Design unbiased population size estimation on Gigapixel images

19th Workshop on Stochastic Geometry, Stereology and Image Analysis

Marcos Cruz

Departamento de Matemáticas, Estadística y Computación

Universidad de Cantabria

Barcelona demonstration 2010

LAVANGUARDIA.es

Un millón de personas inundan Barcelona en una histórica manifestación de rechazo a la sentencia contra el Estatut

ABC

Noticias agencias

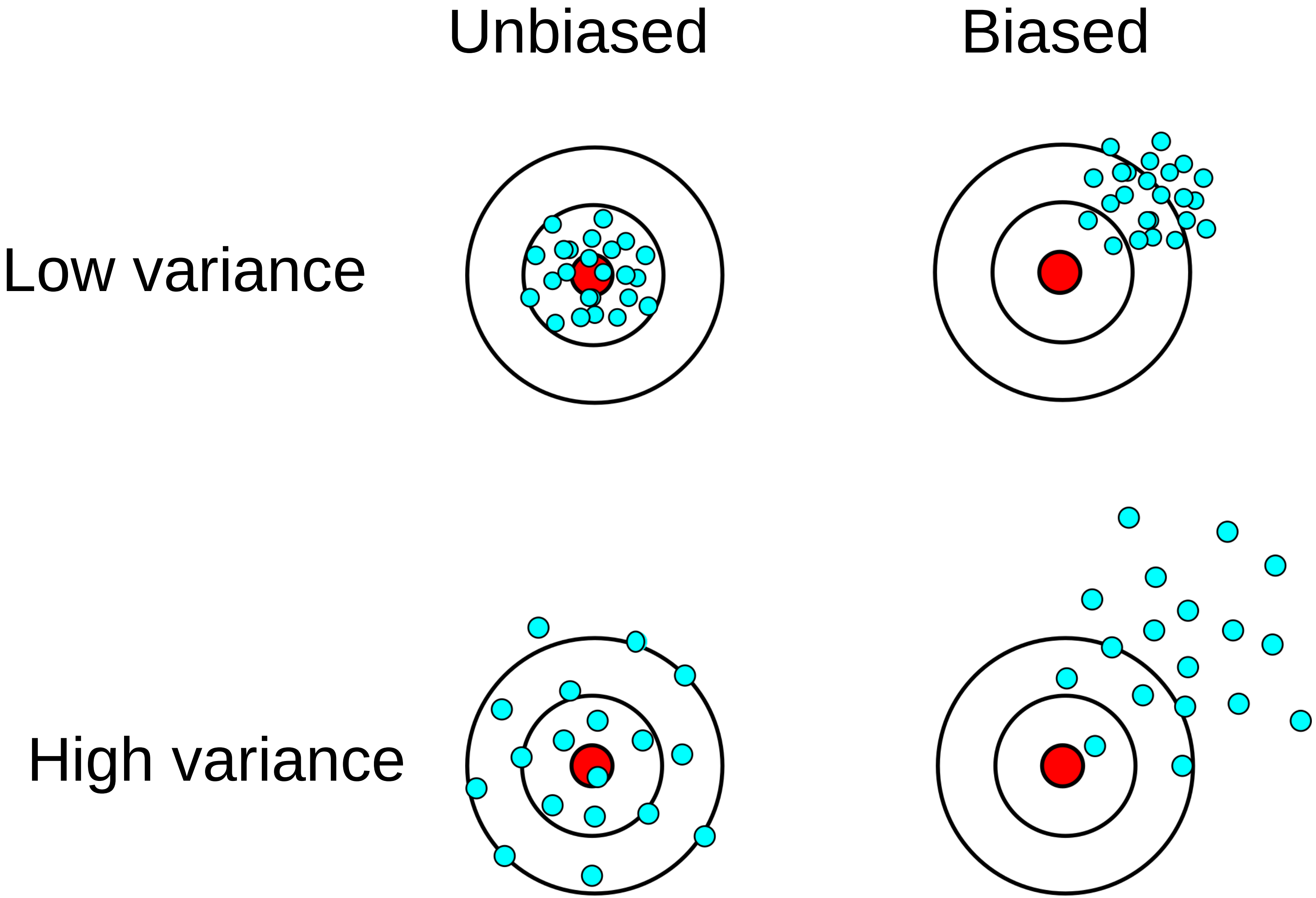
Unas 56.000 personas se manifiestan en defensa del Estatut,según estima Lynce

Why are there huge discrepancies?

- ~~Political Bias~~
- Mathematical Bias
- Results cannot be verified

We need an unbiased and verifiable method

Bias and variance



Number estimation on still images

We assume that all the particles in the image are unambiguously distinguishable for counting

