Towards Fuzzy Calibration

C. V. Jawahar and P. J. Narayanan

International Institute of Information Technology
Gachibowli, Hyderabad 500019. INDIA
jawahar@iiit.net pjn@iiit.net

Abstract. A framework for fuzzy calibration is introduced here. Fuzzy calibration is necessary to account for the imprecision of the camera model that can be computed during calibration. It is the first step of a fuzzy vision framework in which the uncertainties are propagated forward through the different levels of processing until precise values are absolutely necessary. We present the fuzzy calibration framework for weak and strong calibration. We also present some thoughts outlining how such calibration can be used in higher levels of vision processing.

1 Introduction

Calibration is most fundamental to stereo vision. Calibration is the process of computing the parameters of a standard mathematical model that represents the imaging process of a camera. The pin-hole camera model is most popular and adequately represents the projection process of a real camera. For applications that use lenses of low-focal lengths, the lens distortion parameters can be recovered and used to correct distortions such as the radial lens distortion. Different calibration algorithms have been proposed and implemented by Computer Vision researchers in the past couple of decades.

The calibration as used in Computer Vision can be broadly classified into two categories: Strong and Weak [3]. Strong calibration fits a pin-hole model to the measured projection characteristics of the camera. That is, given a point $\bar{X}$ in the 3D space, its projection $\bar{x}$ to the image space can be computed via a matrix M, called the *Projection Matrix*, calculated by strong calibration. Weak calibration, on the other hand, provides the relative geometry between two or more views of the same scene. The *Fundamental Matrix* $\mathbf{F}$ of a pair of views encodes the relations of how the mapping of a pixel in one image is constrained in the other. In presence of more than two views, trifocal tensors and other linear relationships can provide the weak calibration.

Calibration is a hotly debated area among Computer Vision researchers. One school of thought believes strong calibration is essential and relies on it for applications such as 3D metric measurements. Another school discards a priori calibration because of its unreliability. We believe that a middle path that combines the advantages of each can yield the best results. We will use the best available calibration tools, but will not trust their output completely. The ambiguities and uncertainties in the calibration will be modelled and used in the subsequent levels of processing. Fuzzy set theory can provide a theoretical framework for modelling the uncertainty associated with the ambiguity and vagueness of image measurements and is best suited to model the ambiguity of calibration.

Fuzzy notions have been employed extensively in image processing and pattern recognition. However, their applications to higher-level vision have been quite limited in the literature. [1, 6, 7]. The uncertainty incurred in the initial stages of camera calibration can cause errors down the line if propagated as such. We believe that a fuzzy calibration followed by processing of a fuzzy set of correspondences can preserve the ambiguities in the low-level information so that a better decision can be taken at a higher level, if more precise information is available.

Section 2 presents a brief description of the notion of fuzzy calibration and how it differs from the conventional calibration. Strong and weak camera calibration schemes are extended to fuzzy notions in Section 3. In section 4, we describe the scope and applicability of the fuzzy calibration procedures. Section 5 presents a few concluding remarks.

2 Fuzzy Calibration: The First Step

Calibration errors have a great impact on stereo vision systems that use calibration. The calibration parameters are in practice estimated from a large number of observations, typically using a non-linear optimization procedure. The parameters are used as precise values representing the camera, though the estimation process is aware of the mismatch between the model and the measurements. Thus, the uncertainty in the calibration process is ignored completely. This can result in further errors in the subsequent stages of processing. Thus, it is essential for the calibration step to preserve the uncertainties that are found. For a task at a higher level such as stereo vision, a fuzzy set of correspondences can be computed, using the fuzzy calibration data. The uncertainties in the correspondence should reflect those in the calibration as well as new uncertainties introduced by the matching step. The depth values computed from these correspondences will take into account the known uncertainties till then. This process should be carried forward at all levels till a crisp decision is unavoidable. For example, crisp depth values could be calculated when the depth map has to be converted to a triangulated model for display.

