Several inlier structures and outliers... each structure with a different scale... the scales also have to be found. How do you solve it?


Three distinct steps. Some problems not resolved.
* A scale estimate is obtained with all undetected structures contributing as well.
* In the second step, the mean shift is more complicated than it should be.
* The stopping criterion is just a heuristic relation.

A different solution: simpler, avoids the above problems. Will examine the criterions when the algorithm is robust.
Nonlinear objective functions

The $n_1$ inlier measurements $y_i \in \mathbb{R}^l$ with $n_1 < n$, in general satisfy a nonlinear objective function

$$
\Psi (y_i, \beta) \simeq 0, \quad i = 1, \ldots, n_1 \quad \Psi (\cdot) \in \mathbb{R}^k
$$

$\Psi (y_i, \beta)$ outliers $i = (n_1 + 1), \ldots, n$.

We can separate $\Psi$ into a matrix of measurements and the new parameter vector

$$
\Psi (y_i, \beta) = \Phi (y_i) \theta (\beta) \quad \Phi (y) \in \mathbb{R}^{k \times m} \quad \theta (\beta) \in \mathbb{R}^m
$$

and interpret it in a higher dimensional linear space as the relations between $\zeta$ carriers $x_i^{[c]}$ and $\theta$. Gives

$$
x_i^{[c] \top} \theta \simeq 0 \quad c = 1, \ldots, \zeta \quad i = 1, \ldots, n_1.
$$

Ambiguity of the equation is eliminated taking

$$
\theta \top \theta = 1.
$$

This is a necessary condition for Grassmann manifold.


gives a procedure to refine the existing estimate for gpbM. The procedure can be used for our algorithm too and will not be described.

**Ellipse estimation.** $\zeta = 1$.

Input: $[x \ y]^\top \in \mathbb{R}^2$

Nonlin. obj. funct.: $(y_i - y_c)^\top Q (y_i - y_c) - 1 \simeq 0$

where matrix $Q$ is $2 \times 2$ symmetric, positive definite

and $y_c$ is the ellipse center. $y_i$ are measurements!

Carrier: $x = [x \ y \ x^2 \ xy \ y^2]^\top \in \mathbb{R}^5$

with relation between $\theta$ and the input parameters

$$
\theta = [-2y_c^\top Q \ Q_{11} \ 2Q_{12} \ Q_{22}]^\top \quad \alpha = y_c^\top Q y_c - 1
$$

equivalent to

$$
x_i^\top \theta - \alpha \simeq 0.
$$

The scalar $\alpha$ (intercept) was pulled out.

Beside $\theta \top \theta = 1$ the ellipses also have to satisfy the positive symmetry condition

$$
4 \theta_3 \theta_5 - \theta_4^2 > 0 \quad (= 1).
$$

Perturbing $x$ with standard deviation $\sigma$ relative to $x_o$ is

$$
E (x - x_o) = [0 \ 0 \ \sigma^2 \ 0 \ \sigma^2]^\top
$$

since the carrier has $x^2, y^2$ terms.
Fundamental matrix. \( \zeta = 1 \).
Input: \( [x \ y \ x' \ y']^\top \in \mathbb{R}^4 \)
Nonlin. obj. funct.: \( y_i^\top F y_i = [x'_i \ y'_i \ 1] F [x_i \ y_i \ 1]^\top \simeq 0 \)
Carrier: \( x = [x \ y \ x' \ y' \ xx' \ xy' \ x'y' \ yy']^\top \in \mathbb{R}^8 \)
Gives the linear relation
\[
x_i^\top \theta - \alpha \simeq 0 \quad i = 1, \ldots, n_1.
\]

