On Machine Astrometry

Methods for astrometric calibration of frames acquired by imaging detectors

by Filip Hroch

Masaryk University in Brno

DTPA Seminary, March 2012

Astrometry

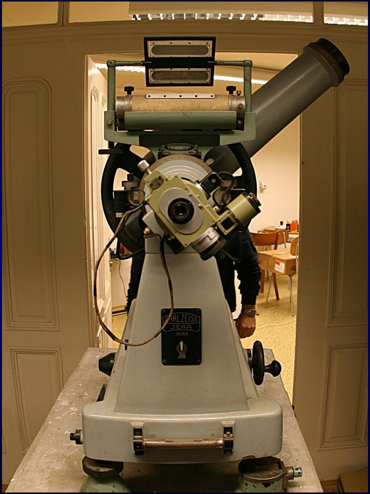
Astrometry is the branch of astronomy focused on precise measurements of the positions and movements of stars and celestial bodies. [1]

Applications

- localization of ”treasures”(discovered supernovas, comets, . . .)
- object motions (since Classical Antiquity!)
- fixing International Celestial Reference System (general usage)

. . . and many others

Fields of Interest

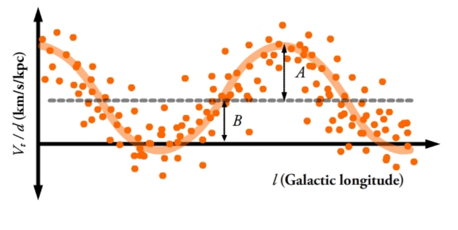


Kladivo Observatory

- Classical Astrometry
- Radio Astrometry
- Astrometric Catalogues
- Plate Astrometry: on photographic plates or digital chips

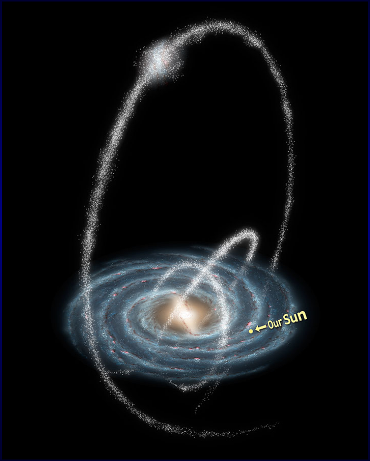
Applications of Plate Astrometry

Milky Way Structure



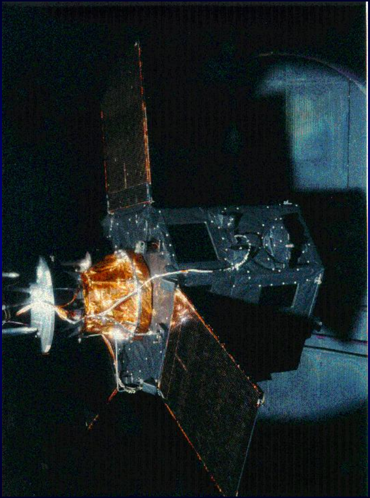
- differential rotation
- stellar streams
- estimate of size and mass
- gallactic halo

Stellar Streams



- Sloan Digital Survey[14]

Astrometry Support


[7]

The Hipparcos Space Astrometry Mission (1989 – 1993)

astrometry from the space, Hipparcos catalogue (1 mas accuracy, 118 218 stars), Tycho catalogue (20 – 30 mas accuracy, 2 539 913 stars)

UCAC3 (2009)

Catalogue covers all sky: total 139 million stars, average density 2000 stars per square degree, mean accuracy 15 to 20 mas ([8]), position is fixed by radio-interferometry to 0.1mas, widely used in navigation, geodetics, etc.

Machine Astrometry

Fully Automatic Astrometry

Main goal

A fully automatics detection of objects and astrometry on celestial frames

- fields of view: arcmin — full sky
- various instruments (deformations)
- executable without any mind (cogwheel machines)
- unaffected by unforeseen data

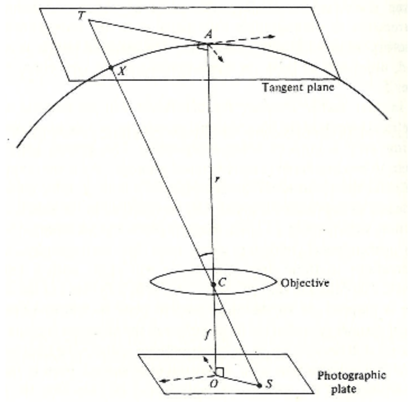
Machine Astrometry

Problem Statement

- pre-ordered lists of mutually equivalent objects (catalogue and stars detected on image) $(\alpha, \delta)_i$ correspond to $(x, y)_i$
- the correspondence on base of projection and transformation
- Modern view to astrometry: transformation, another set of parameters (changes: no plate-overlap method, six-constant plate, center corrections, . . .)
- Applications of modern statistical methods
- Developed matching (unimportant in photographic astrometry)

Gnomonic Projection

Rectilinear Projection



Gnomonic Projection[a]

- from center to tangent plane
- proper telescope projection
- displays great circles as straight lines
- approximation for a small field of view

$$
\begin{aligned}
u &= -(\alpha - \alpha_0)\cos\delta \\
v &= \delta - \delta_0
\end{aligned}
$$