Better results can thus be expected if the uncertainty involved is modelled explicitly and used by all levels of processing. Fuzzy sets and numbers provide an excellent framework for this. It has already been applied quite successfully to many low-level vision tasks. Algorithms that take fuzzy inputs can produce suitable fuzzy outputs representing the uncertainty for the subsequent level of processing. Most vision algorithms can modified easily to handle fuzzy data by replacing each operation by its fuzzy equivalent. Well defined processes are available to convert fuzzy results into crisp ones. The process of handling the uncertainties in vision should rightly begin with *fuzzy calibration*, as it is typically the first step in a vision process.

The right place to start fuzzy vision processing is calibration. Calibration is computed from a few known points specified manually or computed using an appropriate algorithm. It is also possible to boot-strap the process starting with a few manually selected points, as an algorithm can find more known points if an approximate calibration exists. Calibration data consists of a few parameters for the camera model. Conceptually, fuzzy calibration uses fuzzy numbers for each of those parameters. The uncertainties encoded by them will result in different, and more general, constraints. The fuzzy numbers and their associated membership functions can be computed from the error measures used by the linear or non-linear optimization algorithm used for conventional calibration.

3 Fuzzy Calibration Models

In this section, we briefly introduce the models of weak and strong calibration mathematically and discuss how they can be modified to fuzzy models.

3.1 Weak Calibration

The weak calibration of two cameras is given by the fundamental matrix **F** that encodes the constraints between the images of the same world point in two views [3, 4]. If $\bar{x}$ and $\bar{x}'$ are the homogeneous coordinates of the projections of the same point in the left and right views, the following relation holds about them.

$$\bar{x}'^{\mathrm{T}} \mathbf{F} \bar{x} = 0 \tag{1}$$

F is 3×3 matrices of rank 2 and can be computed from 8 or more pairs of matching points using a linear algorithm. In practice, there are linear, nonlinear and statistically optimal algorithms that use a large number of points matching points for computing the fundamental matrix [4, 8].

Let us look at a simple arrangement of two cameras, called a parallel, rectified camera configuration. The fundamental matrix for such a configuration is given by [4]

$$F = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \tag{2}$$

The arrangement may only be approximately parallel in practice. Equation 1 can be expanded for two matching points $[x', y', 1]^{\mathrm{T}}$ and $[x, y, 1]^{\mathrm{T}}$ as

$$x'(xf_{11} + yf_{12} + f_{13}) + y'(xf_{21} + yf_{22} + f_{23}) + (xf_{31} + yf_{32} + f_{33}) = 0. \tag{3}$$

Since the overall scale is unimportant, we can safely assume $f_{23} = 1$ for the parallel rectified situation, since we know that number is non-zero for the ideal situation. Equation 3 reduces to $y' - y = 0$ ideally. However, the parallel, rectified model could be slightly incorrect in a few ways in a practical situation, in spite of the best efforts at making the cameras parallel and rectified.

1. Incorrect alignment of scan lines. That is, Equation 3 reduces to $y' - y + \delta = 0$.
2. Incorrect magnification. Equation 3 reduces to $y' - y(1+\gamma) = 0$ in this case.
3. Non-parallel situation. Equation 3 reduces to $\epsilon x' + y' - y = 0$ in this case.

In all these cases, the values $\delta, \epsilon,$ and γ are very small, ideally zero. The fundamental matrix for a practical parallel, rectified camera arrangement can account for such uncertainties in measurements and can be given by

$$F = \begin{bmatrix} 0 & 0 & \epsilon \\ 0 & 0 & 1 \\ 0 & -(1+\gamma) & \delta \end{bmatrix}. \tag{4}$$

At this stage, two important questions are to be answered: (a) Can we estimate $\delta, \epsilon,$ andγ from observations? (b) Can a precise set of crisp values model the uncertainty in the calibration?

The answer to the first question is in the affirmative. The number of image measurements is typically more than the number of unknowns and one could employ a least mean/median square solution [2–4]. The end results give crisp values for the entries of **F**, forcing the corresponding point to lie on an epipolar line in the second image. A better way to capture such uncertainty is to consider the image measurements as fuzzy measurements, obtained on a discrete 2D grid using projection of a 3D point. The values of δ, ϵ, and γ can be computed as fuzzy numbers using fuzzy regression techniques [5]. We can derive a fuzzy set in the second image where the corresponding feature point should lie using the fuzzy **F** matrix. It will be interesting to investigate what the correspondence implies in this case. We will discuss this problem in the next section. The notion of fuzzy feature measurements allow us to handle the points "in and around" the conventional epipolar line.