Crowd Images

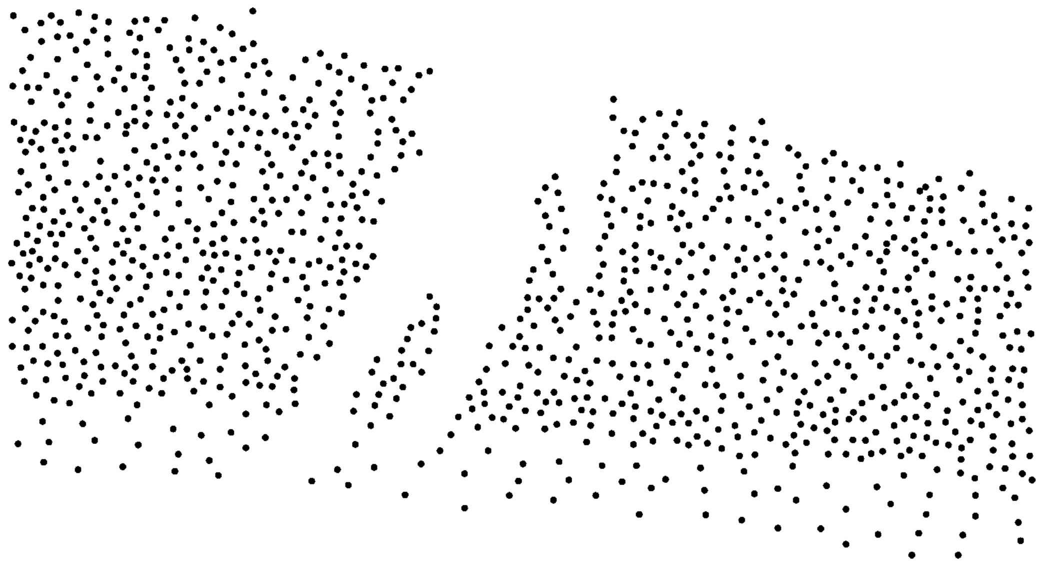
Photo: R. Cancio 1966

Photo: M. Ozer (AFP/Getty Images) 2010

Annotated positions

Ground Truth: N = 1120

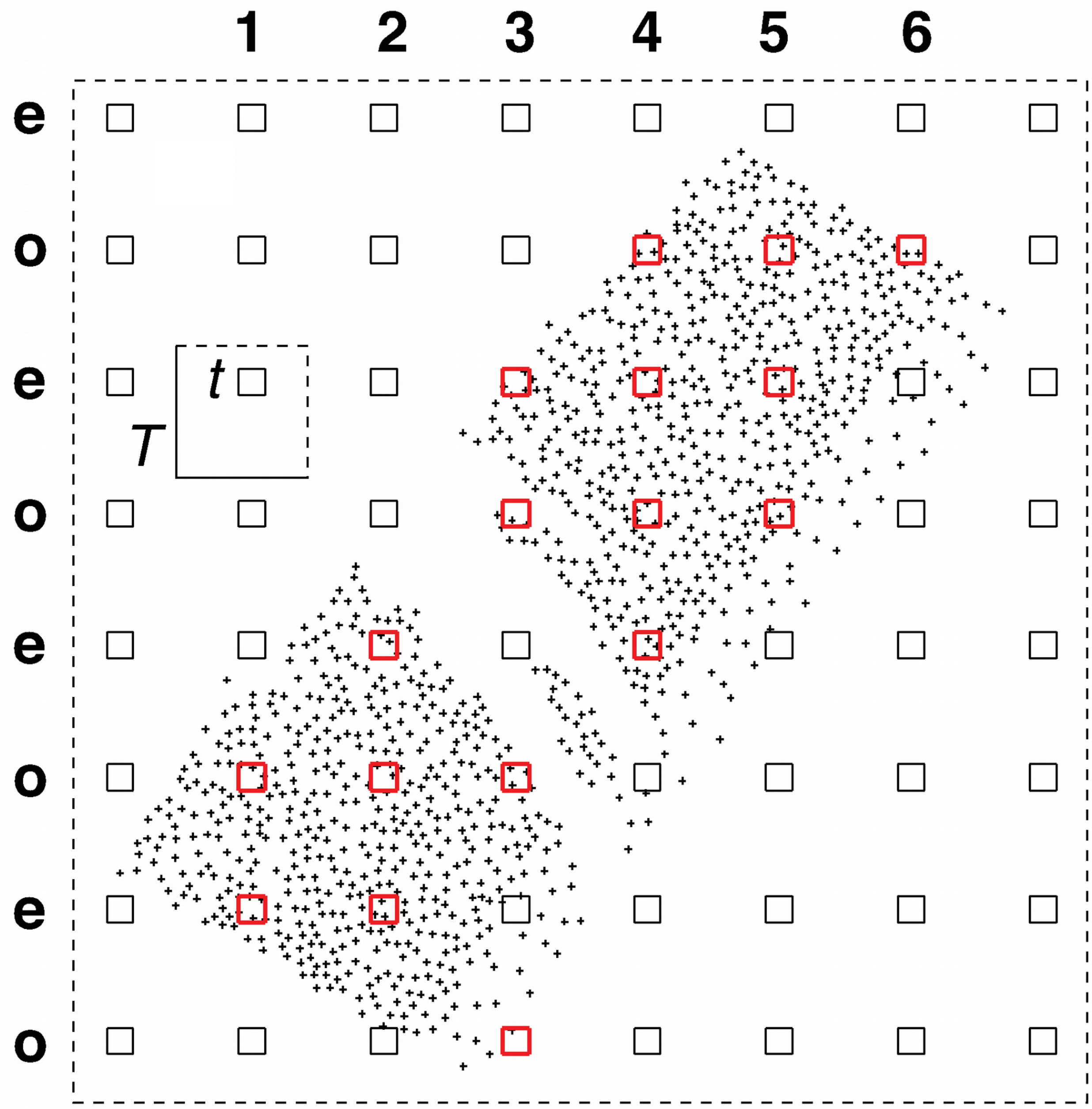
Cruz et al. 2015



Ground Truth: N = 4633

Idrees et al. 2013

Systematic Quadrat sampling



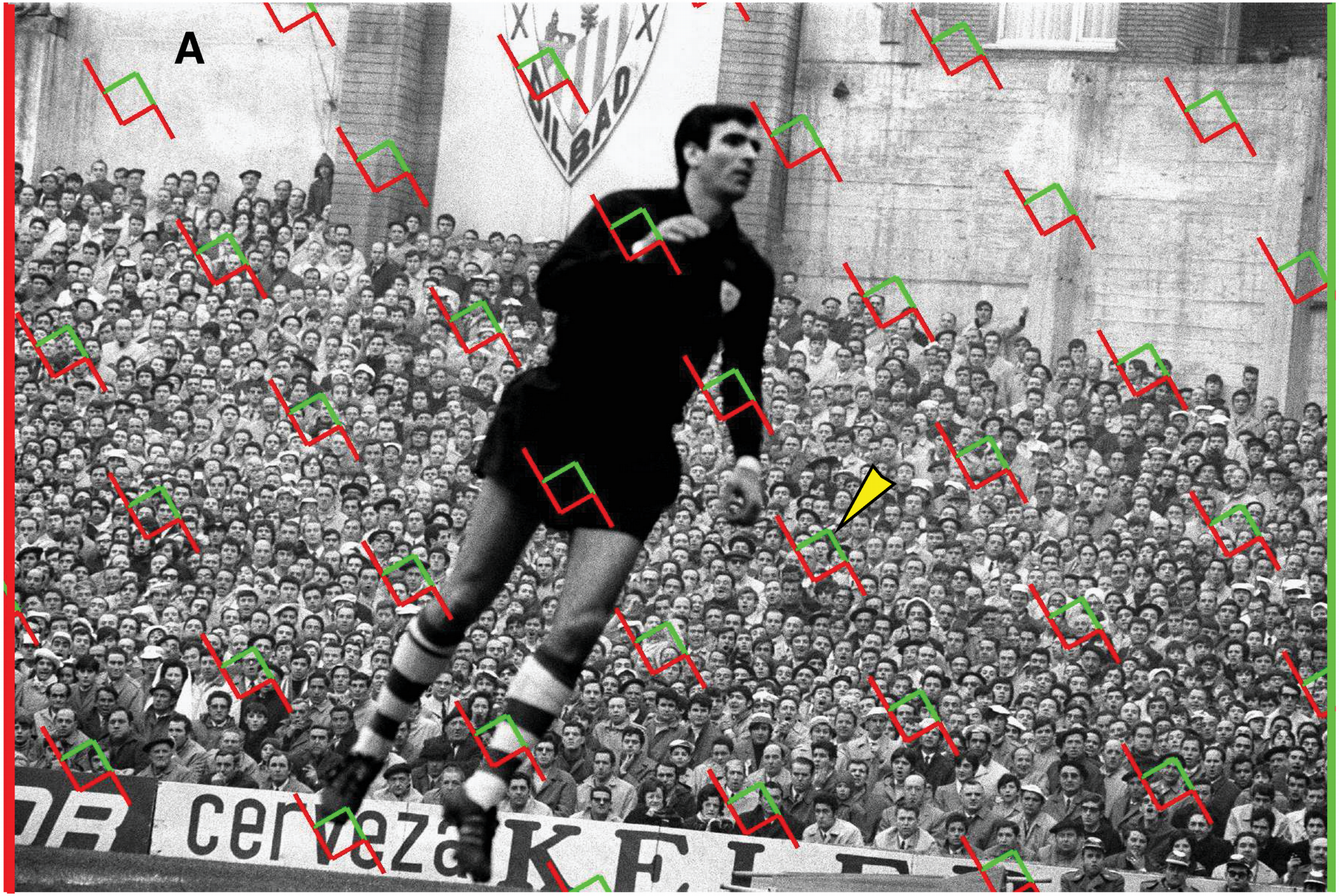
- *T*= Box size
- *t* = Quadrat size
- *Q* = sample size
- Arbitrary fixed rotation
- Uniform random grid

