

# Uncertainty Models in Quasiconvex Optimization for Geometric Reconstruction

Qifa Ke and Takeo Kanade

Department of Computer Science, Carnegie Mellon University

Email: ke@cmu.edu, tk@cs.cmu.edu

## Abstract

*Geometric reconstruction problems in computer vision can be solved by minimizing the maximum of reprojection errors, i.e., the  $L_\infty$ -norm. Unlike  $L_2$ -norm (sum of squared reprojection errors), the global minimum of  $L_\infty$ -norm can be efficiently achieved by quasiconvex optimization. However, the maximum of reprojection errors is the meaningful measure to minimize only when the measurement noises are independent and identically distributed at every 2D feature point and in both directions in the image. This is rarely the case in real data, where the positional noise not only varies at different features, but also has strong directionality. In this paper, we incorporate the directional uncertainty model into a quasiconvex optimization framework, in which global minimum of meaningful errors can be efficiently achieved, and accurate geometric reconstructions can be obtained from feature points that contain high directional uncertainty.*

## 1. Introduction

Given 2D image measurements, the goal of geometric reconstruction in computer vision is to estimate the 3D scene structure and/or the camera motion. Classical examples include triangulation [13], camera resectioning [3, 12], and structure from motion (cf. [6]).

When measurement noises follow independent and identical Gaussian distribution, maximum likelihood estimation can be achieved by minimizing  $F_s$ , the sum of squared reprojection errors. Minimizing  $F_s$ , however, is in general a hard non-convex problem due to the perspective effect in pin-hole camera model. An alternative is to minimize  $F_\infty$ , the maximum reprojection error [5]. In contrast to  $F_s$ , the global minimum of  $F_\infty$  can be achieved by quasiconvex optimization in an efficient way by Second Order Cone Programming [8, 9], or more efficiently by Linear Programming [9]. Previously, global optimal estimates can be achieved only in rare instances, such as *two-view* triangulation using  $L_2$ -norm [13] or  $L_\infty$ -norm [11], *three-view*

triangulation by solving six-degree polynomial equation set, and affine reconstruction using matrix factorization [17]. Minimizing  $F_\infty$  by quasiconvex optimization provides a framework to achieve global optimal estimation in many multi-view geometric reconstruction problems, without the limitation on the number of views and/or the use of affine camera model. While  $F_\infty$  is sensitive to outliers, a robust extension of quasiconvex minimization was presented in [9] to deal with outliers.

However,  $F_\infty$  is a meaningful objective to minimize only when the measurement noises are isotropic and *i.i.d.* (independent and identically distributed) at every 2D feature. In real data, this is rarely the case since the quality of feature matching depends on the image intensity pattern around the feature, which often varies at different feature points and has strong directionality to it. As a result, the uncertainty at each feature often has strong directionality, and is feature-dependent.

To account for the directional uncertainty, we minimize the covariance-weighted reprojection error (the Mahalanobis distance), instead of the Euclidean distance that has been used in [5, 8, 9]. In this paper, we incorporate the directional uncertainty model into the quasiconvex optimization framework. The directional uncertainty can be characterized by the covariance matrix at each 2D feature, as has been used in matrix factorization for affine reconstruction [10, 7]. We show that the point-wise maximum of covariance-weighted reprojection errors is a quasiconvex function, and therefore its global minimum can be obtained by a simple and efficient quasiconvex minimization algorithm that consists of a small number of convex programs (linear programs or second order cone programs).

In each convex program, each feature contributes a convex cone constraint on the unknowns. In our formulation, the shape and size of such convex cone adapts to the directional uncertainty of its underlying feature. A feature with smaller uncertainty in one or two directions will result in a cone with smaller size in corresponding direction(s), thus a stringer constraint on the unknowns. This is in contrast to [5, 8, 9] where all convex cones have common shape/size

and constrain equally on the unknowns. The directional uncertainty also allows us to model line features. As a result, point and line features can be used simultaneously for geometric reconstruction in a common quasiconvex optimization framework.

### 1.1. Geometric reconstruction

Let  $P_i, i = 1, \dots, m$ , denote a set of  $3 \times 4$  camera projection matrices, and  $Z_j, j = 1, \dots, n$ , a set of 3D points. Given  $x_{ij}$ , the images of 3D points  $Z_j$  at cameras  $P_i$ , the goal of geometric reconstruction is to recover  $P_i$  and/or  $Z_j$ . Here coordinates are in homogeneous format. In the following we briefly introduce some classical geometric reconstruction problems (see [6] for more details) that can be solved by quasiconvex optimization [8, 9]:

- **Planar homography estimation.** When the 3D points  $\{Z_j\}$  are on a plane, they can be represented by homogeneous plane coordinates.  $\{Z_j\}$  and their image features  $\{x_j\}$  are related by  $x_j = HZ_j$ , where  $H$  is the planar homography represented by a  $3 \times 3$  matrix. The task here is to estimate  $H$  from  $\{x_j\}$  and  $\{Z_j\}$ . The same problem appears when we are given two images of 3D co-planar points, where the homography relates the points in one image to the points in the other image.
- **Camera resectioning.** We are given 3D points  $\{Z_j\}$  and their images  $\{x_j\}$  in one camera  $P$ . The task is to estimate the camera matrix  $P$ .
- **Multi-view triangulation.** We are given the images  $\{x_{ij}\}$  of 3D points  $\{Z_j\}$  at multiple cameras  $\{P_i\}$ . The task is to estimate  $\{Z_j\}$  from  $\{x_{ij}\}$  and  $\{P_i\}$ .
- **Cameras with known rotations.** This case is similar to multi-view triangulation, except that the location of each camera is unknown and needs to be estimated in addition to  $\{Z_j\}$ . This is an interesting problem as modern inertial sensors can provide accurate camera poses but noisy camera positions. There also exists reconstruction methods where the camera rotations are estimated in a first step [18].
- **Using a reference plane.** When a reference plane is visible in all images, the inter-image planar homography from the given reference plane can be used to compensate the relative camera rotations. Then the problem essentially reduces to the above case of reconstruction with known rotations.

Under pin-hole camera model, the reprojection of  $x_{ij}$  is

$$k_{ij}\hat{x}_{ij} = P_i Z_j \quad (1)$$

For the above problems, Eq. (1) is linear in  $k_{ij}$  and the unknown parameters. Linear least squares technique can then

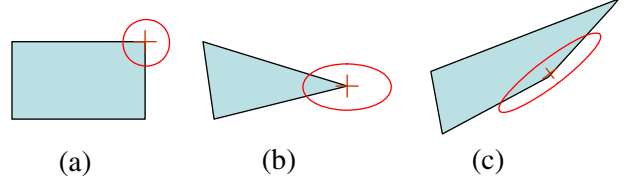


Figure 1. *Uncertainty in feature point locations.* (a): *Scalar uncertainty with covariance matrix  $Q = \text{diag}(\sigma, \sigma)$ ;* (b): *Directional but uncorrelated noises, with covariance matrix  $Q = \text{diag}(\sigma_1, \sigma_2)$ ;* (c): *Directional and correlated noises, with covariance matrix  $Q$  a full  $2 \times 2$  matrix.*

be applied to solve for the unknowns. Coordinate normalization technique may be used to improve the linear estimation. However, the linear algorithm minimizes the algebraic distance  $\|k_{ij}\hat{x}_{ij} - P_i Z_j\|$ , which is not geometrically or statistically meaningful. As a result, its performance is not reliable.

## 2. Quasiconvex optimization with uncertainty

There are three major components in geometric reconstruction: 1) a measure of model fitting error at each 2D feature; 2) a total error that combines the model fitting errors of all features; and 3) an optimization algorithm that minimizes the total error w.r.t. the unknown parameters.

In this section, we show that directional uncertainty of 2D features can be incorporated into the above three components.

### 2.1. Uncertainty of feature position

The accuracy of feature matching depends on the intensity pattern around each feature, which often has strong directionality and is location-dependent. Such directional uncertainty can be characterized by the following inverse covariance matrix (cf. [14, 15]):

$$Q^{-1} = \frac{1}{s} \sum_{(u,v) \in w} \begin{pmatrix} I_u I_u & I_u I_v \\ I_u I_v & I_v I_v \end{pmatrix}, \quad (2)$$

where  $w$  is a small window centered around the feature point in the image  $I$  of the  $i$ -th camera,  $s$  is determined by the intensity pattern inside  $w$ , and  $I_u$  and  $I_v$  are image gradients along  $u$  and  $v$  direction, respectively. A more accurate method to estimate feature position uncertainty is presented in [15], which takes into account not only the image pattern but also the image pixel noises.

Figure 1 shows the three different types feature uncertainty:

- $Q = \text{diag}(\sigma, \sigma)$ : scalar uncertainty that is feature-dependent, but is isotropic and therefore uncorrelated in  $u$  and  $v$  direction;

- $Q = \text{diag}(\sigma_1, \sigma_2)$ : directional ( $\sigma_1 \neq \sigma_2$ ) but uncorrelated in  $u$  and  $v$  direction;
- $Q = \text{full } 2 \times 2 \text{ matrix}$ : directional and correlated in  $u$  and  $v$  direction.

## 2.2. Covariance-weighted reprojection error

For the geometric reconstruction problems listed in Section 1.1, the reprojection of  $\mathbf{x} = (u, v)^\top$  in the image (in Euclidean coordinates), under pin-hole camera model, can be written as:

$$\hat{\mathbf{x}} = \left( \frac{\mathbf{a}^\top \mathbf{X}}{\mathbf{c}^\top \mathbf{X}}, \frac{\mathbf{b}^\top \mathbf{X}}{\mathbf{c}^\top \mathbf{X}} \right)^\top \quad (3)$$

Here  $\mathbf{X}$  is the vector to be estimated.  $\mathbf{a}$ ,  $\mathbf{b}$ , and  $\mathbf{c}$  are known vectors. For example, in the triangulation problem,  $\mathbf{X} = (X, Y, Z, 1)$  is the 3D point to be estimated, and  $\mathbf{a}$ ,  $\mathbf{b}$ , and  $\mathbf{c}$  are the three rows of the camera matrix  $P$ , respectively.