[a][11]

Affine Transformation

Scale, rotation and shift

Center of Rotation

$$x_c = \frac{w}{2},$$
$$y_c = \frac{h}{2}.$$

Normalization ($-1 \leq \hat{x}, \hat{y} \leq 1$):

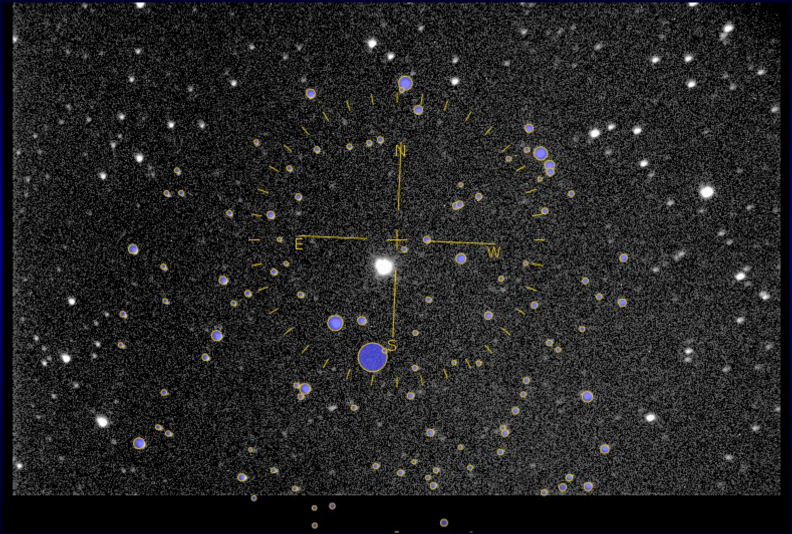
$$\hat{x} = \left(\frac{x}{x_c} - 1 \right),$$
$$\hat{y} = \left(\frac{y}{y_c} - 1 \right)$$

Rotation and scale c

$$\begin{pmatrix} X \\ Y \end{pmatrix} = c \left[\begin{pmatrix} X_0 \\ Y_0 \end{pmatrix} + \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} \right].$$

Parameters: X_0, Y_0, φ, c.

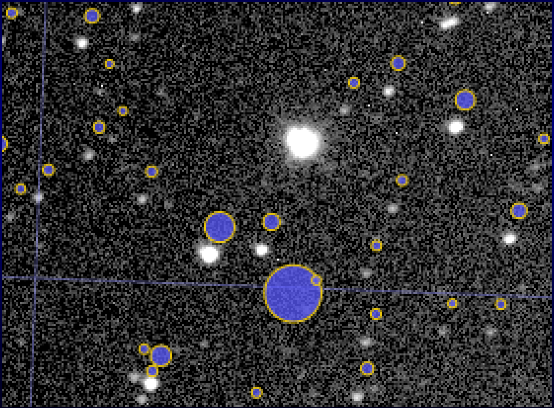
Projection and Transformation Live



Robust Astrometry

Outliers didn't matter[1]

"robustness signifies insensitivity to small deviations from assumptions"



Deviations caused by

- double stars
- moved stars
- defects
- predicable
- parameters

[1] Huber[2]

Robust Statistics

Tail of Outliers[2]

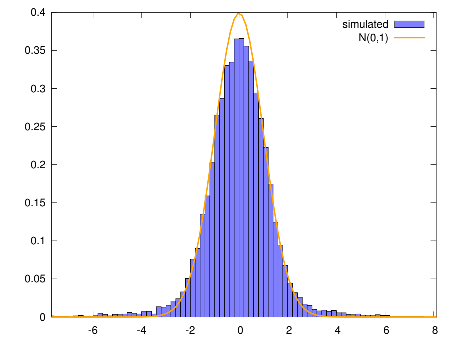
$$F(x) = (1-\varepsilon)\Phi\left(\frac{x-\mu}{\sigma}\right) + \varepsilon\Phi\left(\frac{x-\mu}{3\sigma}\right),$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-t^2/2}\mathrm{d}t$$

ε	s	d	σ
0.0	0.998	0.796	1.008
0.01	1.040	0.819	1.026
0.05	1.179	0.875	1.067
0.1	1.361	0.967	1.146
0.2	1.604	1.112	1.296
1.0	2.994	2.389	3.023

$$d = \frac{1}{N}\sum |x_i - \bar{x}|,$$

$$s = \sqrt{\frac{1}{N}\sum (x_i - \bar{x})^2}$$



$\varepsilon = 0.1$

[2] Example by Tukey(1960), Hubber(1980)

