Error analysis and condition estimation of the pyramidal form of the Lucas-Kanande method in optical flow

Joab Winkler (✉ j.r.winkler@sheffield.ac.uk)
University of Sheffield

Research Article

Keywords: Optical flow, Lucas-Kanade, condition estimation, Gaussian pyramid

Posted Date: July 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1804043/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Error analysis and condition estimation of the pyramidal form of the Lucas-Kanade method in optical flow

Joab R. Winkler

Received: date / Accepted: date

Abstract Optical flow is the apparent motion of the brightness patterns in an image. The pyramidal form of the Lucas-Kanade (LK) method is frequently used for its computation but experiments have shown that the method has deficiencies. Problems arise because of numerical issues in the least squares (LS) problem \( \min \|Ax - b\|_2^2 \), \( A \in \mathbb{R}^{m \times 2} \) and \( m \gg 2 \), which must be solved many times. Numerical properties of the solution \( x_0 = A^+b = (A^TA)^{-1}A^Tb \) of the LS problem are considered and it is shown that the property \( m \gg 2 \) has implications for the error and stability of \( x_0 \). In particular, it can be assumed that \( b \) has components that lie in the column space (range) \( \mathcal{R}(A) \) of \( A \), and the space that is orthogonal to \( \mathcal{R}(A) \), from which it follows that the upper bound of the condition number of \( x_0 \) is inversely proportional to \( \cos \theta \), where \( \theta \) is the angle between \( b \) and its component that lies in \( \mathcal{R}(A) \). It is shown that the maximum values of this condition number, other condition numbers and the errors in the solutions of the LS problems increase as the pyramid is descended from the top level (coarsest image) to the base (finest image), such that the optical flow computed at the base of the pyramid may be computationally unreliable. Two examples of the computation of the optical flow demonstrate the theoretical results, and the implications of these results for extended forms of the LK method are discussed.

Keywords Optical flow · Lucas-Kanade · condition estimation · Gaussian pyramid

1 Introduction

Optical flow is the apparent motion of the brightness patterns in an image. This is different from the motion field, which is the projection onto two dimensions of motion in a three dimensional environment. Both these motions are usually termed optical flow even though they are not necessarily the same [3, §2]. There are several methods for the computation of optical flow and it is assumed in all the methods that it is locally smooth. The methods can be classified as global, for example, the Horn-Schunck method, or local, for example, the Lucas-Kanade (LK) method [4, §2.1], [8, §2], [16], [20, §2], [21]. Local methods are more robust to noise but they do not yield a dense optical flow field, and global methods are more sensitive to noise but they yield a dense optical flow field [8, §1]. The computation of the optical flow has been the focus of much research, particularly in images that are corrupted by significant noise [1], or images in which two or more objects move [5, 24], or images in which there are temporal changes in illumination [11, 15, 22, 23].

Comparisons of the different methods for the computation of optical flow are in [3,4,20]. The conclusions from these comparisons are empirical and based on quantitative measures of the errors in the computed results, but numerical and linear algebraic considerations of the methods are not included in these comparisons. These issues are addressed in this paper by detailed error and condition estimation of the pyramidal form of the LK method. This method requires that many least squares (LS) problems be solved at each level of the pyramid and it is shown that, even in the absence of occlusions, motion discontinuities and other problematic phenomena, the maximum values of the errors and condition numbers of the solutions of the LS
problems (the optical flow) increase as the pyramid is descended from the top level to the bottom level.

Kearney, Thompson and Boley [17] consider condition estimation and error analysis for the computation of the optical flow from two points with pixel coordinates $p_i = (x_i, y_i)$ and $p_j = (x_j, y_j)$ at time $t$, where $p_i$ is near $p_j$ such that a first order Taylor approximation is appropriate. This approximation yields

$$G \omega = -b, \quad (1)$$

where

$$G = \begin{bmatrix} I_x^{(i)} & I_y^{(i)} \\ I_x^{(j)} & I_y^{(j)} \end{bmatrix}, \quad \omega = \begin{bmatrix} u \\ v \end{bmatrix}, \quad b = \begin{bmatrix} f_t^{(i)} \\ f_t^{(j)} \end{bmatrix},$$

$I_x^{(i)} = \frac{\partial I_x}{\partial x}$, $I_y^{(i)} = \frac{\partial I_y}{\partial y}$, $f_t^{(i)} = \frac{\partial f_t}{\partial t}$,

$\omega$ is the optical flow and $I(x) = I^{(i)}(x, y, t)$ is the intensity of the $i$th pixel. The work in this paper follows the analysis in [17] because issues of computational linear algebra, specifically condition estimation and error analysis, are addressed. There are, however, differences because an arbitrary number of points, rather than two points, are considered in this paper since the LK method yields many LS problems, rather than the problem (1), for which the coefficient matrix is square and of order two.

This paper addresses numerical issues in the LS problem and their implications for the LK method:
1. The LS problem
   
   $$x_0 = \arg \min ||Ax - b||^2, \quad A \in \mathbb{R}^{m \times n}, \quad (2)$$
   
   where $||\cdot|| = ||\cdot||_2$, $m > n$ and rank $A = n$, is solved for each window at each level of the pyramid. It is shown that the maximum value of the error in the solution $x_0$, the optical flow, increases as the pyramid is descended from the top level to the lowest level.

