probability of observing a point in a Poisson scene, given that it has already been observed in another view, has been derived. In particular, it has been shown how this probability, as a function of θ_R, is determined by the given direction θ_L and the clutter intensity λ. This provides a theoretical basis for new Bayesian priors in binocular image-matching.

[1] S.L. Burbeck and R.D. Luce. Evidence from auditory simple reaction-times for both change and level detectors. *Perception and Psychophysics*, 32:117–133, 1982.

[2] D.R. Cox and V. Isham. *Point Processes*. Chapman & Hall, 1980.

[3] J. Huang, A.B. Lee, and D. Mumford. Statistics of range images. In *Proc. CVPR*, pages 324–331, 2000.

[4] M.S. Langer. Surface visibility probabilities in 3D cluttered scenes. In *Proc. ECCV*, pages 401–412, 2008.