Homography. \( \zeta = 2 \).
Input: \( [x \ y \ x' \ y']^\top \in \mathbb{R}^4 \)
Nonlinear obj. funct.: \( y'_i \simeq H y_i \) or
\[
[x'_i \ y'_i \ w'_i]^\top \simeq H [x_i \ y_i \ 1]^\top
\]
Direct linear transformation (DLT): with \( \theta = h = \text{vec}[H^\top] \) and \( A_i \) a \( 2 \times 9 \) matrix
\[
A_i h = \begin{bmatrix}
-x_i^\top & 0_3^\top & x'_i y_i^\top \\
0_3^\top & -y_i^\top & y'_i y_i^\top
\end{bmatrix}
\begin{bmatrix}
h_1 \\
h_2 \\
h_3
\end{bmatrix} \simeq 0_2.
\]
Carriers: \( x^{[1]} = [-x \ -y \ -1 \ 0 \ 0 \ 0 \ x'x \ x'y \ x']^\top \)
\( x^{[2]} = [0 \ 0 \ -x \ -y \ -1 \ y'x \ y'y \ y']^\top \)
and the two linear relations can be written as
\[
x_i^{[c]} h \simeq 0 \quad c = 1, 2 \quad i = 1, \ldots, n_1.
\]

Covariance of the carriers
The inliers at input have the same \( l \times l \) covariance \( \sigma^2 C_y \),
where \( \sigma^2 \) is unknown and \( C_y = I_y \) if no additional information is available.
The \( m \times m \) covariances of the carriers are
\[
\sigma^2 C^{[c]}_i = \sigma^2 J_{x^{[c]}_i|y_i} C_y J_{x^{[c]}_i|y_i}^\top \quad c = 1, \ldots, \zeta
\]
where the Jacobian matrix is
\[
J_{x|y} = \frac{\partial x(y)}{\partial y} = \begin{bmatrix}
\frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_l} \\
\cdots & \cdots & \cdots \\
\frac{\partial x_m}{\partial y_1} & \cdots & \frac{\partial x_m}{\partial y_l}
\end{bmatrix}
\]
Each carrier covariance depends on the input point from where it was derived.

Ellipse estimation. \( \zeta = 1 \).
The \( 5 \times 2 \) Jacobian gives the \( 5 \times 5 \) covariance
\[
J_{x_i|y_i}^\top = \begin{bmatrix}
1 & 0 & 2x_i & y_i & 0 \\
0 & 1 & 0 & x_i & 2y_i
\end{bmatrix}
\]
Fundamental matrix. \( \zeta = 1 \).
The \( 8 \times 4 \) Jacobian matrix gives the \( 8 \times 8 \) covariance
\[
J_{x_i|y_i}^\top = \begin{bmatrix}
1 & 0 & 0 & x'_i & y'_i & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & x'_i & y'_i \\
0 & 0 & 1 & x_i & 0 & y_i & 0 \\
0 & 0 & 0 & 1 & 0 & x_i & 0 & y_i
\end{bmatrix}
\]

Homography. \( \zeta = 2 \).
The two \( 9 \times 4 \) Jacobian matrices give the \( 9 \times 9 \) covariances
\[
J_{x_i[1]|y_i}^\top = \begin{bmatrix}
-I_{2\times2} & x'_i I_{2\times2} & 0_2 \\
0_2 & 0_{4\times4} & y_i^\top \\
0_2 & 0_{4\times3} & 0_2 \\
0_2 & 0_{4\times2} & y_i^\top
\end{bmatrix}
\]
\[
J_{x_i[2]|y_i}^\top = \begin{bmatrix}
-I_{2\times2} & y_i I_{2\times2} & 0_2 \\
0_2 & 0_{4\times4} & y_i^\top \\
0_2 & 0_{4\times3} & 0_2 \\
0_2 & 0_{4\times2} & y_i^\top
\end{bmatrix}
\]

If \( \zeta > 1 \), will take for each elemental subset \( \theta, \alpha \), only the largest Mahalanobis distance
\[
d_{\tilde{i}}^{[c]} = \frac{\| x_{\tilde{i}}^{[c]} \theta - \alpha \|}{\sqrt{\theta^\top C_{\tilde{i}}[c] \theta}} \]
\[
\tilde{c}_i = \max_{c = 1, \ldots, \zeta} d_{\tilde{i}}^{[c]}
\]

The carrier vector is \( \tilde{x}_i \), covariance matrix is \( \tilde{C}_i \), variance is \( \tilde{H}_i = \theta^\top \tilde{C}_i \theta \) and the null space is one-dimensional.