2. It is shown that an upper bound of the condition number of the solution $x_0$ of (2) is $\kappa(A) / \cos \theta$, where $\kappa(A)$ is the condition number of $A$ and $\theta$ is the angle between $b$ and its component that lies in the column space (range) $\mathcal{R}(A)$ of $A$. This is an important result because it follows that even if $A$ is well conditioned, that is, $\kappa(A) = O(1)$, the solution $x_0$ may be ill conditioned if $\theta$ is large.

3. The coefficient matrix $A$ contains the spatial intensity gradients and the vector $b$ contains the temporal intensity gradients. The condition number $\kappa(A) / \cos \theta$ of $x_0$ is limited because it quantifies the effect on $x_0$ of errors in $b$, but errors in $A$ are not considered. A condition number of $x_0$ that includes errors in $A$ and $b$ is developed because it is more appropriate for the LK method. This is a non-linear condition number, in contrast to the linear condition number $\kappa(A) / \cos \theta$.

4. The dimensions of the coefficient matrix $A$ satisfy $m \gg n$ and it can therefore be assumed that a significant component of $b$ lies in the space that is orthogonal to $\mathcal{R}(A)$. It is shown that the error in $x_0$ is therefore large and that there is a simple relationship between this error and $\cos \theta$. It is concluded that the angle $\theta$ arises in the error and condition numbers of $x_0$, and it is therefore an important measure to consider when assessing the computational reliability of the optical flow.

5. These issues of condition estimation and error analysis are considered for each level of the pyramid. It is shown that the maximum values of the condition numbers and errors in the solutions of the LS problems (one LS problem for each window at each level of the pyramid) increase as the pyramid is descended from the coarsest image to the finest image. This shows that the advantage of the pyramidal implementation of the LK method - the ability to cope with large displacements between successive images - must be balanced against the numerical measures that quantify the robustness and reliability of the solutions of the many LS problems that are solved in the LK method.

The method of least median of squares (LMedS) is used in [1,18,24] to simultaneously solve the problems of temporal variation of illumination and motion discontinuities induced by the relative motion of two or more objects. The LMedS method requires the solution of the LS problem (2) and the pyramidal form of the LK method is also used in feature tracking [6,26]. More generally, the theory and results in this paper are applicable to all problems in which the LK method is used. There are several different forms of the LK method and they differ in the expressions used for the calculation of the derivatives of the intensity. For example, one form of the LK method requires that $A$ and $b$ in (2) are premultiplied by a diagonal matrix, such that a larger weight is assigned to pixels in the centre of each window, and a smaller weight is assigned to pixels at the borders of the window [19, §4.3.2].

The LK method is described in Section 2 and the error in the solution $x_0$ of the LS problem is considered in Section 3. It is shown that there is a simple relationship between $\cos \theta$ and the error, and an expression for $\cos \theta$ is developed in Section 4. Linear and non-linear condition numbers that consider errors in $b$, and errors in $A$ and $b$, respectively, of the optical flow are considered in Section 5. Examples that show the errors and condition numbers of the optical flow at each level of the
pyramid are in Section 6. The paper is summarised in Section 7 and it is shown that the numerical problems with the solution of the LS problem that are addressed in this paper provide the mathematical motivation for the generalised form of the LK method that includes warping functions [2].

2 The Lucas-Kanade method

The LK method is derived from the assumption of constant intensity $I(x, y, t)$ as a pixel at position $(x, y)$ at time $t$ moves to position $(x + \delta x, y + \delta y)$ at time $t + \delta t$, such that

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t),$$

to first order. It therefore follows that

$$\partial I \delta x + \partial y \partial y + \partial I \partial t = 0,$$

and thus if

$$u = \frac{\partial x}{\partial t} \quad \text{and} \quad v = \frac{\partial y}{\partial t},$$

then

$$I_x u + I_y v + I_t = 0,$$

(4)

where

$$I_x = \frac{\partial I}{\partial x}, \quad I_y = \frac{\partial I}{\partial y} \quad \text{and} \quad I_t = \frac{\partial I}{\partial t}. $$

Equation (4) is one equation in two unknowns (the components $u$ and $v$ of the optical flow), and it therefore possesses an infinite number of solutions. This problem of a non-unique solution is addressed by considering a square window in the images at times $t$ and $t + \delta t$, and if each window contains $p$ points, then the application of (4) to each of these points yields,

$$I_x^{(i)} u + I_y^{(i)} v + I_t^{(i)} = 0, \quad i = 1, \ldots, p,$$

which can be combined into one equation,

$$Ax = b, \quad A \in \mathbb{R}^{p \times 2}, \quad x \in \mathbb{R}^2 \quad \text{and} \quad b \in \mathbb{R}^p,$$

(5)

where the first and second columns of $A$ contain, respectively, the intensity gradients with respect to $x$ and $y$, and $b$ contains the negative of the temporal intensity gradients. Equation (5) is solved in the LS sense (2), which yields the solution,

$$x_0 = A^T b, \quad A^T = (A^T A)^{-1} A^T,$$

(6)

where $A^T$ is the pseudo-inverse of $A$ and $x_0 = [u \ v]^T$ is the optical flow. It follows from (5) that the motion in each window is a translation only, where $u$ and $v$ represent its horizontal and vertical components, respectively.