The uncertainty in the location of each 2D feature point  $\mathbf{x}$  can be taken into account by weighting the reprojection error appropriately. The covariance matrix  $Q$  for the 2D point  $\mathbf{x}$  is a symmetric positive semi-definite matrix, and can be decomposed into the following form by the Singular Value Decomposition:  $Q = U\Sigma U^\top$ , where  $\Sigma = \text{diag}(\sigma_1, \sigma_2)$ , and  $U$  is a  $2 \times 2$  orthonormal matrix. The inverse covariance matrix takes the form of

$$Q^{-1} = U\Sigma^{-1}U^\top \quad (4)$$

Denote  $B = \Sigma^{-1/2}U^\top$ , which is an affine transformation that transforms the input data into covariance-weighted data space where the noises at each feature become isotropic and *i.i.d.*. The transformed coordinates (in Euclidean) of  $\mathbf{x}$  and  $\hat{\mathbf{x}}$  in the image plane are :

$$\mathbf{x}' = (u', v')^\top = B(u, v)^\top \quad (5)$$

$$\hat{\mathbf{x}}' = B\hat{\mathbf{x}} = \frac{1}{\mathbf{c}^\top \mathbf{X}} B \begin{pmatrix} \mathbf{a}^\top \\ \mathbf{b}^\top \end{pmatrix} \mathbf{X} \quad (6)$$

where  $\mathbf{a}^\top$  and  $\mathbf{b}^\top$  follow the notation in Eq. (3). The covariance matrix of the noise in the covariance-weighted data space now becomes isotropic and takes the form of  $\text{diag}(1, 1)$ . Denote

$$A = \left[ B \begin{pmatrix} u \\ v \end{pmatrix} \mathbf{c}^\top - B \begin{pmatrix} \mathbf{a}^\top \\ \mathbf{b}^\top \end{pmatrix} \right]. \quad (7)$$

The covariance-weighted reprojection error function is:

$$f^w(\mathbf{X}) = \|\mathbf{x}' - \hat{\mathbf{x}}'\| = \left\| \frac{A\mathbf{X}}{\mathbf{c}^\top \mathbf{X}} \right\| \quad (8)$$

In this paper, we consider affine or Euclidean reconstruction<sup>1</sup>. The cheirality constraint [4], which states that the 3D

<sup>1</sup>But using a way similar to the method briefed in [5], our algorithm in this paper can be extended to projective reconstruction.

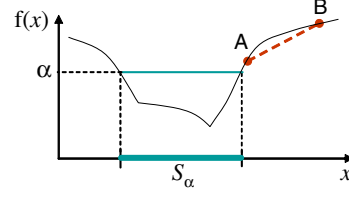


Figure 2. A quasiconvex function on  $\mathbb{R}$ , as all of its  $\alpha$ -sublevel sets  $S_\alpha$  are convex (intervals). This function is constructed as the point-wise maximum of three other functions. This quasiconvex is non-convex, as can be seen from the line segment  $\overline{AB}$  that lies below the function. In a convex function, any line segment that connects two points on the function should lie above the function. Also note that this function is non-differentiable.

points visible in the image must be in front of the camera, can then be expressed as  $\mathbf{c}^\top \mathbf{X} > 0$ . Therefore, Eq. (8) can then be written as:

$$f^w(\mathbf{X}) = \frac{1}{\mathbf{c}^\top \mathbf{X}} \|A\mathbf{X}\| \quad (9)$$

When  $\|\cdot\|$  in Eq. (9) is  $L_2$ -norm,  $f^w(\mathbf{X})$  is the Mahalanobis distance between  $\mathbf{x}$  and  $\hat{\mathbf{x}}$ .

## 2.3. Covariance-weighted total error function

The total error function in the covariance-weighted data space can now be defined using  $L_\infty$ -norm:

$$F_\infty^w(\mathbf{X}) = \max_i f_i^w(\mathbf{X}) \quad (10)$$

## 2.4. Quasiconvex minimization with uncertainty

In this subsection, we show that the covariance-weighted total error function is a quasiconvex function, as a result its global minimum can be obtained efficiently using quasiconvex minimization.

### 2.4.1 Quasiconvexity under uncertainty model

A function with convex sublevel sets is called quasiconvex.

**Definition 1.** (see [2]) A function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is called quasiconvex if its domain  $\text{dom}(f)$  and all its sublevel sets

$$S_\alpha = \{x \in \text{dom}(f) \mid f(x) \leq \alpha\}, \quad (11)$$

for all  $\alpha \in \mathbb{R}$ , are convex.

A convex function has convex sublevel sets, and is therefore quasiconvex. But the reverse is not necessarily true, as can be seen from Fig. 2, where the dash-line  $\overline{AB}$  below the function curve indicates its non-convexity.