3.2 Strong Calibration

The projection equation given the strong calibration matrix M is given by $\bar{x} = M\bar{X}$. The matrix M is a 3×4 matrix as homogeneous coordinates are used to represent points in both the 2D and 3D spaces due to the mathematical ease of dealing with rotations and translations uniformly using them. When the point $\bar{X}$ is specified in a projective or affine space instead of a Euclidean space, corresponding projective or affine calibration matrices can be used. The projection matrix obeys appropriate constraints in each of these spaces and can be computed from the world and image coordinates of a number of points.

We can decompose the projection matrix into the *intrinsic* and *extrinsic* components

$$M = PV = \begin{bmatrix} f_x & 0 & 0 & 0 \\ 0 & f_y & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R \,; T \\ \mathbf{0} \quad 1 \end{bmatrix}, \tag{5}$$

where P is the perspective projection matrix for points specified in a coordinate frame rooted at the camera center, V is the view transformation matrix that transforms the world coordinate frame to the camera coordinate frame. f_x and f_y are the focal lengths in the two directions, and R and T are the rotation matrix and the translation vector respectively that align the world coordinate system to the camera coordinate system.

A typical calibration algorithm [9] gives crisp values to the camera parameters. Achieving precise parameters from measurements has been a difficult task [3, 4]. However, We can treat the camera parameters as fuzzy quantities to account for the uncertainties. A discussion on the complete mathematical formulation of the framework is beyond the scope of this paper. However, we provide a glimpse of the geometric interpretation of the fuzzy calibration parameters now.

Let us look at the region of the world (i.e., the subset of $\mathbb{R}^3$) that maps to a pixel. For that we have to "invert" the projection matrix. We have that $M^{-1} = V^{-1}P^{-1}$. The inverse of the view transformation matrix is given by

$$V^{-1} = \begin{bmatrix} R^{\mathrm{T}} \,;\, \mathbf{0} \\ \mathbf{0} \quad 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} \,;\, -T \\ \mathbf{0} \quad 1 \end{bmatrix},$$

since R is an orthonormal rotation matrix. The inverse of P is not uniquely defined. We can decompose P as $[Q; 0]$ where Q is a 3×3 matrix consisting of the left 3 columns of P (see Equation 5). We now can see that,

$$Q^{-1} = \begin{bmatrix} 1/f_x & 0 & 0 \\ 0 & 1/f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$

Now, we can "invert" the projection matrix as

$$M^{-1}\bar{x} = V^{-1}P^{-1}\bar{x} = V^{-1} \begin{bmatrix} Q^{-1}\bar{x} \\ k \end{bmatrix} = \begin{bmatrix} R^{\mathrm{T}} \,;\, \mathbf{0} \\ \mathbf{0} \quad 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} \,;\, -T \\ \mathbf{0} \quad 1 \end{bmatrix} \begin{bmatrix} Q^{-1}\bar{x} \\ k \end{bmatrix} \tag{7}$$

where k is a constant that cannot be computed. Together, the above equation gives the parametric equation of the imaging ray for the pixel $\bar{x}$, the parameter being k. Equation 7 gives the unbounded volume of space that projects to the pixel in the global coordinate frame. If P and V give precise projection and transformation, the volume that projects to the pixel is a ray called the *imaging ray*. If uncertainties in the focal lengths or the rotation and translation estimates can be encoded using fuzzy numbers in the corresponding matrices, these equations specify an *imaging volume*, which is the generalization of the imaging ray. We can compute a membership functions in the fuzzy framework for each ray in the imaging volume.

Comments We presented fuzzy models for weak and strong calibration in this section. Fuzzy fundamental matrix can be computed from crisp or fuzzy points in multiple images, relaxing the strong epipolar assumption. This can account for the possible inaccuracies of imaging models and the rectification process. Fuzzy strong calibration can account for uncertainties in estimation of parameters such as the focal length, rotations, and translations. We can compute the geometrical imaging volume with associated distribution of uncertainties within it as the possible locations of the 3D scene point that projects to each pixel.