The solution $x_0$ assumes that there are errors in $b$ only, and that $A$ does not contain errors. The qualifications that follow from this assumption must be addressed because $A$ and $b$ contain, respectively, the spatial and temporal derivatives of the intensity $I(x, y, t)$. It follows that if there are errors in the intensity, then errors in its spatial and temporal derivatives must be considered in the analysis, and thus conclusions to be drawn from the solution (6) are limited. The inclusion of errors in $A$ and $b$ requires that the total least squares problem be considered [12, §6.3], and the effect of perturbations in $A$ and $b$ on the optical flow $x_0$ is considered by the development of a non-linear condition number of $x_0$ in Section 5.2.

The computation of the translation $x_0$ between two images of a dynamic scene occurs in several applications, including remote sensing and medical imaging, and it has been shown that noise and the first order approximation (3) introduce bias in the estimate of $x_0$ [7,17,25,27]. Robinson and Milanfar [25] use prior knowledge of the image and the translation to design gradient-estimation filters but the bias due to noise is not considered, and thus the LK method yields poor results for images whose signal-to-noise ratio is less than about 25 dB. Pham et al. [27] reduce the bias by iteratively computing the translation and upsampling the matrices that store the components $u$ and $v$ of the optical flow, which is the pyramid procedure proposed by Lucas and Kanade [21], and refined by Baker and Matthews [2]. Pham et al. [27] report that a pyramid with three levels yields an optical flow field with a very small bias.

The first order approximation (3) requires that

$$|\delta x|, |\delta y|, |\delta t| \ll 1,$$

and if this condition is not satisfied, a pyramid is formed in which the images at the $i$th level are formed by passing the images at the $(i-1)$th level through a low pass filter and then downsampling the filtered images. The coefficients of the low pass filter are [9,10],

$$(\frac{1}{16} \ 1 \ 4 \ 6 \ 4 \ 1),$$

(7)

which is approximately equal to a Gaussian filter with standard deviation $\sigma = 1$. The pyramidal implementation is shown in Figure 1 for two images $F$ and $G$ that are taken at times $t$ and $t + \delta t$ respectively. The optical flow is computed at the top of the pyramid, level 4 (the coarsest image) in Figure 1, which is the initial condition for the computation of the optical flow at level 3 of the pyramid. This operation yields two matrices, one
for each component of the optical flow, and each matrix is upsampled to yield matrices of twice the number of rows and columns and then passed through a modified form of the filter (7),

\[
\begin{pmatrix}
\frac{1}{8} & 4 & 4 \\
4 & \frac{1}{8} & 6 \\
4 & 6 & 1
\end{pmatrix}
\begin{pmatrix}
1 & 4 & 6 & 4 & 1
\end{pmatrix}.
\] (8)

This process is continued to the base of the pyramid, level 1, and the optical flow at this level is the desired optical flow. The difference in the scale factors for the filters for downsampling the images and upsampling the optical flow field arises because all the pixels in the images that are downscaled are included in the computation, but rows and columns of z eros are added to the matrices that store the components \(u\) and \(v\) before they are upscaled, such that a matrix of order \(M \times N\) is increased to a matrix of order \(2M \times 2N\). The computation of all the pixel values in these larger matrices cannot include the added rows and columns of zeros and thus only one quarter of the pixels in the matrices of order \(2M \times 2N\) are included in the computation.

The matrices of the components of the optical flow must be upsampled by the filter (8) and then multiplied by two in order to preserve the values of its components between successive levels of the pyramid. In particular, two adjacent pixels have unit separation before upsampling, but their separation increases to two after upsampling. The preservation of the values of the components of the optical flow in the upsampled images therefore requires that they be multiplied by two.

The discussion above shows that the pyramidal form of the LK method requires that the LS problem be solved for the optical flow many times at each level of the pyramid, and thus the error in its solution must be considered. This issue is addressed in Section 3 and it is shown in Section 4 that this error is closely related to the value of the angle \(\theta\) between \(b\) and its component that lies in the column space \(\mathcal{R}(A)\) of \(A\). The computed optical flow must be numerically stable and it is therefore necessary to consider its numerical condition. This issue is addressed in Section 5, and linear and non-linear condition numbers are considered.

### 3 The error in the solution of the LS problem

This section considers the error \(e\) in the solution of the LS problem (6),

\[
e = \frac{\|b - Ax_0\|}{\|b\|} = \frac{\| (I - AA^T) b \|}{\|b\|} = \frac{\| (I - \Sigma \Sigma^T) c \|}{\|c\|},
\] (9)

where the singular value decomposition of \(A\) is \(U \Sigma V^T\), \(U\) and \(V\) are orthogonal matrices of orders \(m\) and \(n\) respectively,

\[
\Sigma = \begin{bmatrix} \Sigma_1 & 0 \end{bmatrix} \in \mathbb{R}^{m \times n}, \quad \Sigma_1 = \text{diag} [\sigma_1 \sigma_2 \cdots \sigma_n],
\]

the singular values \(\sigma_i\) satisfy \(\sigma_i \geq \sigma_{i+1}, i = 1, \ldots, n-1\), and

\[
e = U^T b.
\] (10)