In the following we show that both the covariance-weighted reprojection error function  $f^w(\mathbf{X})$  and the total error function  $F_\infty^w(\mathbf{X})$  are quasiconvex.

**Algorithm:**  $\min_{\mathbf{X}} F_{\infty}^w(\mathbf{X})$ .

- 1: Given  $F^* \in [l, h]$ , and the tolerance  $\varepsilon > 0$ .
- 2: **while**  $(h - l) > \varepsilon$  **do**
- 3:    $\alpha := (h + l)/2$ .
- 4:   Construct covariance-weighted convex cones  
 $S_i^w = \{\mathbf{X} \mid \|\mathbf{A}_i \mathbf{X}\| \leq \alpha \mathbf{c}_i^T \mathbf{X}\}$ .
- 5:   Solve convex feasibility problem (15).
- 6:   **if** (15) is feasible, **then**  $h := \alpha$ ;
- 7:   **else**  $l := \alpha$ .
- 8: **end while**

Figure 3. Bisection algorithm to pin down the optimal value by searching in the one-dimensional range domain of  $F_{\infty}^w(\mathbf{X})$ .

**Result 1.**  $f^w(\mathbf{X})$ , the covariance-weighted reprojection error function defined in Eq. (9), is a quasiconvex function.

*Proof.* The  $\alpha$ -sublevel set of  $f^w(\mathbf{X})$  is:

$$\begin{aligned} S^w &= \{\mathbf{X} \mid f^w(\mathbf{X}) \leq \alpha\} \\ &= \{\mathbf{X} \mid \|\mathbf{A} \mathbf{X}\| \leq \alpha \mathbf{c}^T \mathbf{X}\} \end{aligned} \quad (12)$$

where matrix  $\mathbf{A}$  is defined in Eq. (7).

The inequality  $\|\mathbf{A} \mathbf{X}\| \leq \alpha \mathbf{c}^T \mathbf{X}$  defines a convex cone [2]. As a result,  $S^w$  is a convex set.  $\square$

**Result 2.**  $F_{\infty}^w(\mathbf{X})$ , the point-wise maximum of covariance-weighted reprojection error  $f_i^w(\mathbf{X})$ , is a quasiconvex function.

*Proof.* The  $\alpha$ -sublevel set of  $F_{\infty}^w(\mathbf{X}) = \max_i f_i^w(\mathbf{X})$  is:

$$\begin{aligned} S_{\alpha}^{\infty} &= \{\mathbf{X} \mid \max_i f_i^w(\mathbf{X}) \leq \alpha\} \\ &= \{\mathbf{X} \mid f_i^w(\mathbf{X}) \leq \alpha\} \\ &= \bigcap_{i=1}^m S_i^w \end{aligned} \quad (13)$$

Here  $S_i^w = \{\mathbf{X} \mid \|\mathbf{A}_i \mathbf{X}\| \leq \alpha \mathbf{c}_i^T \mathbf{X}\}$  is the  $\alpha$ -sublevel set of the covariance-weighted reprojection error function  $f_i^w(\mathbf{X})$ . From Result 1, we know that  $\{S_i^w\}$  are all convex sets. As a result, their intersection  $S_{\alpha}^{\infty}$  is also a convex set. Therefore,  $F_{\infty}^w(\mathbf{X})$  is a quasiconvex function.  $\square$

## 2.4.2 Minimization algorithm

A quasiconvex function may be non-differentiable, and gradient-decent-like approaches are not applicable for its minimization. See Fig. 2 again for an example. However, the fact that its sublevel sets are convex makes it possible to obtain its global minimum using convex programming.

Denote  $F^*$  the global minimum of  $F_{\infty}^w(\mathbf{X})$ . For a given value  $\alpha$ , if the sublevel set  $S_{\alpha}^{\infty} = \{\mathbf{X} \mid F_{\infty}^w(\mathbf{X}) \leq \alpha\}$  is non-empty, then we must have  $F^* \leq F_{\infty}^w(\mathbf{X}) \leq \alpha$ . On

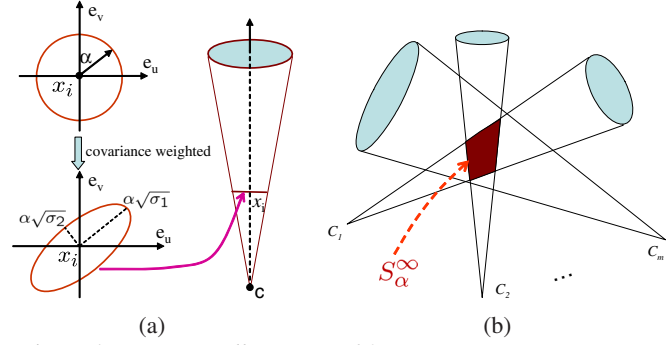


Figure 4. Geometric illustration of 3D reconstruction using quasiconvex optimization. (a): The shape and size of each cone are determined by  $\alpha$  and covariance matrix  $\mathbf{Q}_i$ ; (b): The algorithm seeks minimum  $\alpha$  such that the cones have non-empty common intersection  $S_{\alpha}^{\infty}$ . Note that cones have different shapes and sizes. The last cone  $C_m$  has large directional uncertainty and has little constraint on determining  $S_{\alpha}^{\infty}$  in the direction that has large uncertainty.