Design unbiased estimator:

$$\hat{N} = \frac{T^2}{t^2} Q$$

Cruz et al. 2015

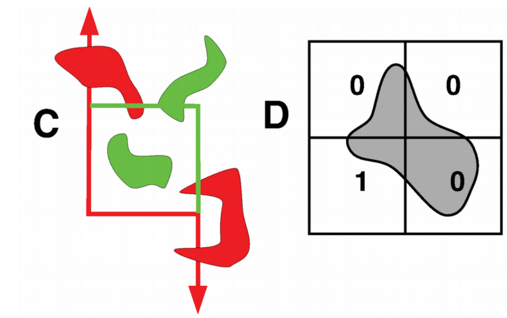
Systematic Quadrat sampling



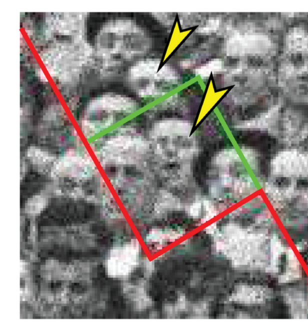
Procedure

- Define counting unit: **head**, body, eye…
- Define unbiased counting rule

Unbiased counting frame



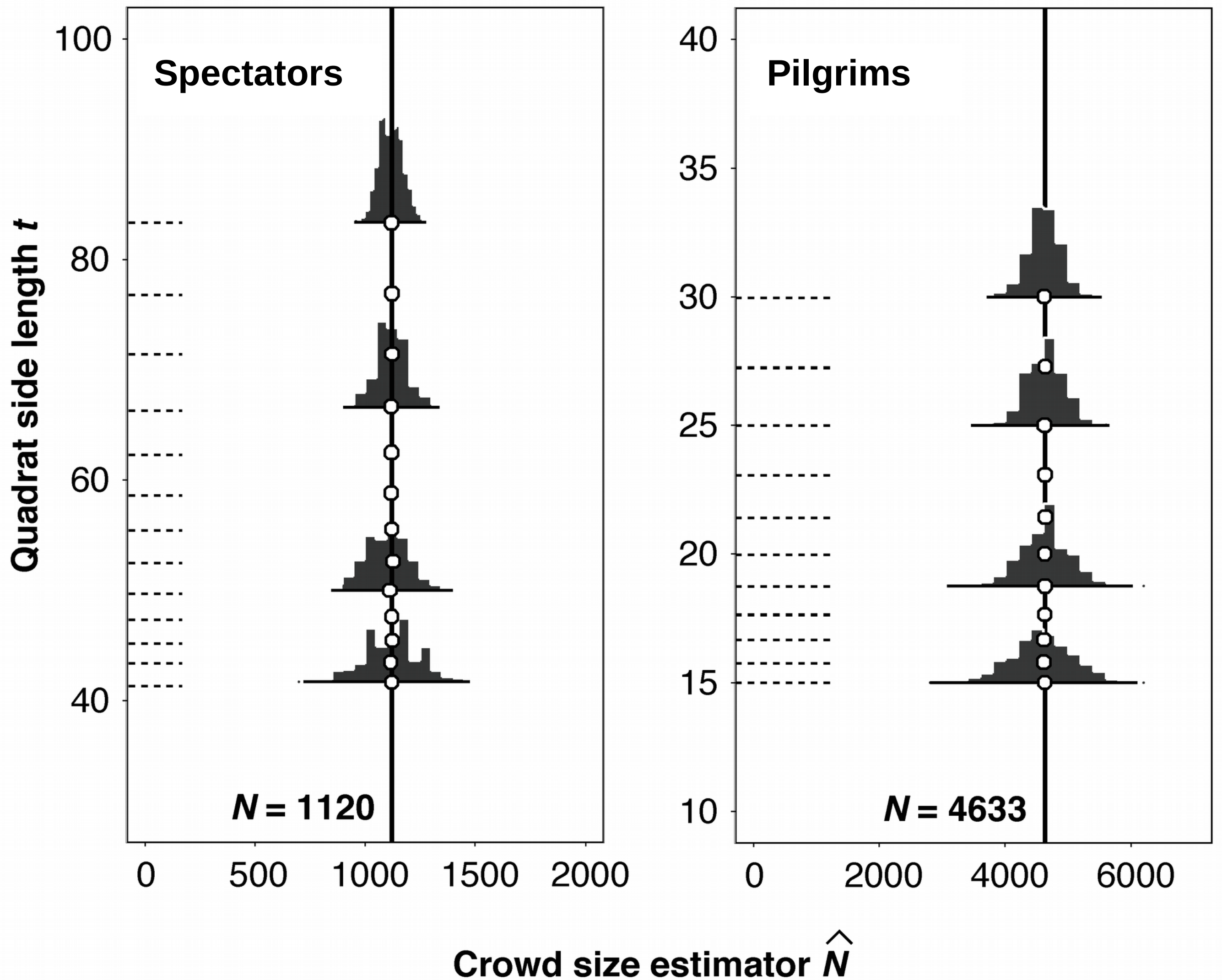
E



H.J.G. Gundersen 1977

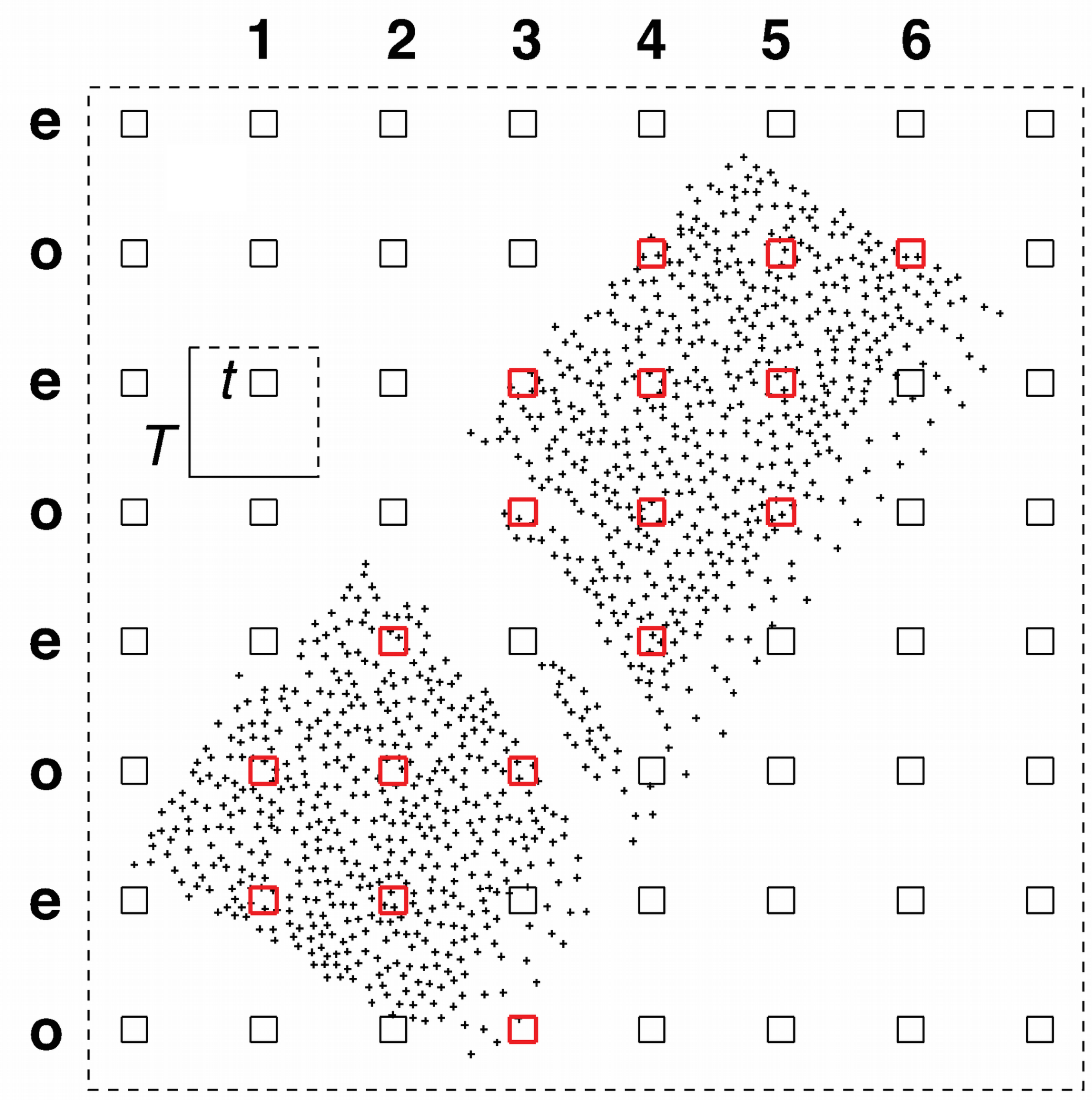
Advantages of the proposed method

Design unbiased and no false detections



Cruz et al. 2015

Variance can be estimated



- Naive approach: assume independent quadrat counts
  - Easy to calculate
  - Unrealistic, bad estimation
- Our approach: non-independent quadrat counts
  - Covariogram model
  - Better performance

Cruz et al. 2015

Variance estimator

- *Cavalieri* variance estimator:
  - First term estimates between stripes variance contribution
  - Second term estimates within stripes variance contribution

$$\mathrm{var}_{\mathrm{Cav}}\left(\widehat{N}\right) = \frac{1}{6}\cdot\frac{(1-\tau)^2}{\tau^4(2-\tau)}\cdot[3(C_0-v_n)-4C_1+C_2]+\frac{v_n}{\tau^4},$$