The matrix \(U\) is partitioned into two matrices \(U_1\) and \(U_2\),

\[
U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}, \quad U_1 \in \mathbb{R}^{m \times n}, \quad U_2 \in \mathbb{R}^{m \times (m-n)},
\]

and it follows from the orthogonal property of \(U\) that

\[
U_1^T U_1 = I_n, \quad U_2^T U_2 = I_{m-n}, \quad U_1^T U_2 = 0_{n,m-n}, \quad U_2^T U_1 = 0_{m,n-n},
\] (11)

and

\[
U_1 U_1^T + U_2 U_2^T = I_m,
\] (12)

where the columns of \(U_1\) form an orthonormal basis for \(\mathcal{R}(A)\) and the columns of \(U_2\) form an orthonormal basis for the space that is orthogonal to \(\mathcal{R}(A)\). It follows from (10) that \(e\) can be written as

\[
e = \begin{bmatrix} e_1 \\ \bar{e}_2 \end{bmatrix} = \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} b, \quad e_1 \in \mathbb{R}^n, \quad \bar{e}_2 \in \mathbb{R}^{m-n},
\] (13)

where \(e_1\) and \(\bar{e}_2\) are the components of \(b\) that lie in, respectively, the spaces spanned by the columns of \(U_1\) and \(U_2\). It follows from (9) that

\[
e^2 = \frac{\sum_{i=n+1}^{m} c_i^2}{\sum_{i=1}^{m} c_i^2} = \frac{\|\bar{e}_2\|^2}{\|e\|^2} = \frac{\|\bar{e}_2\|^2}{\|e_1\|^2 + \|\bar{e}_2\|^2},
\] (14)

and thus the error in \(x_0\) is a measure of the proportion of \(b\) that lies in the space that is orthogonal to \(\mathcal{R}(A)\). Equation (14) yields

\[
e^2 = 1 - \frac{\sum_{i=1}^{m} c_i^2}{\sum_{i=1}^{m} c_i^2} = 1 - \frac{\|e_1\|^2}{\|e\|^2} = 1 - \cos^2 \theta,
\] (15)

where

\[
\cos \theta = \frac{\|e_1\|}{\|e\|},
\] (16)

which shows that a large value of \(\theta\) leads to a large error in the solution \(x_0\) of the LS problem.

### 4 The angle between \(b\) and \(\mathcal{R}(A)\)

It is shown in this section that \(\theta\), where \(\cos \theta\) is defined in (16), is the angle between \(b\) and the column space \(\mathcal{R}(A)\) of \(A\). This geometric interpretation follows from the decomposition of \(b\) into a component \(r_1 \in \mathcal{R}(A)\) and a component \(r_2 \in \mathcal{R}(A)^\perp\),

\[
b = r_1 + r_2.
\]
Since \( r_1 \in \mathcal{R}(A) \) and \( r_2 \in \mathcal{R}(A)^{\perp} \), there exist vectors \( t_1 \) and \( t_2 \) such that
\[
r_1 = U_1 t_1 \quad \text{and} \quad r_2 = U_2 t_2,
\]
and thus
\[
b = U_1 t_1 + U_2 t_2.
\]
It follows from (11) that the premultiplication of this equation by \( U_1^T \) and then \( U_2^T \) yields
\[
b = U_1 U_1^T b + U_2 U_2^T b,
\]
where the first and second terms on the right are, respectively, the orthogonal projection of \( b \) onto \( \mathcal{R}(A) \) and \( \mathcal{R}(A)^{\perp} \). It follows from (11) that the second term is orthogonal to the first term, which confirms that \( r_1 \) is orthogonal to \( r_2 \). Also, it follows from (12) that
\[
\|U_1 U_1^T b\|^2 + \|U_2 U_2^T b\|^2 = \|b\|^2,
\]
where
\[
\|U_1 U_1^T b\| = \|U_1^T b\| = \|e_1\|
\]
and thus \( \cos \theta \) is equal to the ratio of the magnitude of the orthogonal projection of \( b \) onto \( \mathcal{R}(A) \) to the magnitude of \( b \). Equation (15) shows that there is a simple relationship between the error in \( x_0 \) and the angle between \( b \) and its component that lies in \( \mathcal{R}(A) \). The limiting cases are:

1. \( b \in \mathcal{R}(A) \): It follows that \( \cos \theta = 1 \) and \( U_2 = 0 \), and thus \( e = 0 \).
2. \( b \in \mathcal{R}(A)^{\perp} \): It follows that \( \cos \theta = 0 \) and \( U_1 = 0 \), and thus \( e = 1 \), its maximum value.

The ratio \( \cos \theta \) provides a geometric interpretation of the error in the solution of the LS problem, and it is shown in Section 5 that \( \cos \theta \) also arises in expressions for the linear and non-linear condition numbers of \( x_0 \).