General Principles of Robust Statistics

Fisher's maximum likelihood:[3]

$$L = \prod_{i=1}^{N} p(x_i|\tilde{x}) = \prod_{i=1}^{N} f(x_i - \tilde{x}),$$

assumption $p(x|\tilde{x}) = f(x - \tilde{x})$ and substitution $\rho(x) = -\ln f(x)$

$$\ln L = -\sum_{i=1}^{N} \rho(x_i - \tilde{x})$$

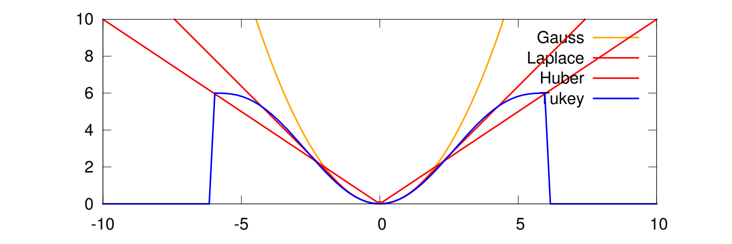
A standard way to look for minimum, $\psi(x) \equiv \rho'(x)$

$$\frac{\mathrm{d} \ln L}{\mathrm{d}\tilde{x}} = 0, \quad \sum_{i=1}^{N} \psi(x_i - \tilde{x}) = 0$$

[3] tilde $\tilde{x}$ denotes a robust estimator with contrast to the least square's $\bar{x}$

Remarkable Distributions

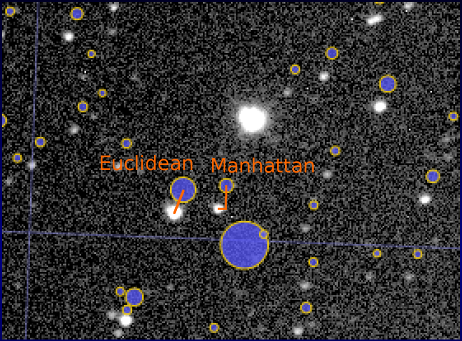
	$f = p(x\|0)$	$\rho = -\ln f$	$\psi = \rho'$
Gauss	$e^{-x^2/2}$	$\frac{x^2}{2}$	x
Laplace	$e^{-\|x\|}$	$\|x\|$	$\mathrm{sgn}(x)$
Huber	$e^{-\rho}$	$\begin{cases} -ax - a^2/2, & x < -a \\ x^2/2, & -a < x < a \\ ax - a^2/2, & x > a \end{cases}$	$\begin{cases} -a \\ x \\ a \end{cases}$
Tukey	$e^{-\rho}$	$\begin{cases} x^6/6c^4 - (x^2/2)(1 - x^2/c^2), \\ 0 \end{cases}$	$\begin{cases} x(1 - x^2/c^2)^2 \\ 0 \end{cases}$



Computational Methods of Robust Statistics

- Zero estimate by median: $\tilde{x}_0 = \mathrm{med}(x_i)$
- Estimate scatter (median for one-dimension, simplex for multidimensional)
$$s = \mathrm{med}(|x_i - \tilde{x}_0|)$$
- Robust estimator
$$\sum_i \psi\left(\frac{x_i - \tilde{x}}{s}\right) \to \mathrm{min.}$$
by Newton's method, Levendberg-Marquart (Minpack)
- Deviation
$$\sigma^2 = \frac{N}{N-1} s^2 \frac{(1/N)\sum_i \psi^2[(x_i - \tilde{x})/s]}{\{(1/N)\sum_i \psi'[(x_i - \tilde{x})/s]\}^2}$$

Measure for deviations



Euclidean distance

$$\sqrt{\sum_i [(u_i - x_i)^2 + (v_i - y_i)^2]}$$

Manhattan distance

$$\sum_i [|u_i - x_i| + |v_i - y_i|]$$

Distance by the robust distribution

$$\sum_i \rho(u_i, v_i, x_i, y_i), \quad \sum_i [\rho(u_i - x_i) + \rho(v_i - y_i)]$$

(Can be the product of probabilities the best choice?)

Robust Astrometry

Maximum likelihood:

$$L = \prod_{i=1}^{N} f\left(\frac{u_i - x_i}{s}\right) \cdot f\left(\frac{v_i - y_i}{s}\right).$$

Solution ($\Delta x_i = u_i - x_i, \Delta y_i = v_i - y_i$):

$$\begin{aligned}
\sum_i \psi\left(\frac{\Delta x_i}{s}\right) &= 0,\\
\sum_i \psi\left(\frac{\Delta y_i}{s}\right) &= 0,\\
\sum_i \left[\psi\left(\frac{\Delta x_i}{s}\right) x_i' + \psi\left(\frac{\Delta y_i}{s}\right) y_i'\right] &= 0,\\
\sum_i \left[\psi\left(\frac{\Delta x_i}{s}\right) y_i' - \psi\left(\frac{\Delta y_i}{s}\right) x_i'\right] &= 0.
\end{aligned}$$

Deviations in Robust Astrometry

Residual sum

$$S_0 = \sum_{i=1}^{N} \left\{ \rho\left(\frac{\Delta x_i}{s}\right) + \rho\left(\frac{\Delta y_i}{s}\right) \right\},$$

Covariance matrix

$$c_{kl} = \left(\frac{N}{N-M}\right) s^2 \frac{\frac{1}{N}\sum_i \psi(d_i)^2}{\left\{\frac{1}{N}\sum_i \psi'(d_i)\right\}^2} a_{kl}^{-1},$$