Scale estimation for an inlier structure
Take \( M \) elemental subsets. Then \( n_{r1} \ll n \) points from
\[
\min_M \sum_{i=1}^{n_{r1}} \tilde{d}[i] \]
gives the initial set.
Taking \( \epsilon = 5\% \) is enough. All structures have \( n_{r1} = 0.05n \), where \( n \) is the total number of input points.
Initial set has \( n_{r1} \) points. The largest distance is \( \tilde{d}_{[r1]} \).
Increase to \( 2 \times \tilde{d}_{[r1]} = \tilde{d}_{[r2]} \). Find \( n_{r2} \) and continue...
Assume that at \( b \times \tilde{d}_{[r1]} = \tilde{d}_{[rb]} \) will satisfy
\[
2[n_{rb} - n_{r(b-1)}] \leq \frac{n_{r(b-1)}}{b - 1}
\]
from where \( \tilde{\sigma} = \tilde{d}_{[rb]} \) because the border is not precise.

\[
y = x \in \mathbb{R}^2 \quad x \in \mathbb{R}^5
\]
search is in \( \mathbb{R} \), the null space
**Inlier estimation with mean shift**

Take $N = M/10$ elemental subsets from the inlier points. Do mean shift. Converges to the closest mode.

\[
\begin{align*}
[\hat{\theta}, \hat{\alpha}] &= \arg \max_{\theta, \alpha} \frac{1}{n \sigma} \sum_{i=1}^{n} \kappa \left( (z - \tilde{z}_i)^\top \tilde{B}_i^{-1} (z - \tilde{z}_i) \right) \\
&= \frac{1}{n \sigma} \arg \max_{\theta} \left( \arg \max_{z} f_\theta(z) \right)
\end{align*}
\]

with the variance $\tilde{B}_i$ being

\[
\tilde{B}_i = \hat{\sigma}^2 \tilde{H}_i = \hat{\sigma}^2 \theta^\top \tilde{C}_i \theta
\]

\[
= \hat{\sigma}^2 \theta^\top \tilde{J}_{\tilde{x}_i|\tilde{y}_i} \tilde{J}_{\tilde{x}_i|\tilde{y}_i}^\top \theta
\]

and $\kappa(u) = K(u^2), u \geq 0$. The Epanechnikov kernel

\[
\kappa(u) = \begin{cases} 
1 - u & (z - \tilde{z}_i)^\top \tilde{B}_i^{-1} (z - \tilde{z}_i) \leq 1 \\
0 & (z - \tilde{z}_i)^\top \tilde{B}_i^{-1} (z - \tilde{z}_i) > 1
\end{cases}
\]

has $g(u) = -\kappa'(u) = 1$ for $u \leq 1$ and zero for $u > 1$. All currently available input points are processed $\tilde{z}_i = \tilde{x}_i^\top \theta$, $i = 1, \ldots, n$ in an iteration around $z = z_{old}$

\[
z_{new} = \left[ \sum_{i=1}^{n} g(u) \right]^{-1} \left[ \sum_{i=1}^{n} g(u) \tilde{z}_i \right]
\]

Number of inliers: $n_{in}$. TLS estimate: $\theta^{tls}$ and $\hat{\alpha}^{tls}$.

**Strength based merge process**

The strength of a structure $s$ is defined as

\[
\tilde{d} = \frac{\sum_{i=1}^{n_{in}} d_i}{n_{in}} \quad s = \frac{n_{in}}{d} = \frac{n_{in}^2}{\sum_{i=1}^{n_{in}} d_i}
\]

The $j$th structure has: $n_j$ inliers and strength $s_j$.

Before $l = 1, \ldots, (j - 1)$ structures were detected:

$n_l$ inliers and strength $s_l$.

TLS estimate for $j$ and $l$ together gives strength $s_{j,l}$ and will be fused if

\[
s_{j,l} > \frac{n_j s_j + n_l s_l}{n_j + n_l} \quad \text{for an } l.
\]

For the nonlinear objective function merges should be done in the input space, whenever is possible.

300 inliers. 300 outliers. Gaussian inlier noise $\sigma_g = 20$.

$M = 500$. Three similar inlier structures.