5 The numerical condition of the optical flow

This section considers the numerical condition of the optical flow \( x_0 \), the solution of the LS problem. The condition number of \( A \), \( \kappa(A) = \sigma_1/\sigma_2 \), where \( \sigma_1 > \sigma_2 \), is frequently used as a measure of the stability of \( x_0 \), but it requires qualification:

1. The condition number \( \kappa(A) \) is a function of \( A \) and it is independent of \( b \), but \( x_0 = A^T b \) is a function of \( A \) and \( b \). This independence of \( b \) follows because \( \kappa(A) \) is the upper bound of the ratio of the relative error in \( x_0 = A^T b \) to the relative error in \( b \) with respect to all vectors \( b \in \mathcal{R}(A) \) and perturbations \( \delta b \in \mathbb{R}^m \).
2. The property \( b \in \mathcal{R}(A) \) is not, in general, satisfied in the LK method and thus a measure of the stability of \( x_0 \) with respect to perturbations in \( b \) that is appropriate when \( b \notin \mathcal{R}(A) \) must be developed.
Although these points show that the use of $\kappa(A)$ as a measure of the stability of $x_0$ has disadvantages, it is used by some researchers. For example, the computational reliability of $x_0$ is measured by the assignment of a threshold to $\lambda_1 = \sigma_1^2$ and/or $\lambda_2 = \sigma_2^2$ [4, §2.1], [18, §II.A], [19, §4.3.2], [20, §2]. This measure is used because $\sigma_2$ is the minimum distance of $A$ from singularity but it is not a measure of the stability of $x_0$ because it is independent of $b$. These disadvantages of $\kappa(A)$ are overcome by the effective condition number of $x_0$, which is a function of $A$ and $b$, and considered in Section 5.1.

The effective condition number is a linear condition number because it is assumed that the entries of $A$ (the spatial derivatives of the intensity) are not subject to error, and only errors in $b$ (the temporal derivatives of the intensity) are subject to error. A more realistic condition number of the optical flow requires that errors in the spatial and temporal derivatives of the intensity be considered. This inclusion leads to a non-linear extension of the effective condition number, and it is considered in Section 5.2.

5.1 The effective condition number

This section considers the effective condition number of the solution $x_0$ of the LS problem, which must be solved many times in the LK method. It is assumed for generality that the coefficient matrix $A$ is of order $m \times n$ where $m \geq n$, rather than $m \times 2$.

The relative errors of $x_0$ and $b$ are, respectively,
\[
\Delta x_0 = \frac{\|\delta x_0\|}{\|x_0\|} \quad \text{and} \quad \Delta b = \frac{\|\delta b\|}{\|b\|},
\] (17)
and the effective condition number of $x_0$ is defined as
\[
\eta(A, b) = \max_{\delta b \in \mathbb{R}^m} \frac{\Delta x_0}{\Delta b}.
\]

Theorem 1 The effective condition number $\eta(A, b)$ of $x_0$ is
\[
\eta(A, b) = \frac{\|c\|}{\sigma_n \|\Sigma^\dagger c\|}.
\] (19)

Proof It follows from (6) and (17) that
\[
\|\delta x_0\| = |A^\dagger \delta b| \leq |A^\dagger| \|\delta b\| = |A^\dagger| \|b\| \Delta b.
\] (20)
The division of this inequality by $\|x_0\| = \|A^\dagger b\|$ yields (18), from which (19) follows.

The condition that must be satisfied for $\kappa(A)$ to be a measure of the stability of $x_0$ follows from (20) because $\|b - Ax_0\| = 0$ if and only if $b \in \mathcal{R}(A)$. This condition is necessarily satisfied by all vectors $b$ if $A$ is square, but by some vectors $b$, not all vectors $b$, if $m > n$. If $b \in \mathcal{R}(A)$, then $b = Ax_0$ and thus $\|b\| \leq \|A\| \|x_0\|$, and (20) yields
\[
\|\delta x_0\| \leq |A^\dagger| \|b\| \Delta b \leq |A^\dagger| \|A\| \|x_0\| \Delta b.
\]

This leads to the definition of the condition number $\kappa(A)$ of $A \in \mathbb{R}^{m \times n}$, $m \geq n$, as the upper bound of the ratio of the relative error of $x_0$ to the relative error of $b$ if and only if $b \in \mathcal{R}(A)$,
\[
\kappa(A) = \max_{\delta b \in \mathbb{R}^m, b \in \mathcal{R}(A)} \frac{\Delta x_0}{\Delta b} = \|A^\dagger\| \|A\| = \frac{\sigma_1}{\sigma_n}.
\]
Furthermore, it follows from (10) and (19) that
\[
\max_{b \in \mathcal{R}(A)} \eta(A, b) = \max_{c \in \mathcal{R}(A)} \frac{\|c\|}{\sigma_n \|\Sigma^\dagger c\|} = \frac{\sigma_1}{\sigma_n},
\]
and thus $\kappa(A)$ is the maximum value of $\eta(A, b)$ with respect to all vectors $b \in \mathcal{R}(A)$.

The superiority of the effective condition number $\eta(A, b)$ with respect to the condition number $\kappa(A)$ as a measure of the stability of $x_0$ follows because $\eta(A, b)$ is a function of $A$ and $b$, but $\kappa(A)$ is only a function of $A$. The example of regression in [28, §3] shows the difference between these condition numbers. The dependence of $\eta(A, b)$ on $A$ and $b$ is clearly advantageous, but $\eta(A, b)$ must be used with care because it follows from (18) that the denominator contains the term $\|x_0\|$.