the other hand, if  $S_{\alpha}^{\infty}$  is empty, then we can conclude that  $F^* > \alpha$ . Such simple relationship between the global minimum  $F^*$  and the  $\alpha$ -sublevel set  $S_{\alpha}^{\infty}$  enables us to apply the bisection algorithm, as shown in Fig. 3, to pin down the global minimum of  $F_{\infty}^w(\mathbf{X})$ . Each step of the bisection algorithm checks whether  $S_{\alpha}^{\infty}$  is empty or not by solving the following feasibility problem:

$$\begin{aligned} \text{find } \mathbf{X} \\ \text{s.t. } \mathbf{X} \in S_{\alpha}^{\infty} \end{aligned} \quad (14)$$

Given that  $S_{\alpha}^{\infty}$  is a convex set as a result of the quasiconvexity of  $F_{\infty}^w(\mathbf{X})$ , problem (14) is a convex feasibility problem:

$$\begin{aligned} \text{find } \mathbf{X} \\ \text{s.t. } \|\mathbf{A}_i \mathbf{X}\| \leq \alpha \mathbf{c}_i^T \mathbf{X}, \\ \mathbf{c}_i^T \mathbf{X} > 0 \quad i = 1, \dots, m \end{aligned} \quad (15)$$

Specifically, (15) is a Linear Program (LP) when  $\|\cdot\|$  is a  $L_1$  norm, and a Second Order Cone Program (SOCP) when  $\|\cdot\|$  is  $L_2$  norm. Both LP and SOCP are well studied and existing efficient algorithms and implementations are ready to use (e.g., the Matlab built-in function *linprog*, or SeDuMi [16]).

The bisection algorithm converges quickly in  $\lceil \log_2((h - l)/\varepsilon) \rceil$  iterations, and obtains the global minimum within the predefined accuracy of  $\varepsilon$ .

## 2.4.3 Geometric interpretation

The covariance-weighted quasiconvex minimization algorithm has an intuitive geometric interpretation. We use triangulation as an example to illustrate.



For a feature point  $\mathbf{x}_i$ , the  $\alpha$ -sublevel set  $S_i^w$  of the covariance-weighted reprojection error function  $f_i^w(\mathbf{X})$  is a second order convex cone in the 3D space in front of the camera. The shape and size of such convex cone are determined by  $\alpha$  and the covariance matrix  $\mathbf{Q}_i = \mathbf{U} \text{diag}(\sigma_1, \sigma_2) \mathbf{U}^\top$ , as shown in Fig. 4(a). The construction of convex cone  $S_i^w$  is the following. First a circle in the image plane with radius  $\alpha$  is scaled by  $\sqrt{\sigma_1}$  and  $\sqrt{\sigma_2}$  in  $u$  and  $v$  direction, respectively. This results in an ellipse with axes of  $\alpha\sqrt{\sigma_1}$  and  $\alpha\sqrt{\sigma_2}$  respectively. The ellipse is then rotated by the rotation matrix  $\mathbf{U}$ . The final convex cone, dubbed as *elliptical cone*, is formed by connecting the camera optical center and the rotated ellipse.

The  $\alpha$ -sublevel set  $S_\alpha^\infty$  of the cost function  $F_\infty^w(\mathbf{X})$  is the intersection of all  $\alpha$ -sublevel sets  $\{S_i^w\}$ . The feasibility problem (15) is therefore equivalent to determining if there exists common intersection of the convex elliptical cones  $\{S_i^w\}$ , as shown in Fig. 4(b). The bisection algorithm then seeks a minimum value  $\alpha$  such that the common intersection of the convex elliptical cones  $\{S_i^w\}$  is non-empty.

Since  $\alpha$  is common for all convex cones, the relative shape and size of each individual cone is actually determined by its associated covariance matrix. Therefore the effect of each convex cone constraint on the final estimation  $\mathbf{X}$  is weighted by the inverse covariance matrix. Consider an extreme case when uncertainty goes to infinity. In such case, the corresponding convex cone is scaled to infinite size and does not have any constraint on the estimate of  $\mathbf{X}$ . On the other hand, if the uncertainty is zero, the convex cone becomes a ray, and we must constraint  $\mathbf{X}$  on the ray, which is a strong constraint. When directional uncertainty is presented, the effect of constraints from different directions are determined by  $\sigma_1$  and  $\sigma_2$ , respectively.

### 3. Experiments

We use planar homography estimation and multi-view triangulation as two applications of our algorithm, and use synthetic and real data to evaluate its performance. We compare the performances of three algorithms: the quasiconvex minimization of  $F_\infty^w$  with uncertainty model, the quasiconvex minimization of  $F_\infty$  without utilizing feature uncertainty, and the normalized linear algorithm.

