Efficient and unbiased crowd size estimation

Marcos Cruz, Javier González-Villa, Luis Manuel Cruz-Orive

Abstract:

Crowd size estimation is a long standing problem. Exhaustive manual counting is tedious, slow, difficult to verify and unfeasible for large populations. Automatic detection and counting algorithms are generally biased and usually have high and unpredictable variance. An alternative is to multiply population density with some reference area but, unfortunately, sampling details, handling of edge effects, etc., are seldom described, whereby the statistical properties of the estimators are generally unknown. We address the problem by using principles of geometric sampling. These principles are old and solid, but largely unknown outside the areas of microscopy and stereology. The proposed design is design unbiased irrespective of population size, pattern, perspective artifacts, etc. The implementation is very simple—it is based on the random superimposition of coarse quadrat grids. The essential ingredient is proper, well defined sampling. We propose (and check via Monte Carlo resampling) a new theoretical variance prediction formula. As far as efficiency, for the different cases studied counting about 50 (100) individuals in 20 quadrats, yielded relative standard errors of about 8% (5%) with a few minutes work. This fact may effectively break the barrier hitherto imposed by the current lack of reliable automatic detection algorithms, because semiautomatic sampling and manual counting may be an attractive option.