It therefore follows that $\eta(A, b)$ is ill conditioned if $x_0$ is ill conditioned, and it is well conditioned if $x_0$ is well conditioned. It is shown in [29, §4] that $\eta(A, b)$, and therefore $x_0$, are ill conditioned if the discrete Picard condition is satisfied [13, §4.5]. This condition states that if the constants $|c_i|$ decay to zero faster than the singular values $\sigma_i$ decay to zero, that is,
\[
\frac{|c_i|}{\sigma_i} \rightarrow 0 \quad \text{as} \quad i \rightarrow n,
\] (21)
then $x_0$ is ill conditioned. This condition can be derived from the solution $x_0$ of the LS problem,
\[
x_0 = A^\dagger b = \sum_{i=1}^{n} \left( \frac{c_i}{\sigma_i} \right) v_i,
\] (22)
where $v_i$ is the $i$th column of $V$. If $b$ is perturbed to $b + \delta b$, then the solution $x_0$ is perturbed to
\[
x_0 + \delta x_0 = \sum_{i=1}^{n} \left( \frac{c_i + \delta c_i}{\sigma_i} \right) v_i,
\] (23)
and there exists a constant $s$ such that the perturbations $|\delta c_i| \approx \epsilon$, $i = 1, \ldots, n$, satisfy
\[
|c_i| > \epsilon, \quad i = 1, \ldots, s,
\]
\[
|c_i| < \epsilon, \quad i = s + 1, \ldots, n.
\]
These inequalities follow because the perturbations $|\delta c_i|$ are approximately constant but the coefficients $|c_i|$ decay to zero because the discrete Picard condition (21) is satisfied, and thus from (23),

$$x_0 + \delta x_0 = \sum_{i=1}^{n} \left( \frac{c_i + \delta c_i}{\sigma_i} \right) v_i \approx \sum_{i=1}^{s} \left( \frac{c_i}{\sigma_i} \right) v_i + \sum_{i=s+1}^{n} \left( \frac{\delta c_i}{\sigma_i} \right) v_i.$$  

Since the discrete Picard condition is satisfied, $x_0$ can be approximated with a small error by only considering the first $s$ singular values in (22),

$$x_0 = A^t b \approx \sum_{i=1}^{s} \left( \frac{c_i}{\sigma_i} \right) v_i,$$

and thus

$$\delta x_0 \approx \sum_{i=s+1}^{n} \left( \frac{\delta c_i}{\sigma_i} \right) v_i.$$  

It follows that

$$\|\delta x_0\| \approx \epsilon \left( \sum_{i=s+1}^{n} \frac{1}{\sigma_i^2} \right)^{\frac{1}{2}} \approx \epsilon \frac{\epsilon}{\sigma_n},$$

and thus the norm of the perturbation $\delta x_0$ is approximately proportional to the magnitude of the noise and inversely proportional to the smallest singular value of $A$, from which it follows that $x_0$ is ill conditioned if the discrete Picard condition (21) is satisfied. It is seen that $x_0$ is ill conditioned if it is dominated by the large singular values of $A$, in which case $\delta x_0$ is dominated by the small singular values of $A$.

The satisfaction of the discrete Picard condition is necessary for the application of Tikhonov regularisation because it guarantees that the regularisation error is small and the regularised form of $x_0$ is numerically stable [28, §5]. Tikhonov regularisation is used for the removal of blur from images because experiments have shown that many images satisfy the discrete Picard condition [14]. This condition forms the prior information that guarantees that Tikhonov regularisation is effective for image deblurring, and more generally, the ill conditioned property of $\eta(A, b)$ limits its practical use if prior information is not available. The condition number $\eta(A, b)$ is, however, an important condition measure because it defines the conditions between $A$ and $b$ for which $x_0$ is well conditioned, and the conditions for which $x_0$ is ill conditioned.

Equation (21) defines the condition for which $x_0$ is ill conditioned, and it is well conditioned if

$$\frac{|c_{i+1}|}{\sigma_{i+1}} \gg \frac{|c_i|}{\sigma_i}, \quad i = 1, \ldots, n - 1,$$

because it follows that $\eta(A, b) \approx \|c\|/|c_n|$, which is independent of the singular values of $A$. The condition

$$|c_{i+1}| \gg |c_i|, \quad i = 1, \ldots, n - 1,$$

yields a different value of $\eta(A, b)$,

$$\eta(A, b) \approx \left( \sum_{i=1}^{m} c_i^2 \right)^{\frac{1}{2}} |c_n|.$$  

The condition number $\eta(A, b)$ may be ill conditioned but its upper bound is numerically stable, and an expression for this bound is developed in Theorem 2. This bound includes the term $\cos \theta$ that is defined in (16), and thus the error in the solution $x_0$ of the LS problem is related to the numerical condition of $x_0$.

**Theorem 2** If $A \in \mathbb{R}^{m \times n}$, $m \geq n$, then the effective condition number (19) of the solution $x_0$ of the LS problem is

$$\eta(A, b) = \frac{1}{\sigma_n} \left( \sum_{i=1}^{m} c_i^2 \sigma_i^2 \right)^{\frac{1}{2}} \leq \frac{1}{\sigma_n} \left( \sum_{i=1}^{n} c_i^2 \sigma_i^2 \right)^{\frac{1}{2}},$$

where $c = \{c_i\}_{i=1}^{m}$ is defined in (10), and thus $\eta(A, b)$ satisfies

$$\eta(A, b) \leq \begin{cases} \frac{\kappa(A)}{\cos \theta} & \text{if } m = n, \\ \frac{\kappa(A)}{\cos \theta} & \text{if } m > n. \end{cases}$$  

**Proof** If $m = n$, (19) yields

$$\eta(A, b) = \left( \frac{\sigma_1}{\sigma_n} \right) \left( \frac{\sum_{i=1}^{m} c_i^2}{\sum_{i=1}^{n} \frac{c_i^2}{\sigma_i^2}} \right)^{\frac{1}{2}} \leq \frac{\sigma_1}{\sigma_n} = \kappa(A),$$

and the result (24) for $m = n$ follows.