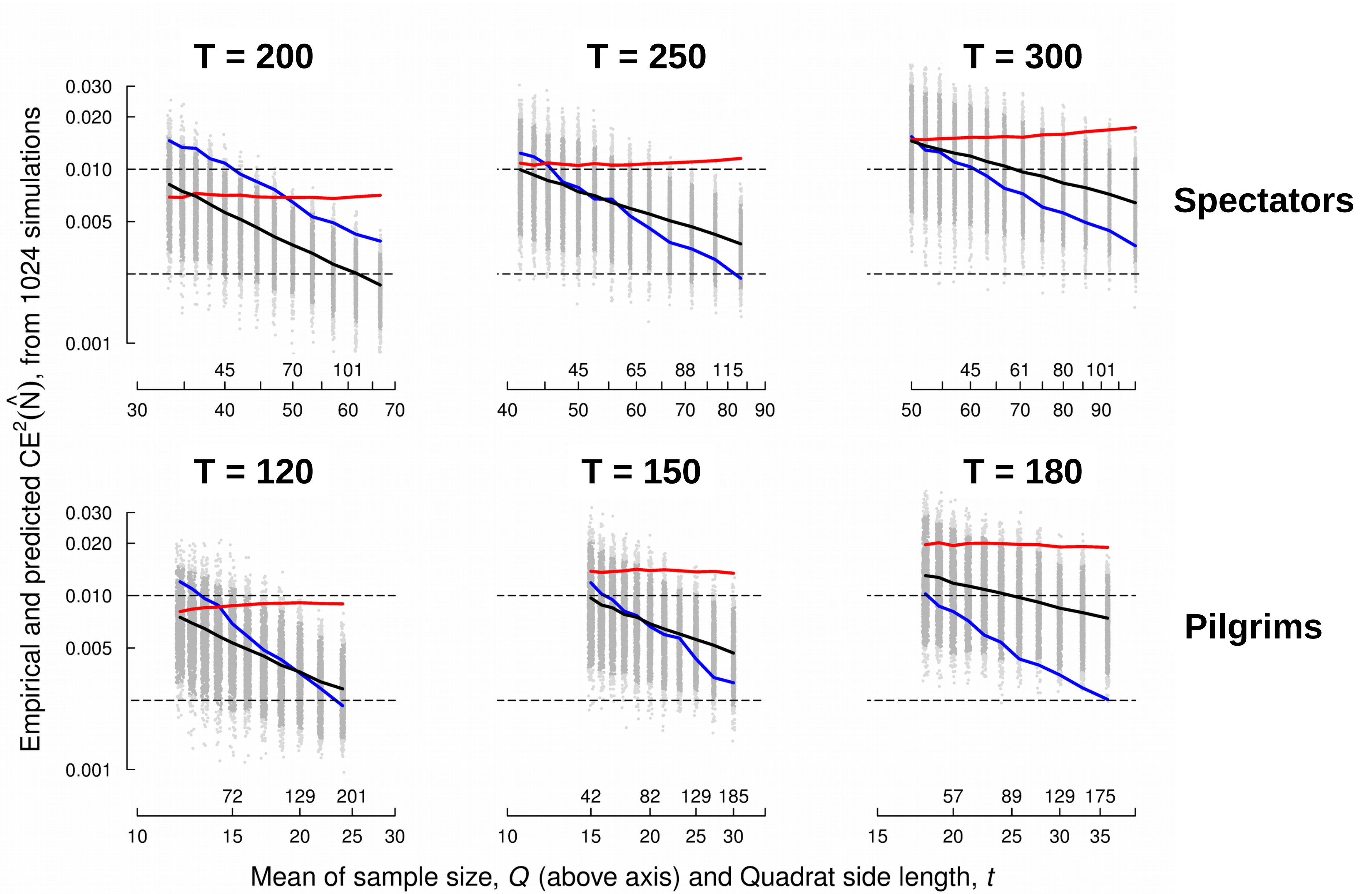
$$C_k = \sum_{j=1}^{n-k} Q_j Q_{j+k}, \quad k=0,1,2.$$