Deviations of parameters $\tilde{\sigma}_i$

$$\tilde{\sigma}_i^2 = \frac{S_0}{N-P} c_{ii} \qquad \text{for } i = 1 \dots P,$$

The Algorithm of Robust Astrometry

On Input: starlist x_i, y_i, catalogue: α_i, δ_i, center of projection α_0, δ_0

Iterative algorithm:

1. projection
2. estimate parameters
3. inverse to correct center
4. next iteration

On Output: $\alpha_0, \delta_0, c, \varphi$, a statistical description

Nights in Monte Carlo

- Dice: generating $\mathscr{U}$ uniform distribution $1, \ldots 6$
- 100,000 - 1,000,000 events
- Normal $x \in N(0,1)$ distribution from Uniform $y \in \mathscr{U}$: Quantile function (inverse to cumulative function):

$$\begin{aligned} \Phi(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-t^2/2} \mathrm{d}t = \frac{1}{2}\left(1 + \operatorname{erf}(\tfrac{x}{\sqrt{2}})\right) \\ \Phi^{-1}(x) &= \sqrt{2}\, \operatorname{erf}^{-1}(2\Phi(y) - 1) \end{aligned}$$

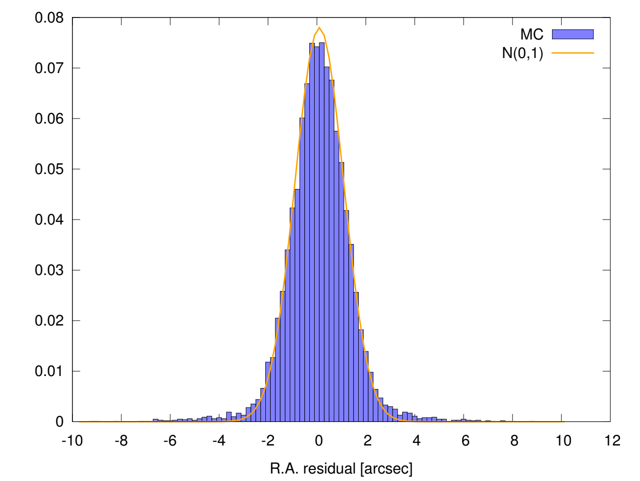
Shapiro–Wilk test

$$W = \frac{\left(\sum_{i=1}^{N} a_i x_i\right)^2}{\sum_{i=1}^{N} (x_i - \bar{x})^2}$$

. . . the null hypothesis is that the population is normally distributed, if the p-value is less than the chosen alpha level . . .

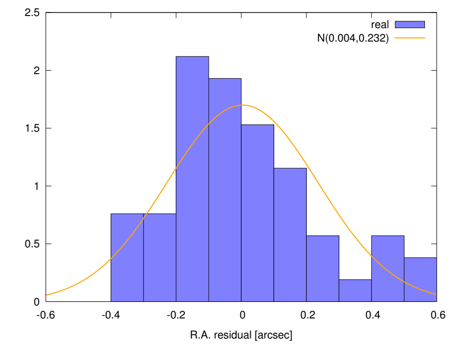
Results of Simulated Astrometry

$\varepsilon = 0.1$, Monte Carlo simulation (MC) $\sigma_\alpha = 1$ robust $\sigma_\alpha = 0.81$, least-square $\sigma_\alpha = 1.85$



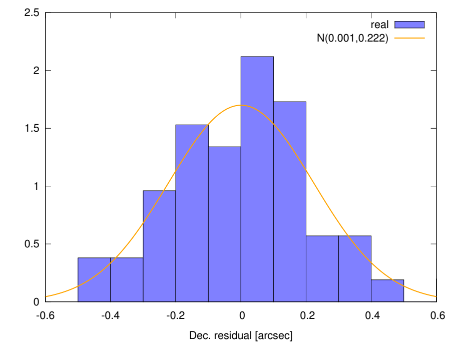
Results of Astrometry on Real Data

Right Ascension deviations



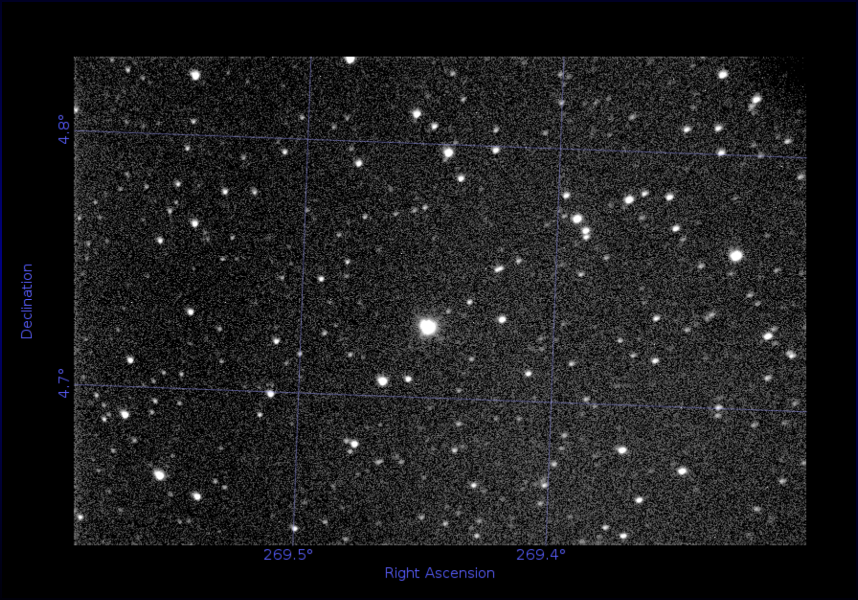
Shapiro-Wilk: $p = 0.01$ (rejected as Normal distribution)