Consider now the situation $m > n$,  

$$\eta(A, b) = \left( \frac{\sigma_1}{\sigma_n} \right) \left( \frac{\sum_{i=1}^{m} c_i^2}{\sum_{i=1}^{n} \frac{c_i^2}{\sigma_i^2}} \right)^{\frac{1}{2}} \leq \left( \frac{\sigma_1}{\sigma_n} \right) \left( \sum_{i=1}^{n} \frac{c_i^2}{\sigma_i^2} \right)^{\frac{1}{2}} = \kappa(A) \frac{\|c\|}{\|c_n\|},$$

and the result (24) for $m > n$ follows from (16).  

Equation (24) shows that the condition number of $x_0$ increases rapidly as $\cos \theta \to 0$. For example, if a window of order $7 \times 7$ is used, then $A \in \mathbb{R}^{49 \times 2}$ and
the error in $x_0$ is equal to zero if and only if $b$ lies in the two dimensional subspace $\mathbb{R}_{\text{col}A}^2$ spanned by the columns of $A$, in the space $\mathbb{R}^4$, and the error in $x_0$ is due to the component of $b$ that lies in the space that is orthogonal to $\mathbb{R}_{\text{col}A}^2$. The condition number of $x_0$ is large if $\frac{1}{\cos \theta} > 1$, even if $\sigma_1/\sigma_2 = O(1)$ and thus the measure $\lambda_2 = \sigma_2^2$ of the stability of $x_0$ is incorrect because it does not include the term $\frac{1}{\cos \theta}$ in the expression for the condition number of the solution of the LS problem.

Example 1 Consider the matrix $A$ and vector $b$,

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$ 

The singular values of $A$ are $\sigma_1 = 1$ and $\sigma_2 = \sqrt{3}$, and thus $\kappa(A) = \sqrt{3}$. It may therefore be thought that the solution $x_0$ of the LS problem is well conditioned but this is incorrect because $b$ is orthogonal to $\mathcal{R}(A)$.

$$b^T A = [0 \ 0 \ 0].$$

It follows that $\cos \theta = 0$ and $x_0 = [0 \ 0]^T$.

The left singular matrix $U$ of $A = U \Sigma V^T$ is

$$U = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix},$$

and thus

$$U_1 = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \quad \text{and} \quad U_2 = \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$
and thus (27) becomes
\[
\rho(A, b) = \eta(A, b) + \kappa(A) \frac{\| (I - A A^T) b \|}{\sigma_n \| x_0 \|} 
\leq \eta(A, b) + \kappa(A) \frac{\| b \|}{\sigma_n \| A^T b \|} 
= \eta(A, b) + \kappa(A) + \kappa(A) \eta(A, b),
\]
and thus the inclusion of errors in the spatial derivatives of the intensity is significant because the effect of \( \eta(A, b) \) is magnified since it is multiplied by \( \kappa(A) \).

It was stated in Section 5.1 that the effective condition number \( \eta(A, b) \) is ill conditioned if the discrete Picard condition is satisfied, in which case (27) and its upper bound (28) cannot be computed reliably. This problem is addressed by using the upper bound (24) for the value of \( \eta(A, b) \) in (28). The condition number (27) is, however, useful because it shows that the inclusion of perturbations in the spatial derivatives of the intensity causes a significant increase in the condition number of the optical flow.

6 Example

This section contains two examples that show the errors and condition numbers of the optical flow at each level of the pyramid. The Middlebury data set [3] is usually used for the comparison of the results from different methods for the computation of the optical flow, and a popular measure for this comparison is the angular error (AngErr) between the computed optical flow \((u, v)\) and the ground truth optical flow \((u_{GT}, v_{GT})\).

The AngErr is defined as the angle between the vectors \((1, u, v)\) and \((1, u_{GT}, v_{GT})\).

\[
\text{AngErr} = \cos^{-1} \left( \frac{1 + (u \times u_{GT}) + (v \times v_{GT})}{\sqrt{1 + u^2 + u_{GT}^2} \sqrt{1 + v^2 + v_{GT}^2}} \right),
\]

and it is shown in [3, §4.1] that there are problems with this error measure. The constant one is added to the numerator and denominator to avoid the division %0 that occurs when \( u = v = u_{GT} = v_{GT} = 0 \). Any non-zero constant can be used to avoid this division and problems therefore arise when the computed optical flow and/or the ground truth optical flow are non-zero because the computed angular error is not a true measure of the exact angular error. These problems with the angular error lead to the absolute error (AbsErr), which is defined as

\[
\text{AbsErr} = \sqrt{(u - u_{GT})^2 + (v - v_{GT})^2}.
\]

This error measure does not suffer from the problems of the angular error and it is therefore a more practical error measure.