Final structure: $\hat{\sigma} = 46.1$ and 329 inliers.
**Final classification**

Continue until you don’t have enough points to start another initial set.

The structures are sorted by their strengths in descending order. The inlier structures will be always at the beginning of the list.

The user is able to specify where to stop and retain only the inliers, which have denser structures.

\[ \theta_1 x_i + \theta_2 y_i - \alpha \simeq 0 \quad i = 1, \ldots, n_{in} \]

\[ M = 300 \] for all multiple line estimations.

100 points, \( \sigma_y = 5 \)
200 points, \( \sigma_y = 10 \)
300 points, \( \sigma_y = 20 \)

400 unstructured outliers

The correct inliers are recovered 97 times from 100. Three times the ”100 points line” didn’t recover.

\[ \text{scale : } 11.5 \quad 16.3 \quad 40.1 \quad 322.2 \]

\[ \text{structure : } 136 \quad 202 \quad 306 \quad 318 \]

\[ \text{strength : } 11660 \quad 9996 \quad 7202 \quad 832. \]

The correct inliers are recovered 97 times from 100. Three times the ”100 points line” didn’t recover.

\[ \text{scale : } 12.06 \pm 2.94 \quad 18.90 \pm 4.32 \quad 35.56 \pm 10.35 \]

\[ \text{inliers : } 121.4 \pm 17.8 \quad 225.7 \pm 32.1 \quad 288.3 \pm 44.9 \]
\[ \theta_1 x_i + \theta_2 y_i - \alpha \simeq 0 \quad i = 1, \ldots, n_{in} \]

100 points, \( \sigma_g = 20 \)
200 points, \( \sigma_g = 10 \)
300 points, \( \sigma_g = 5 \)

\[ (y_i - y_c)^\top Q (y_i - y_c) - 1 \simeq 0 \quad i = 1, \ldots, n_{in} \]

\[ M = 2000 \] for all multiple ellipse estimations.

400 unstructured outliers

200 points, \( \sigma_g = 3 \)
300 points, \( \sigma_g = 6 \)
400 points, \( \sigma_g = 9 \)

500 unstructured outliers

The correct inliers are recovered 96 times from 100. Four times the "top structure" didn’t recover.

\[ \text{scale} : \quad 57.7 \quad 17.4 \quad 8.9 \quad 428.4 \]
\[ \text{structure} : \quad 164 \quad 199 \quad 293 \quad 344 \]
\[ \text{strength} : \quad 28971 \quad 10134 \quad 2627 \quad 719. \]

The correct inliers are recovered 98 times from 100. Two times the "200 points ellipse" didn’t recover.

\[ \text{scale} : \quad 8.63 \pm 3.02 \quad 17.77 \pm 3.73 \quad 49.72 \pm 15.20 \]
\[ \text{inliers} : \quad 286.2 \pm 24.8 \quad 202.7 \pm 20.5 \quad 151.2 \pm 25.2 \]

\[ \text{scale} : \quad 9.16 \quad 12.77 \quad 13.42 \quad 146.17 \quad 152.70 \]
\[ \text{inliers} : \quad 181 \quad 339 \quad 353 \quad 396 \quad 117 \]
\[ \text{strength} : \quad 25486 \quad 24785 \quad 21188 \quad 2349 \quad 944. \]

The correct inliers are recovered 98 times from 100. Two times the "200 points ellipse" didn’t recover.

\[ \text{scale} : \quad 7.83 \pm 2.15 \quad 13.99 \pm 3.81 \quad 17.86 \pm 4.37 \]
\[ \text{inliers} : \quad 194.2 \pm 27.2 \quad 319.0 \pm 47.3 \quad 430.5 \pm 59.5 \]
$$(\mathbf{y}_i - \mathbf{y}_c)^\top \mathbf{Q} (\mathbf{y}_i - \mathbf{y}_c) - 1 \simeq 0 \quad i = 1, \ldots, n_{in}$$

200 points, $\sigma_y = 9$
300 points, $\sigma_y = 6$ 500 unstructured outliers
400 points, $\sigma_y = 3$

The correct inliers are recovered 95 times from 100. Five times the ”200 points ellipse” didn’t recover.