$$v_n = \frac{(1-\tau)^2}{3-2\tau}\cdot\sum_{i=1}^{n}(Q_{oi}-Q_{ei})^2.$$

$$\tau = t/T$$

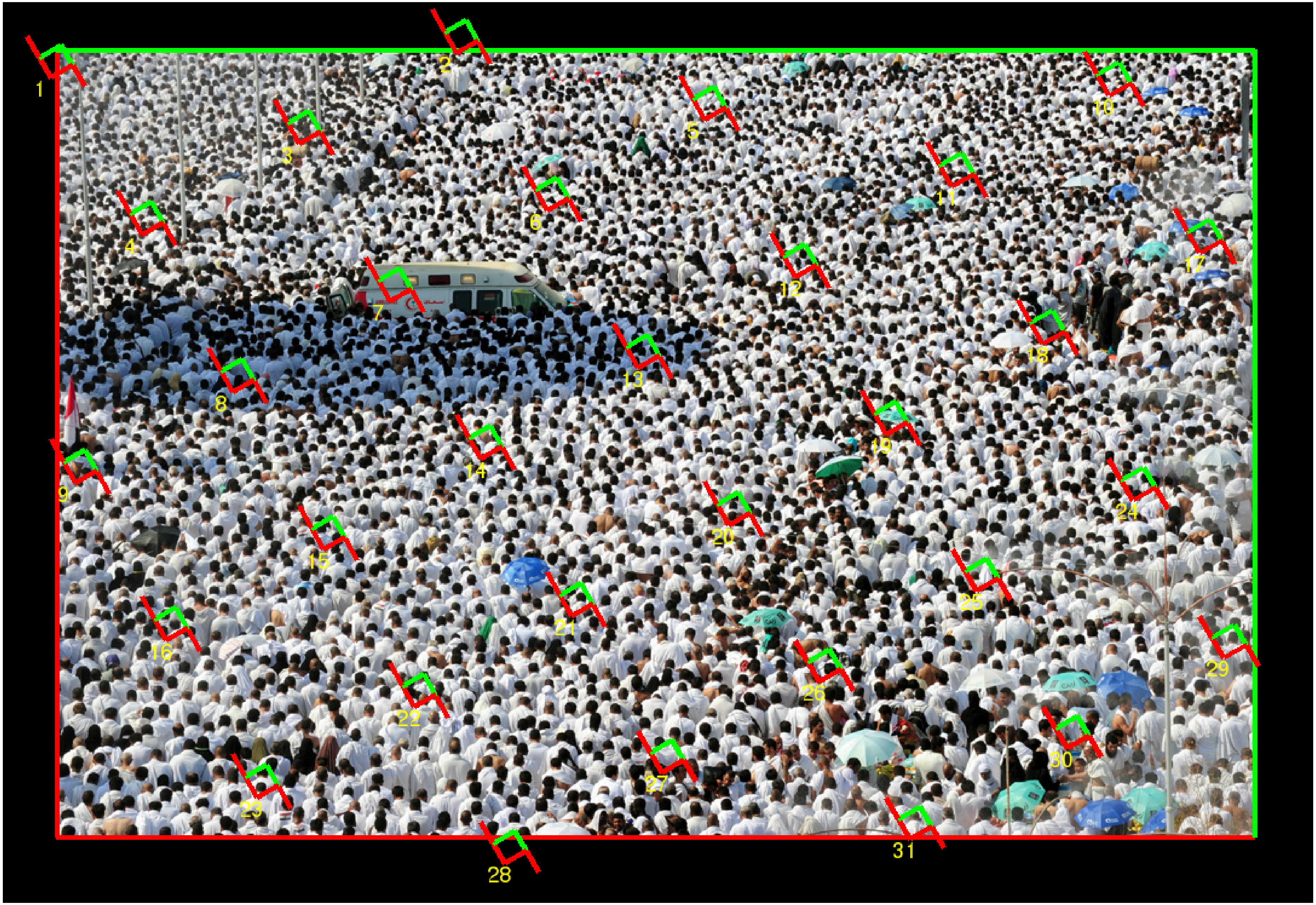
Cruz et al. 2015

Predictable and low Variance



Cruz et al. 2015

Independent of population size



Reference area not needed

Avoids ambiguous reference area definition