Declination deviations

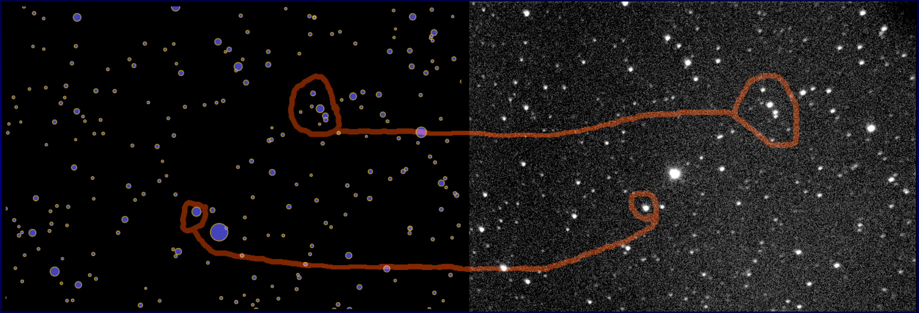


Shapiro-Wilk: $p = 0.85$ (accepted as Normal distribution with 85% significance)

Calibrated Plate



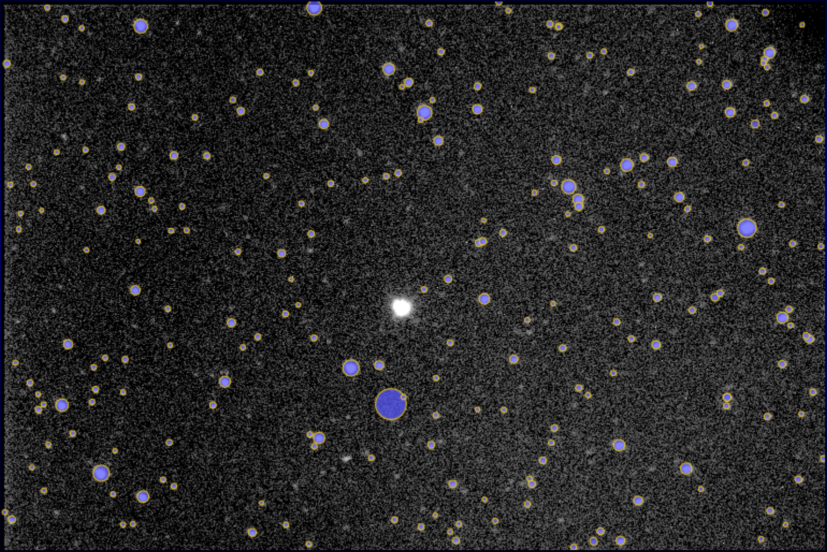
Searching for mutually matching objects



Looks trivial . . .

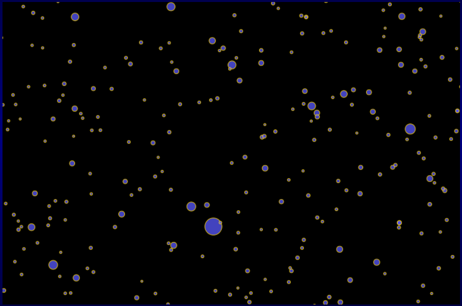
Problem Statement

Catalogue Overlay



Problem Statement

Graphical View



Problem Statement

Machine View

α	δ	x	y
269.32877180	4.75990450	691.460	301.210
269.46645390	4.70561230	323.246	171.016
269.55306180	4.66534920	90.749	72.835
269.48413950	4.83174120	289.181	506.877
269.44389680	4.79628500	391.942	409.028
269.39211090	4.77216420	525.995	340.089
269.54502150	4.82300590	127.843	489.439
269.56777920	4.68875310	54.229	136.187
269.37191650	4.78052750	580.081	360.176
269.53804240	4.65767420	129.737	50.999
269.45703000	4.81114700	358.725	449.405
269.33678980	4.83109000	677.341	490.346
269.54718390	4.79450530	119.210	414.265

Nightmare for machines . . .