#### 3.1. Synthetic data

##### 3.1.1 Planar homography estimation

Fig. 5(a) shows the setup to generate the synthetic data for homography estimation, where the camera images the points on the "ground plane". This simulates the case where a camera mounted on a vehicle is looking at the ground plane at some angle. Note that in this case  $h_{33}$  in homography  $\mathbf{H}_T$  may become very small. In all the algorithms being compared, we do not assume  $h_{33} = 1$ .

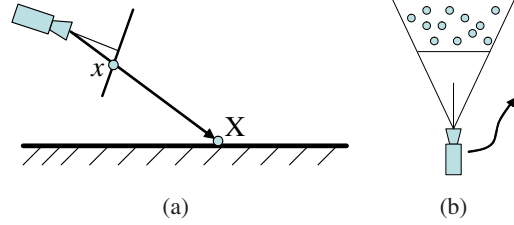


Figure 5. The setup for synthetic data generation. In both cases, the image coordinates are normalized such that the camera focal length  $f = 1$ . (a): Camera looking at points on the ground plane; (b): Camera rotating and translating, and looking at 3D points within a depth range.

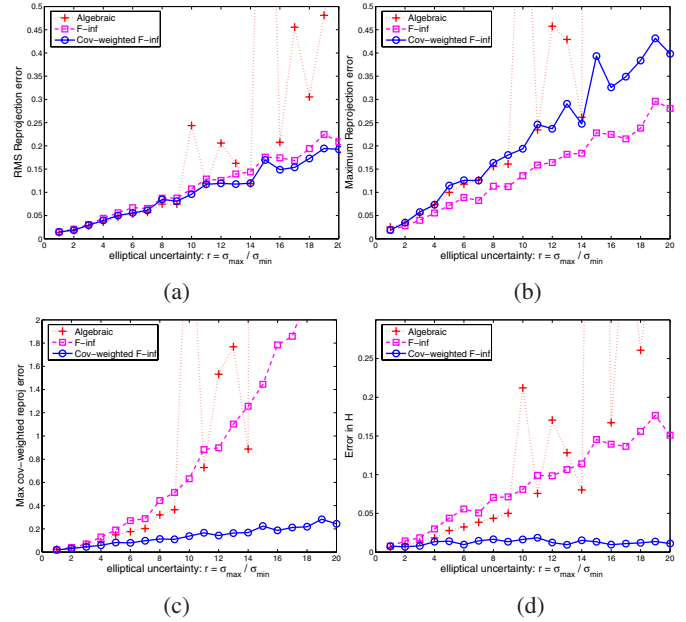


Figure 6. Results from planar homography estimation. (a): RMS error; (b): Maximum of reprojection error; (c): Maximum of covariance-weighted reprojection error; (d): Error in  $\mathbf{H}$ .

We randomly generate twenty 3D points  $\mathbf{X}_i$  on the ground plane, and compute the 2D images  $\mathbf{x}_i$  of these 3D points. We then add elliptical Gaussian noise to  $\mathbf{x}_i$ . The noise perturbed points are denoted by  $\tilde{\mathbf{x}}_i$ . The ellipticity of the noise is measured by  $r = \sqrt{\sigma_{\max} / \sigma_{\min}}$ , where  $\sigma_{\max}$  and  $\sigma_{\min}$  are the major and minor axes of the uncertainty ellipse, respectively. The orientation of the ellipse is randomly selected for each point.

We compare the performances using four criteria:

- Maximum reprojection error

$$F_\infty(\mathbf{X}) = \max_i d(\tilde{\mathbf{x}}_i, \mathbf{H}\mathbf{X}_i) \quad (16)$$

where  $d(\cdot, \cdot)$  denotes the Euclidean distance;

- Root of Mean Squares (RMS) of reprojection errors;

- Maximum covariance-weighted reprojection error  $F_{\infty}^w(\mathbf{X})$  as defined in Eq. (10);
- Error in H defined as:

$$e_H = \sqrt{\frac{1}{N} \sum_{i=1}^N d^2(\mathbf{x}_i, \mathbf{H}\mathbf{X}_i)} \quad (17)$$

where  $d(\mathbf{x}_i, \mathbf{H}\mathbf{X}_i)$  is the Euclidean distance between the ground truth 2D point  $\mathbf{x}_i$  and the reprojection point  $\mathbf{H}\mathbf{X}_i$ . This error metric compares the estimated H to the ground truth  $\mathbf{H}_T$  using ground truth points. If there is not error in the estimated H, then  $e_H = 0$ .

We apply our algorithm to estimate H from this synthetic data. We repeat the experiments for 20 times, and report the above error measures using mean of these 20 runs. Fig. 6 shows the results, where  $\sqrt{\sigma_{\min}} = 0.01$  and  $\sqrt{\sigma_{\max}}$  varies from 0.01 to 0.2, i.e., the ellipticity  $r$  varies from 1 (isotropic) to 20.