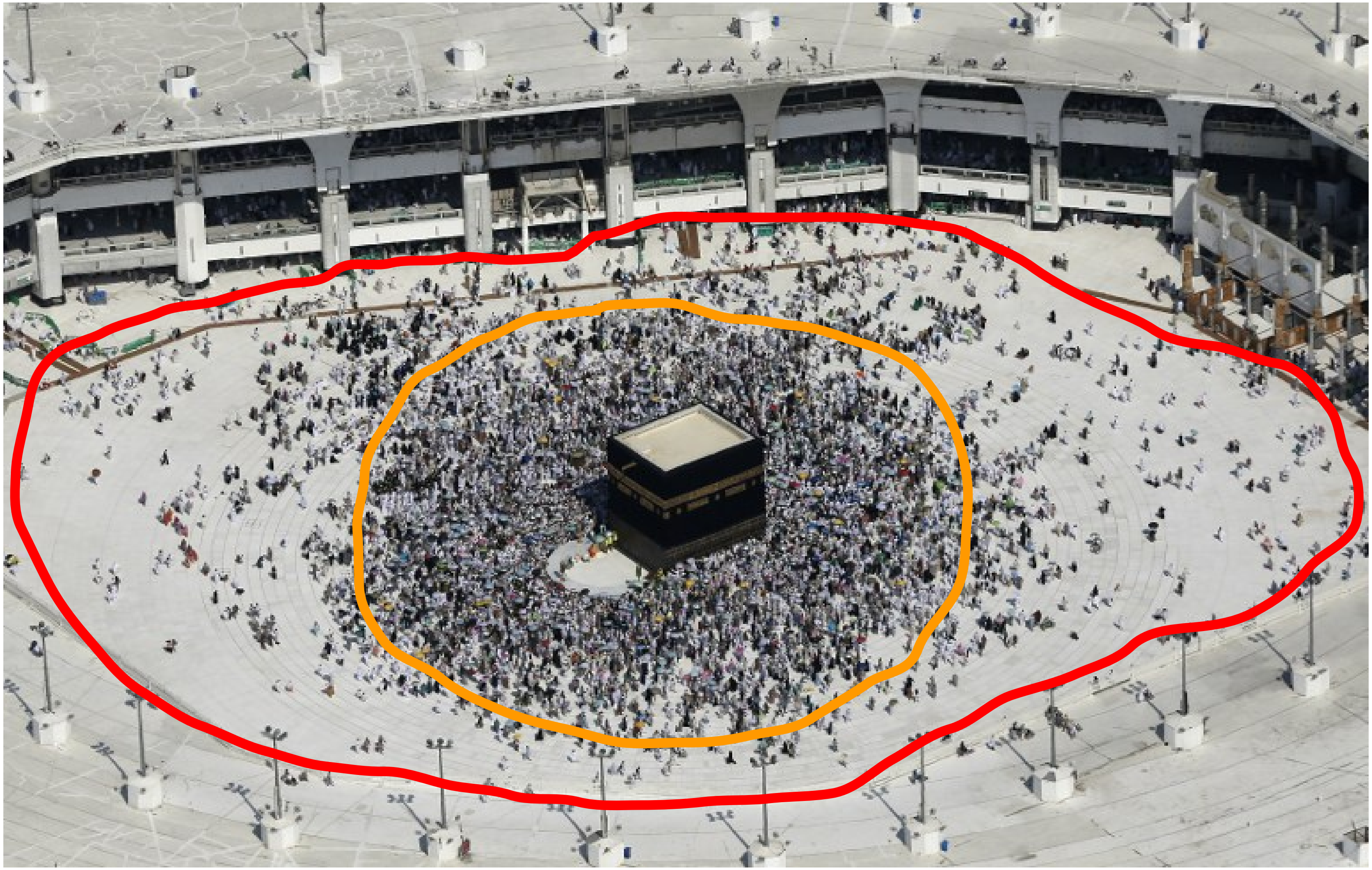


Photo: Ahmad Gharabli/AFP, 2016

Disadvantages: not automatic

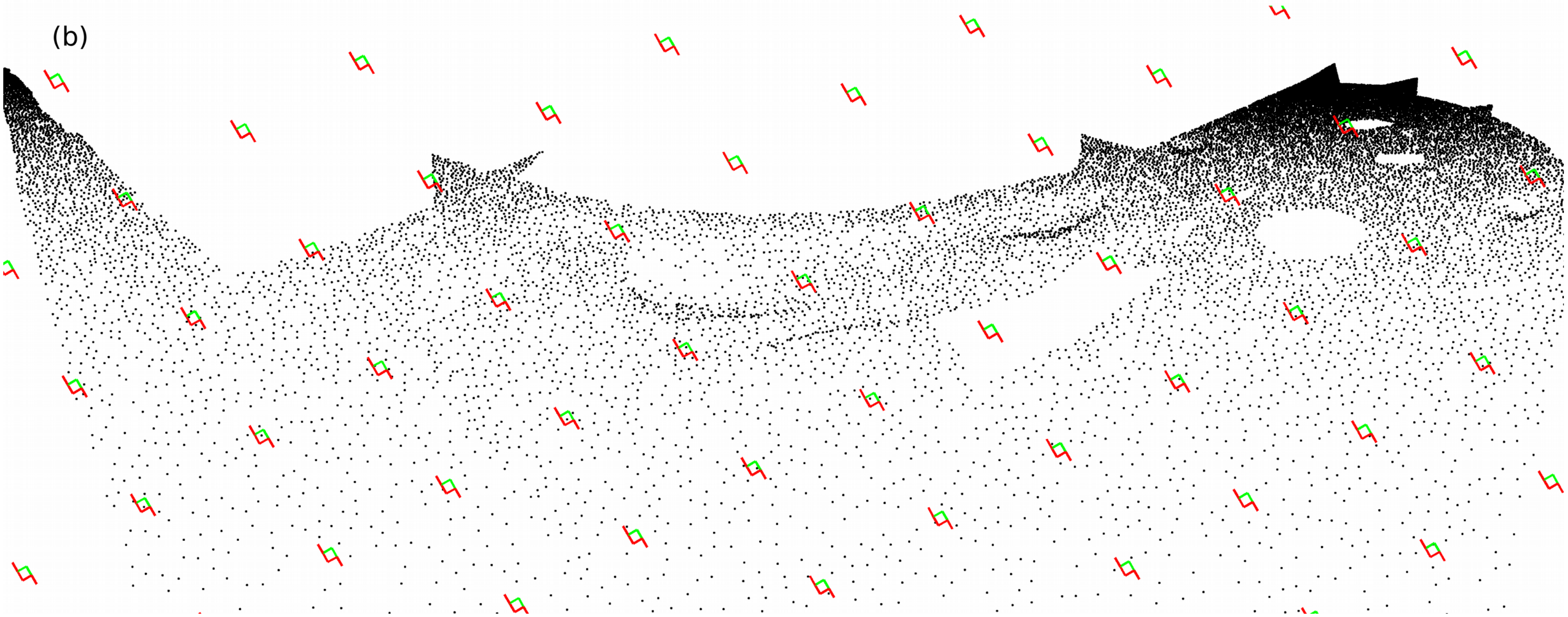
- Recent automatic alternatives:
  - Ground Truth: 4633
  - Idrees et al. 2013 estimation: 2550
  - Bansal & Venkatesh 2015 estimation: 1557.9