The Matching Problem

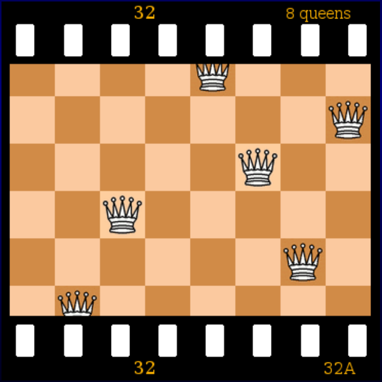
- crucial (pre-astrometry) problem
- hell of permutations:

N	$N!$
5	120
10	3628800
15	1307674368000
20	2432902008176640000
25	15511210043330985984000000
30	265252859812191058636308480000000

- generating permutations
- recognizing true permutations
- insensitivity to geometrical transformations

Eight queens problem

The eight queens puzzle is the problem of placing eight chess queens on an 8×8 chessboard so that no two queens attack each other. [2]



Features:

- rules for queens
- non-analytical
- back-tracking algorithm

Eight queens problem

The Backtracking algorithm prototype

```
Function Try i-th queen
  Map possible positions for i-th queen
  For All the positions:
     Is acceptable?:
       Move i-th queen to the position

       Having free queens?
         Call Try (i+1)-queen
         Did fail Try ?
           Undo the move.
  end for
End function Try
```

Matching by Backtracking

General Backtracking Algorithms

- recursive check all possibilities
- depth is limited by heuristics
- similar problems: crosswords, sudoku, stable marriage (secretary, roommates) problem, Constraint satisfaction problem

Matching Backtracking Algorithms

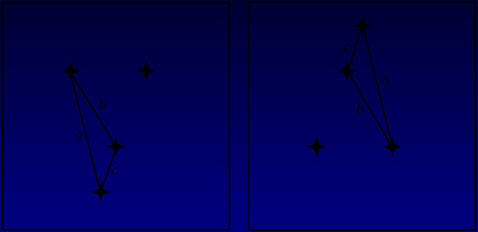
- recursive check all star list permutations
- heuristic use a metrics

Complications for spherical or rectangular coordinates:

- supported sets: $\mathbb{N}^2 \to \mathbb{R}^2$ (both on an interval)
- polluted by statistical errors
- projected and transformed

Triangle Space

The Definition



Lets:

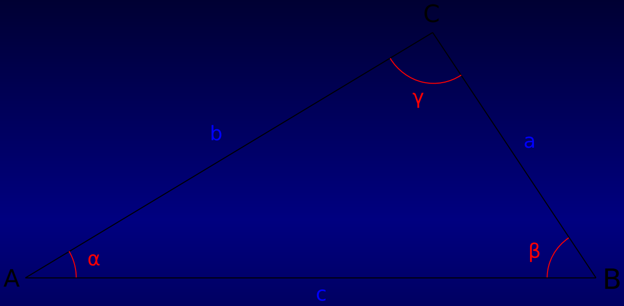
$$d_1 \leq d_2 \leq d_3$$

$$u \equiv \frac{d_1}{d_3}, \qquad v \equiv \frac{d_2}{d_3}$$

- Using triangle similarity $\triangle ABC \sim \triangle DEF$
- u, v is independent on affine transformation
- needs projection

Triangle Space

Properties

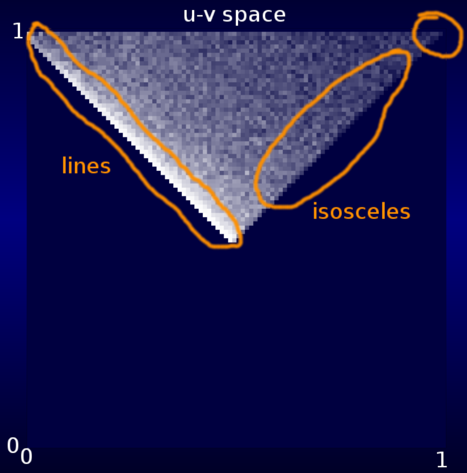


- $u, v \in \mathbb{R} | 0 \leq u, v \leq 1$ (due to ordering)
- Law of cosines:

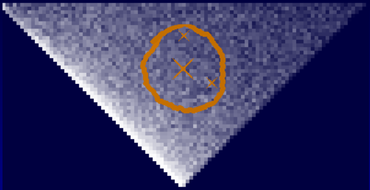
$$1 = u^2 + v^2 - 2uv\cos\gamma$$

- equilateral: $u = v = 1$
- isosceles: $u = v < 1$
- reduces to line, limit: $v \leq 1 - u$, $d_3 \leq d_1 + d_2$ (right just only for line)

Star Distribution in Triangle Space



Backtracking in Triangle Space



Principle

- equivalent triangles has close u, v
- but use of single triangles is unreliable, in neighborhood is $p_k =$ triangles (for uniform distribution)
- use of a sequence of stars > 3
- searching by back-tracking

Analogy

- chess desk: u, v-space
- possible moves: given by star list
- rejections: by neighborhood

Matching Algorithm

```
Function Sequence for next star
  For all unused stars in catalogue:
     Compute u1,v1
     For all unused stars on image:
        Compute u2,v1
        Is Acceptable and distance({u1,u2} - {v1,v2}) < limit?
            Has the sequence required length?
              Got Solution!
            Call Sequence for next+1 star
            Did fail the Sequence?
              Skip the star
End function Sequence
```

Properties of Matching Algorithm

- possibly generates all permutations of set N (common to both lists)!
- the number is reduced by rejecting over distance limit $d_{\max}$
- the solution is checked by dispersion of scale

$$\sigma_c^2 = \text{mad}\left(\frac{d_{\text{star},i}}{d_{\text{catalogue},i}}\right)$$

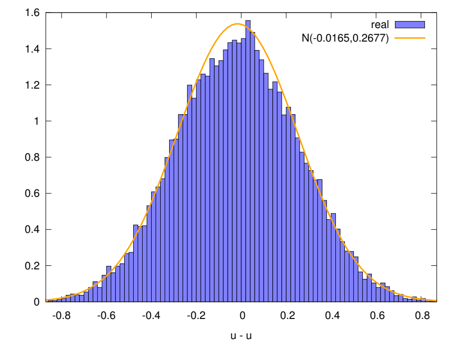
Estimate of search radius

- x, y, α, δ has normal distribution $N(0,1) = e^{-x^2/2}$
- u, v has approximately $N(0, \sigma > 1)$
- $\sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2}$ has χ^2 (?)
- for 90% probability, the radius is determined as....