As we can see from (a), both  $F_{\infty}$  and  $F_{\infty}^w$  have similar RMS error. Normalized linear algorithm has similar RMS error when  $r$  is small, but becomes unreliable when  $r \geq 10$ .

Fig. 6(b) shows that minimizing  $F_{\infty}$  gives lowest maximum reprojection error, while Fig. 6(c) shows that minimizing  $F_{\infty}^w$  gives lowest maximum covariance-weighted reprojection error. This indicates that quasiconvex minimization indeed achieves the global minimum of  $F_{\infty}$  and  $F_{\infty}^w$ , respectively.

From Fig. 6(d), which compares the estimated H against ground truth  $\mathbf{H}_T$  using the metric  $e_H$ , we can see that  $F_{\infty}^w$  performs the best, and its performance does not degrade at all with the increase of  $r$ . This indicates that  $F_{\infty}^w$  is the proper metric to minimize. We also find that normalized linear algorithm performs better than  $F_{\infty}$  when  $r$  is small, but when  $r$  is large, its performance becomes unreliable.

### 3.2. Multiview triangulation

Fig. 5(b) shows the set up to generate the synthetic data for multi-view triangulation. The camera is rotating and translating, and takes ten consecutive views of twenty 3D points located at different depth.

Fig. 7 shows the average results of 20 runs. Again it shows that the covariance-weighted reprojection error is the right metric to use, as can be seen by the fact that  $F_{\infty}^w$  gives the best 3D estimation  $\mathbf{Z}$  when compared to the ground truth  $\mathbf{Z}_T$  using the following metric:

$$e_{3D} = \frac{\|\mathbf{Z} - \mathbf{Z}_T\|_2}{\|\mathbf{Z}_T\|_2} \quad (18)$$

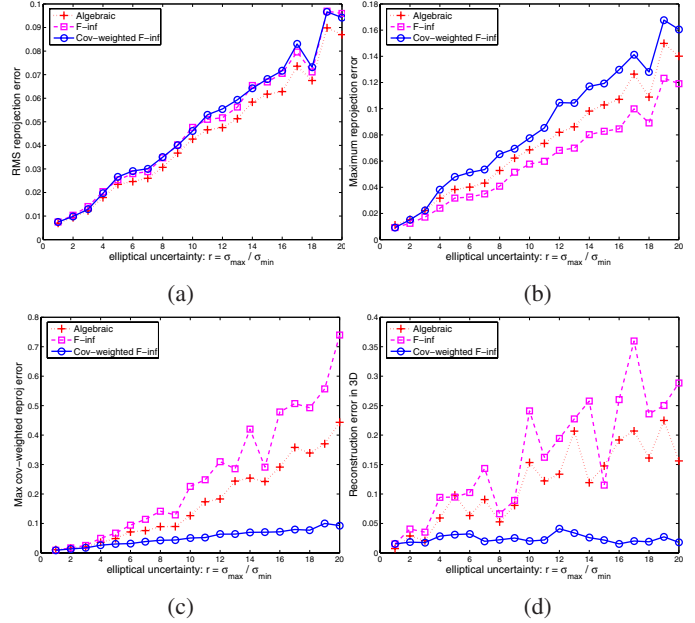


Figure 7. Results from multi-view triangulation. (a): RMS error; (b): Maximum of reprojection error; (c): Maximum of covariance-weighted reprojection error; (d): Reconstruction error compared against ground truth 3D.

### 3.3. Infinite elliptical uncertainty

The performance of  $F_{\infty}^w$  does not degrade even when the ellipticity of noises  $r$  goes to essentially infinity, as can be seen from Table 1. This fact indicates that the normal optical flow can be modelled by directional (infinity) uncertainty. As a result, the point feature and line feature can be simultaneously used in the quasiconvex optimization for many geometric reconstruction problems.

	Homography	Triangulation
Algebraic	833.8092	1.4380
$F_{\infty}$	710.3891	3.1084
$F_{\infty}^w$	0.0093	0.0146

Table 1. Results under infinite elliptical uncertainty  $r = 10^5$ . The table shows  $e_H$  for homography estimation and  $e_{3D}$  for triangulation.

### 3.4. Real data

We apply our algorithm to estimate the inter-image homography using real image data. Fig. 8(a) and (b) shows two input images. Here (a) is the first image. The other images are obtained by applying a known and gradually-changed planar homography to the first image. This way we have the ground truth of the planar homography for evaluation purpose.