Gigapixel image

- Gigapanoramic picture Puerta del Sol 2015:

  http://lab.elespanol.com/estaticos/gigapan_sol/

Count*Em*1 on Gigapixel images

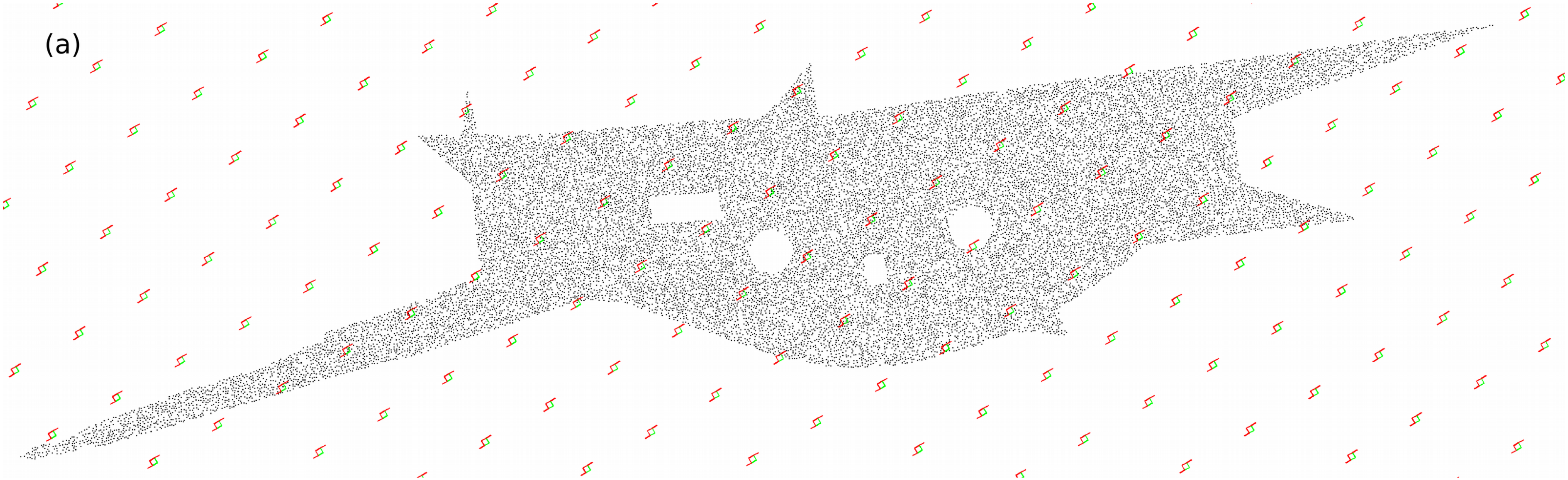


$$N = 2 \times 10^4 \text{ simulated points}$$

$$CE_E \approx 0.3$$

Cruz & González-Villa 2017

Count*Em*1 without perspective

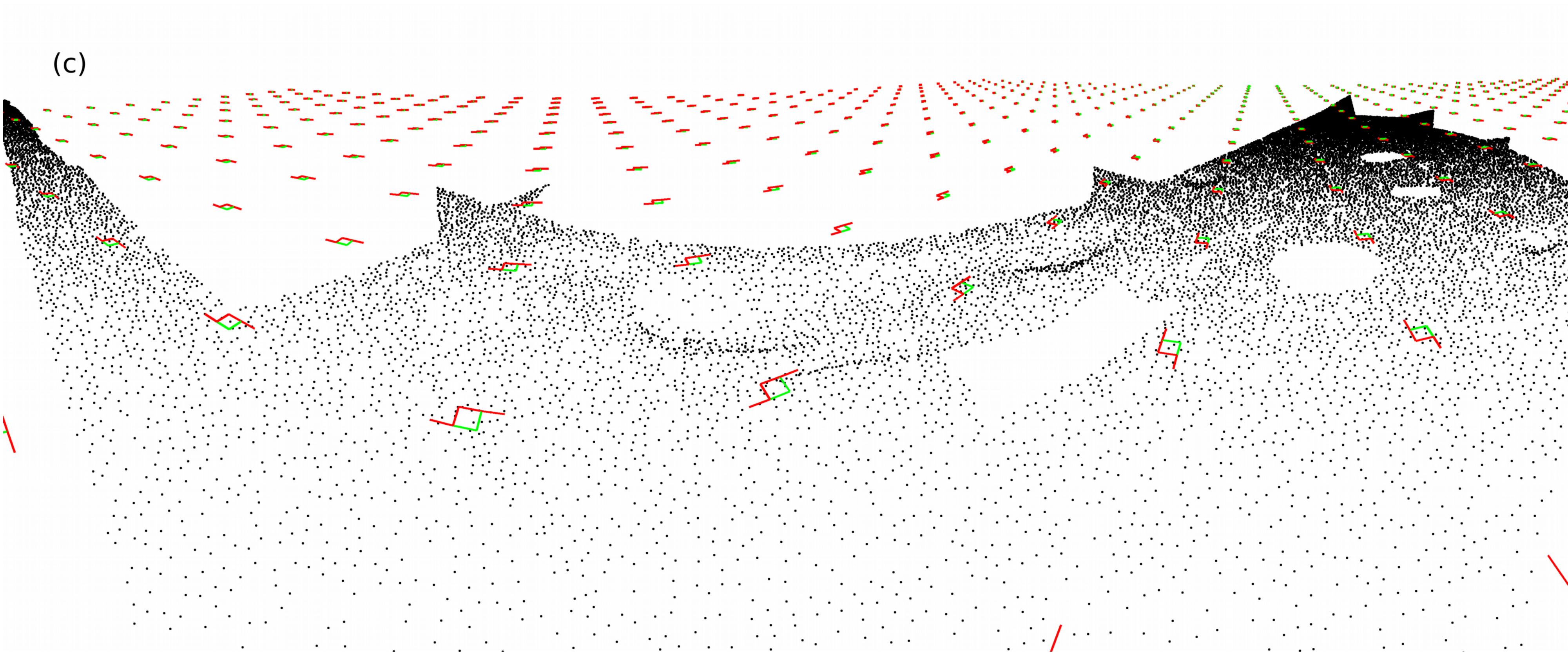


$$N = 2 \times 10^4 \textit{ simulated points}$$

$$CE_E \approx 0.07$$

Cruz & González-Villa 2017

Count*Em*2: Project grid onto image

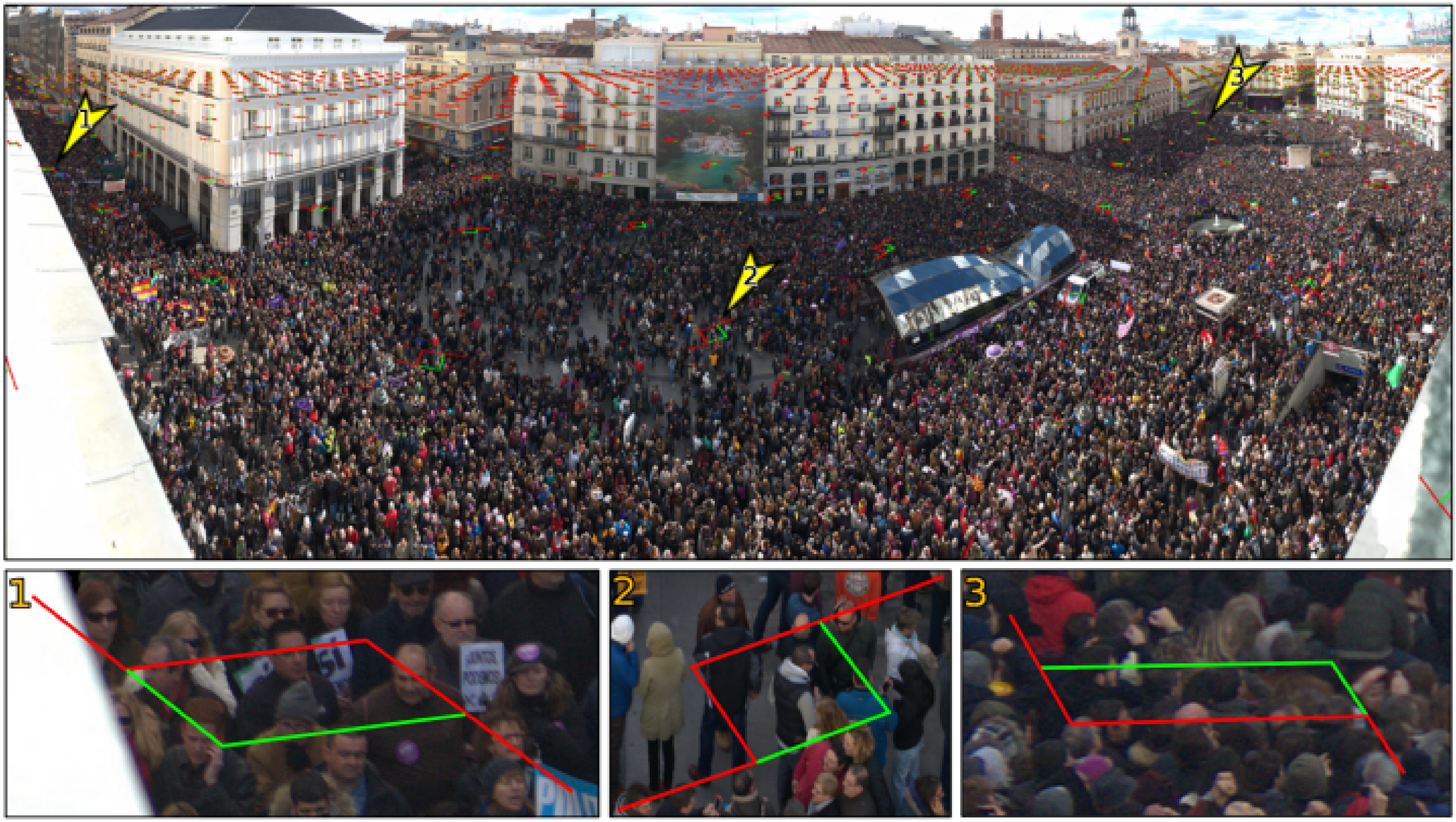


$$N = 2 \times 10^4 \; simulated\ points$$

$$CE_E \approx 0.07$$

Cruz & González-Villa 2017

Count*Em*2 on Gigapixel image



$$\widehat{N} = 29500$$

$$ce_{Cav} \approx 0.09$$

Cruz & González-Villa 2017

References

- http://countem.unican.es
- M Cruz, D Gómez, LM Cruz-Orive, PLOS ONE, 2015; 10(11):e0141868. DOI: 10.1371/journal.pone.0130314
- M Cruz, J González-Villa, 2017; submitted to IEEE Transactions on Image Processing
- HJG Gundersen, J.Microsc. 1977;111(2):219–223. DOI: 10.1111/j.1365-2818.1977.tb00062.x
- H Idrees, I Saleemi, C Seibert, M Shah, CVPR 2013 IEEE Conference proceedings, 2013, 2547–2554. DOI: 10.1109/CVPR.2013.329
- A Bansal, KS Venkatesh, IJCEE, 2015 DOI:10.17706/ijcee.2015.7.5.316-324