$\text{scale} : \quad 6.34 \pm 1.42 \quad 13.31 \pm 3.45 \quad 19.42 \pm 5.72$
$\text{inliers} : \quad 404.4 \pm 24.9 \quad 315.7 \pm 29.8 \quad 197.1 \pm 17.0$

**Ellipse estimation in a real image**

Image size 200 × 150.
EDISON segmentation system based on the mean shift is applied (top/right) with the default spatial $\sigma_s = 7$ and range $\sigma_r = 6.5$ bandwidth.
Canny edge detection (bottom/left) with the thresholds of 100 and 200. The strongest three ellipses are drawn superimposed over the edges.
The ellipses drawn superimposed over the original image (bottom/right) are correct.
Fundamental matrix estimation.

$M = 1000$ for all fundamental matrix estimations.

The 546 points pairs were extracted with SIFT with the default distance ratio of 0.8.

The first three structures are

- 160 pairs with $\hat{\sigma}_1 = 0.42$ (red)
- 147 pairs with $\hat{\sigma}_2 = 0.59$ (green)
- 60 pairs with $\hat{\sigma}_3 = 0.34$ (blue).

Fundamental matrix estimation.

The SIFT finds 173 point correspondences.

The first structure is 70 pairs with $\hat{\sigma} = 0.38$. 
Homography estimation

$M = 1000$ for all homography estimations.

The SIFT finds 168 point correspondences.

The first two structures are
66 pairs with $\hat{\sigma}_1 = 1.37$ (red)
34 pairs with $\hat{\sigma}_2 = 4.3$ (green).

Homography estimation

The SIFT finds 495 correspondences.

The first three structures are
160 pairs with $\hat{\sigma}_1 = 1.25$ (red)
98 pairs with $\hat{\sigma}_2 = 1.59$ (green)
121 pairs with $\hat{\sigma}_3 = 1.77$ (blue).
Conditions for robustness

For the algorithm to be robust the four parameters $M$, the number of trial; $n_{out}$, amount of outliers; $\hat{\sigma}$, the noise of the inlier structure; $\epsilon$, the initial set interact in a complex manner.

200 inliers, $\sigma_g = 9$. 400 outliers.
Change the parameters, one at a time

- $n_{out} = 100, 400, 800$  
- $\sigma_g = 3, 9, 15$  
- $M = 50 - 4000$  
- $\epsilon = 1 - 40\%$.

In each condition do 100 trials and return the average of the counted true inliers over the total points classified as inliers.

How important is the number of trials $M$.

Not much, when $M$ exceeds some value depending on the input data.
Increasing to outliers to $n_{out} = 900$ ($\sigma_g = 9$), has stronger effect than increasing the noise ($n_{out} = 400$) $\sigma_g = 15$. 
The initial set should be quasi-correct. $M = 2000$.

From the initial set 24 points are between $\pm 2\sigma_g = \pm 18$.

$\epsilon = 5\%$ is equivalent to $n_{r1} = 30$ points.

The inliers are 33% from the total number of points.


Average real initial set with automatic increase. $\sigma_g = 9$.

$\sigma_g$ changing: initial sets  
... but the final results in fact decrease for large noise.

Taking $\epsilon = 5\%$ is enough most of the time.
The algorithm is more robust when

- An input data is preprocessed and part of the outliers are eliminated. Then the probability of the output being satisfactory increases.

- The number of trials $M$, beyond a value depending on the input data, does not matter.

- On the average, the probability of success is given by how "tough" the input data is. The $\epsilon = 5\%$ generally is enough as the starting point. The output can be improved only through further preprocessing of the input data.

Open problems...

Estimating the $m \times k$ matrix $\Theta$ and a $k$-dimensional vector $\alpha$. Reducing to $k$ independently runs of the algorithm is it enough?

An image contains both lines and conics together with outliers. How do you approach it, if all inlier structures should be recovered?

You have to process robustly a very large image. Will hierarchical processing from many small images to the large image find all the relevant inlier structures?

The covariances of the input points are not equal.

Estimate first all the inlier structures with the same $\hat{\sigma}$. After that, for each structure separately, do scale estimation. Will this procedure work all the time?
Thank You!