Statistical properties of differences

Distribution of $u_c - u_0$

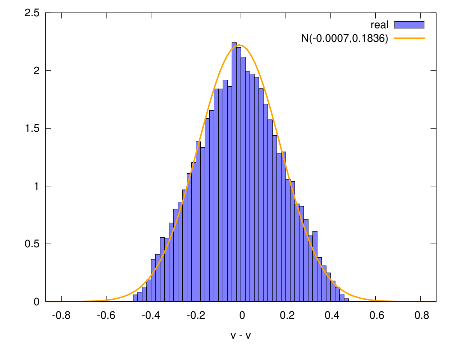
$$N(-0.0165, 0.2677)$$



Shapiro-Wilk: $p = 10^{-7}$ (rejecting)

Distribution of $v_c - v_0$

$$N(-0.0165, 0.2677)$$



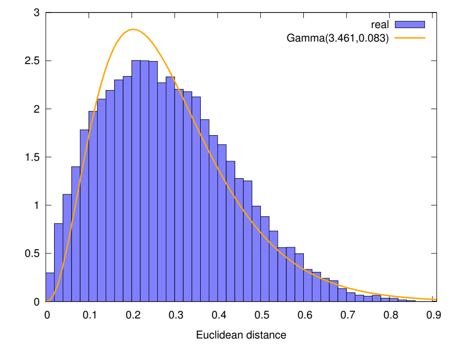
Shapiro-Wilk: $p = 10^{-25}$ (rejecting)

Statistical Properties of Distance Measures

Euclidean distance

$$\sqrt{(u_c - u_0)^2 + (v_c - v_0)^2}$$

$$\Gamma(-0.0165, 0.2677)$$

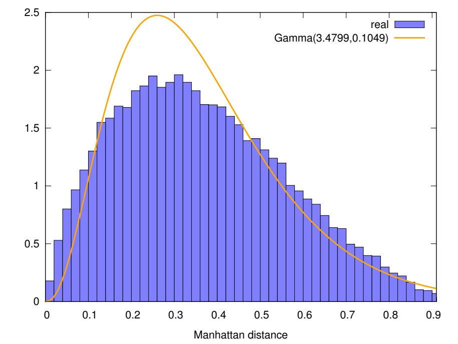


Shapiro-Wilk: $p = 0$ (rejecting)

Manhattan distance

$$|v_c - v_0| + |u_c - u_0|$$

$$\Gamma(-0.0165, 0.2677)$$



Shapiro-Wilk: $p = 0$ (rejecting)

Matching Live

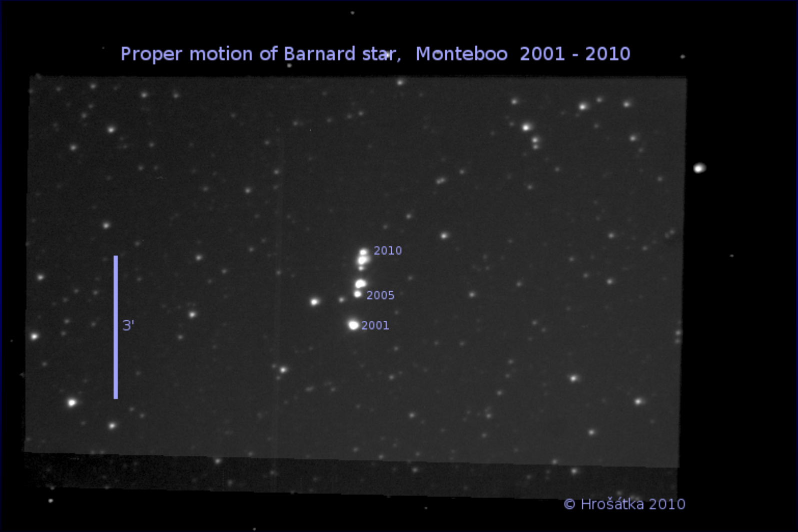
Resume

- robust astrometry (transformation + statistical methods)
- matching on base of back-tracking
- support for Virtual observatory
- Are Genetic Algorithms false track?
- matching vs. registering

- Astrometrica (new project, similar matching)
- Sextractor (classical project, triangles)
- Focas (classical project, triangles, Iraf)