Examples 2 and 3 consider the following numerical measures:

1. The condition number \( \kappa(A) \), the upper bound of the effective condition number, \( \kappa(A) / \cos \theta \), and the non-linear condition number \( \rho(A, b) \).

2. The angle \( \theta \) between \( b \) and its component that lies in \( R(A) \).

3. The error \( e \) in the solution \( x_0 \) of the LS problem.

The examples use the images shown in Figure 2 [3].

**Example 2** The five measures listed above (the three condition numbers, the angle \( \theta \) and the error \( e \) in \( x_0 \)) were computed in each window for windows of side length \( 2s + 1 \), \( s = 2, 3, \ldots, 8, 9 \), at each level of a pyramid that has four levels, for the images in Figure 2. The mean and standard deviation of \( \log_{10} \kappa(A) \), the upper bound \( \log_{10} \kappa(A) / \cos \theta \) of the effective condition number, and the non-linear condition number \( \log_{10} \rho(A, b) \), at each level of the pyramid and for each window size, are shown in Figure 3. The surface plots in the left column show that, as expected, the values of the mean of \( \rho(A, b) \) at each of the 32 points in the grid are larger than the corresponding values of \( \kappa(A) \) and \( \kappa(A) / \cos \theta \). Also, the means of \( \rho(A, b) \) and \( \kappa(A) / \cos \theta \) increase, but not monotonically, as the pyramid is descended from the coarsest image (level 4) to the finest image (level 1). Similarly, these means decrease as the window size increases, at each level of the pyramid. The variation of the mean of these condition numbers, Figure 3 (left), is similar to the variation of their standard deviation, Figure 3 (right).

The mean and standard deviation of the angle \( \theta \), where \( \cos \theta \) is defined in (16), and the error \( \log_{10} e \), where \( e \) is defined in (9), are shown in Figure 4. It is seen that, at each level of the pyramid, the mean of \( \theta \) is approximately constant for windows of side length 7 pixels or more. The surface plots in Figure 4 (left) show that the variation of the mean of \( \theta \) and \( \log_{10} e \) are very similar, but the surface plots Figure 4 (right) show that the variation of their standard deviation is significantly different. It is seen that, for each measure, the standard deviation is significant, which implies that there is considerable variation of each measure.

**Example 3** The LK method with a pyramid of three levels was applied to the images shown in Figure 2, with a window of size \( 15 \times 15 \). The condition number \( \log_{10} \kappa(A) \), upper bound \( \log_{10} \kappa(A) / \cos \theta \) of the effective condition number, the non-linear condition number \( \log_{10} \rho(A, b) \), angle \( \theta \) and error \( \log_{10} e \) were computed for every LS problem in the windows for each level of the pyramid. The results are shown in Figures 5, 6, 7, 8 and 9, and it is seen that they increase as the pyramid is descended. Furthermore, the variance in each of
these measures increases as the pyramid is descended and the figures are therefore consistent with the graphs of the standard deviation in Figures 3 and 4. It follows that the computational reliability of the optical flow decreases as the pyramid is descended.

Figure 8 shows that the values of the local maxima of the angle $\theta$ increase significantly as the pyramid is descended. In particular there are many LS problems for which $\theta$ is about 80 degrees, from which it follows that the error in the solutions of these problems is large. This result is consistent with the plots of the condition numbers in Figures 6 and 7. Furthermore, it follows from (15) large values of $\theta$ are associated with large errors $e$, which is confirmed in Figure 9.

7 Summary

This paper has considered numerical properties of the solution $x_0$ of the LS problem and it has been shown that the large values of the condition numbers $\eta(A, b)$ and $\rho(A, b)$, and error $e$, of $x_0$ arise because a significant component of $b$ lies in the space that is orthogonal
Fig. 4 The variation of the mean (left) and standard deviation (right) of the angle $\theta$, where $\cos \theta$ is defined in (16), and the error $\log_{10} \epsilon$, where $\epsilon$ is defined in (9), with the size of the window and number of levels in the pyramid, for Example 2.

Fig. 5 The condition number $\log_{10} \kappa(A)$ at level three (top), level two (middle) and level one (base) of the pyramid, for Example 3.
to $\mathcal{R}(A)$ and thus the terms $1/\cos \theta$ and $\cos \theta$, in, respectively, the expressions for the condition numbers and error are significant. The physical reason for these large values follows from the assumption that the optical flow in each window at each level of the pyramid is constant, that is, the motion in each window can be represented by a translation. It follows that an improved form of the LK method requires that a richer class of motions be allowed in each window, and this is achieved by the introduction of warping functions in the LK method [2]. These warping functions are defined by parametric models, for example, affine and projective models, and the values of the parameters of these models are computed in the optimisation procedure in an extended form of the LK method.

Conflict of interest

The author declares that he does have not relationships or interests that have direct or potential influence, or impart bias, on the work.

Data availability

The data used in this paper are taken from the Middlebury data set [3], which is in the public domain and frequently used in computer vision.

References

Fig. 7 The non-linear condition number $\log_{10} \rho(A, b)$ at level three (top), level two (middle) and level one (base) of the pyramid, for Example 3.

Fig. 8 The angle $\theta$ at level three (top), level two (middle) and level one (base) of the pyramid, for Example 3.

Fig. 9 The error $\log_{10} e$ at level three (top), level two (middle) and level one (base) of the pyramid, for Example 3.