The features in the first image are tracked through the sequence, and the inverse covariance matrix for each feature

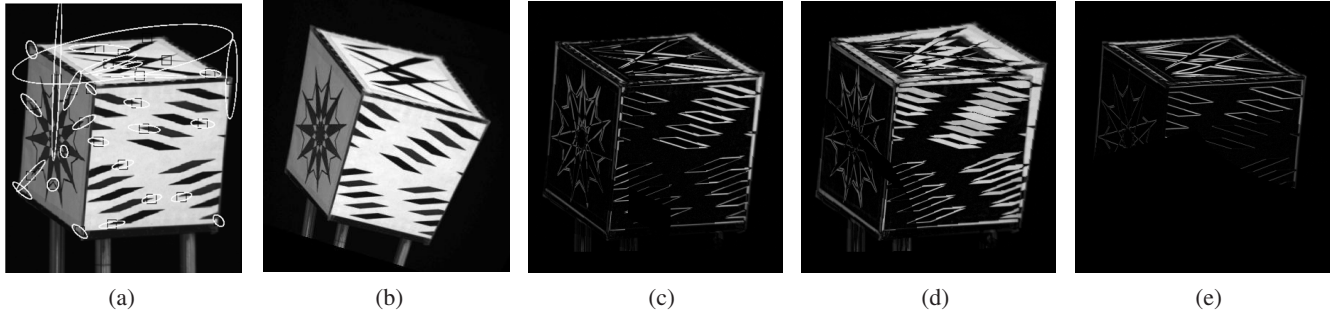


Figure 8. Planar homography estimation. (a) is the first image, where elliptical uncertainties for some feature points are shown. (b) is the last image. (c)–(d) are the residual images by applying the estimated homographies estimated by different methods (note that residuals are scaled up for visibility): (c): Normalized linear algorithm, average pixel residual  $e = 15.0798$ ; (d): Minimizing  $F_\infty$ ,  $e = 28.2082$ ; (e): Minimizing  $F_\infty^w$ ,  $e = 5.8896$ .

is computed by the Hessian matrix (Eq. (2)). The elliptical uncertainties for some feature points are shown in Fig. 8(a).

Fig. 8(c) to (e) show the image residual by applying the inverse-warping using the estimated homographies. As we can see, minimizing  $F_\infty^w$  gives the homography that has the lowest intensity residual. It correctly down weight the features with large directional uncertainty on the top of the box to produce a correct estimate of  $H$ . On the other hand, both normalized linear algorithm and the minimization of  $F_\infty$  give worse results, as can be seen by the large residuals on the top of the box, where there exist features with large directional uncertainty. Normalized linear algorithm performs better than minimizing  $F_\infty$  in this case.

## 4. Conclusion

In this paper we incorporate the directional uncertainty in 2D feature into the quasiconvex optimization framework. We show that the global minimum of  $F_\infty^w$ , the maximum of covariance-weighted reprojection errors, is a quasiconvex function and can be globally minimized. We have shown good performance of our algorithm using both synthetic and real data.

The directional uncertainty can be used to model line feature. Point and line features can therefore be unified in the quasiconvex optimization framework that can be applied to many geometric reconstruction problems.

The uncertainty model also allows us to extend our work to deal with outliers, by adding a scalar weight  $w$  to the reprojection error at each iteration step, where  $w$  is determined by the reprojection error using some robust function [1].

## References

- [1] M. Black and A. Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *International Journal of Computer Vision*, 19(1):57–92, 1996.
- [2] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [3] R. Haralick, C. Lee, K. Ottenberg, and M. Nolle. Review and analysis of solutions of the 3-point perspective pose estimation problem. *IJCV*, 13(3):331–356, 1994.
- [4] R. I. Hartley. Chirality. *International Journal of Computer Vision*, 26(1):41–61, 1998.
- [5] R. I. Hartley and F. Schaffalitzky.  $L_\infty$  minimization in geometric reconstruction problems. In *CVPR 2004*.
- [6] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [7] M. Irani and P. Anandan. Factorization with uncertainty. In *ECCV (1)*, pages 539–553, 2000.
- [8] F. Kahl. Multiple view geometry and the  $l_\infty$ -norm. In *ICCV 2005*.
- [9] Q. Ke and T. Kanade. Quasiconvex optimization for robust geometric reconstruction. In *ICCV 2005*.
- [10] D. D. Morris and T. Kanade. A unified factorization algorithm for points, line segments and planes with uncertainty models. In *ICCV*, pages 696–702, 1998.
- [11] D. Nistér. *Automatic dense reconstruction from uncalibrated video sequences*. Phd thesis, Royal Institute of Technology KTH, Stockholm, Sweden, 2001.
- [12] L. Quan and Z. Lan. Linear  $n \geq 4$ -point pose determination. In *ICCV 1998*.
- [13] R. Hartley and P. Sturm. Triangulation. *Computer Vision and Image Understanding*, 68(2):146–157, 1997.
- [14] J. Shi and C. Tomasi. Good features to track. In *CVPR'94*.
- [15] R. M. Steele and C. Jaynes. Feature uncertainty arising from covariant image noise. In *CVPR*, 2005.
- [16] J. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12:625–653, 1999.
- [17] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: A factorization method. *IJCV*, 9(2), 1992.
- [18] M. Uyttendaele, A. Criminisi, S. Kang, S. Winder, R. Hartley, and R. Szeliski. High-quality image-based interactive exploration of real-world environments. In *IEEE Computer Graphics and Applications*, 2004.