Barnard's star

Monteboo 2001-2010

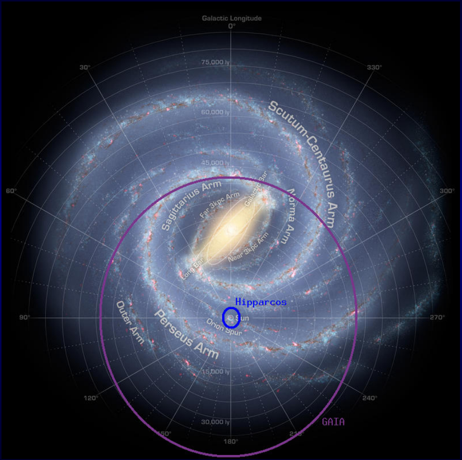


Close To The Edge

- expansion: more projections (spherical, flipping),
- implementation of photometric calibration in analogy to astrometry
- wide-field cameras
- support for Virtual Observatory
- MonteBoo archive
- Detection of UFOs

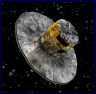
GAIA

Successor of Hipparcos



Mission Objectives[a]

- accuracy down to 20μas (!) (distance up to 50 kpc)
- 1 billion stars (Milky Way + local group)
- launch 2013 (five years duration)
- create a three-dimensional structural map
- ESA collaboration

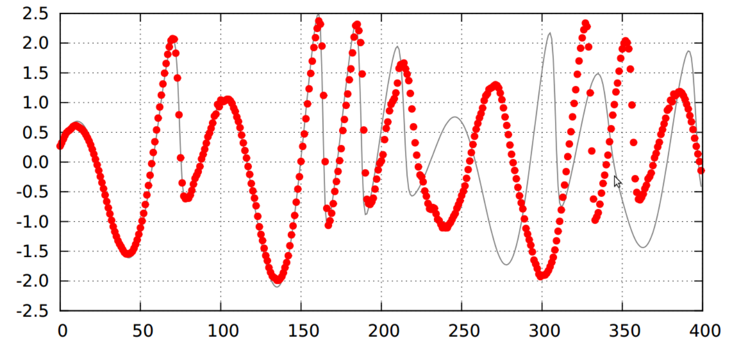


[a] [13]

Czech Center of GAIA

Advertising

- principal investigator: P.Koubský
- working group: V. Votruba, Z.Janák

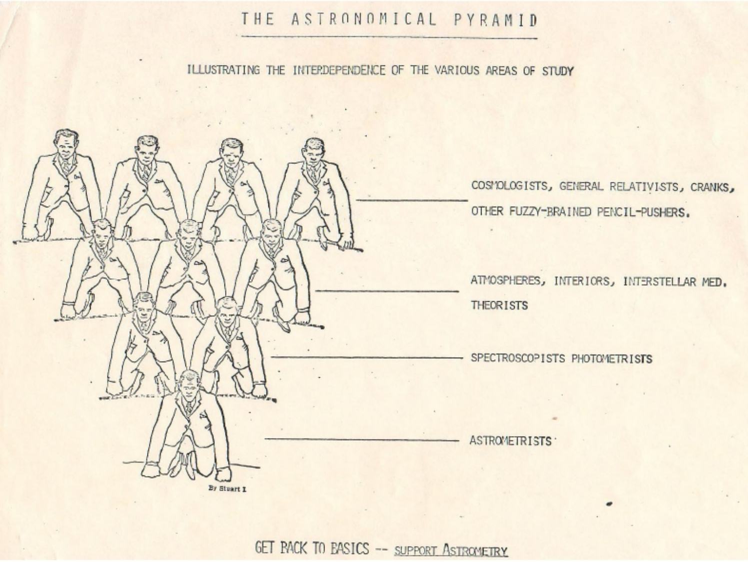


Data mining of light curves (neural networks, clustering, . . .)[4]

[4] [1]

The Astronomical Pyramid

Pencil-pushers on shoulders of astrometrists or vice versa?



[10]

References I.

- http://en.wikipedia.org/wiki/Astrometry
- http://en.wikipedia.org/wiki/Eight_queens_puzzle
- http://www.mlahanas.de/Greeks/PtolemyAstronomy.htm
- http://www.univie.ac.at/hwastro/rare/1515_ptolemae.htm
- http://adsabs.harvard.edu/abs/1978JHA.....9...42W
- http://www.astro.uni-bonn.de/~geffert/ge/arg/arg.htm
- http://www.rssd.esa.int/index.php?project=HIPPARCOS
- http://ad.usno.navy.mil/ucac/readme_u3.html
- http://ad.usno.navy.mil/ucac/ctio.html
- http://www.scholarpedia.org/article/Astrometry
- http://en.wikipedia.org/wiki/Gnomonic_projection
- http://chandra.si.edu/photo/2012/a520/
- http://sci.esa.int/science-e/www/object/index.cfm?fobjectid=31197
- http://en.wikipedia.org/wiki/List_of_stellar_streams

References II.

- Z.Janák, Diploma thesis, http://is.muni.cz/th/106451/prif_m/diplomka.pdf
- P. J. Huber: Robust Statistics, Willey (2004)
- http://what-when-how.com/space-science-and-technology/optical-astrometry-from-space/

http://www.physics.muni.cz/~hroch/amachine.pdf
http://munipack.physics.muni.cz