4 Fuzzy Vision: Beyond Calibration

In this section, we outline how the fuzzy calibration can be used for other kinds of vision processing. There are multiple ways to handle the imprecision of the calibration process and to develop algorithms using it. We present simple approaches to some problems in this section. We use the notion of *fuzzy correspondence* for the examples given here. Fuzzy correspondence gives a set of points in the right image with associated membership functions for each pixel in the left image, based on fuzzy feature measurements that encode similarity [7]. We can study fuzzy correspondences and fuzzy depth maps with the help of the rich literature available on fuzzy arithmetic.

Fuzzy Correspondences and Weak Calibration: In a typical situation, we have two problems: obtaining the fuzzy correspondences and estimating the fundamental matrix from them. Here, we outline an iterative algorithm which carries out both simultaneously starting with a set of crisp pixel matches given by a standard matching algorithm.

Let $\{\mathbf{x_1}, \mathbf{x_2}, \ldots \mathbf{x_n}\}$ be a set of pixels in the first image and $\{\mathbf{y_1}, \mathbf{y_2}, \ldots, \mathbf{y_n}\}$ be the pixels in the second image. We can assign memberships to the corresponding pairs of matches based on a similarity measure for the feature points. Such correspondences need not be geometrically valid and the second point may not lie "near" the epipolar line. If μ_i is the membership of the pair $(\mathbf{x_i}, \mathbf{y_i})$ in the fuzzy correspondence, weak calibration can be obtained by minimizing $J = \sum_i \mu_i(d(x_i, Fy_i) + d(y_i, F^T x_i))$, where $d(x_i, Fy_i)$ gives the distance of the point x_i to the epipolar line given by Fy_i. The minimization of the above objective function is carried out in two stages. First, compute the fundamental matrix based on the eigenvectors of $X^{\mathrm{T}}MX$ where M is a diagonal matrix with fuzzy memberships as the diagonal elements and X is a measurement matrix. More detailed description of a non-fuzzy implementation of this computation can be seen in [4]. Second, assign memberships based on the nearness of the point to the epipolar line. $\mu_i = \frac{k}{(d(x_i, Fy_i) + d(y_i, F^T x_i))}$, for a constant k. It can be shown that the above iterative algorithm converges and provides an optimal estimate of the fuzzy correspondence and the weak calibration simultaneously.

Fuzzy 3D Point from Stereo We saw how fuzzy strong calibration can compute the imaging volume corresponding to a pixel in Section 3.2. When the correspondence between two pixels of an image pair is known, their imaging volumes can be intersected like in stereo vision. The correspondence used itself could be fuzzy as described above, adding an additional level of uncertainty. The three levels of imprecision are: (a) the pixel of the left image represents a set of imaging rays with associated membership values μ_i^{L}, (b) the correspondence pairs it up with a set of pixels in the right image with its own membership values μ_{ij}^{c}, and (c) each corresponding pixel represents a set of imaging rays with memberships μ_j^{R} using the right camera's calibration data. We can intersect the viewing volumes taking into account the three uncertainty measures. The result will be a region in space (a fuzzy subset over $\mathbb{R}^3$) with an associated confidence measure for each point in the region. The distribution of the confidence measure is a function of the three membership functions given above. An appropriate T-norm may be used for this purpose. A more detailed treatment of this is beyond the scope of this paper. This results in a disparity/depth map where the disparity/depth estimate at each point is a fuzzy set, or otherwise we have multiple depth estimates with varying amount of certainty.

5 Conclusions and Future Work

We proposed a fuzzy calibration scheme to incorporate the uncertainty associated in the imaging process into the subsequent levels of processing. Calibration is indeed the first step in many vision processes. It is important to keep track of the uncertainties in calibration and use it at higher levels. A fuzzy framework is ideally suited for this. We intend to investigate fuzzy notions for higher level vision processes such as modelling using multiple cameras, new view generation, etc. We believe the fuzzy framework will yield better results to such problems.

References

1. M. Ben-Ezra, S. Peleg, and M. Werman. Efficient Computation of Most Probable Motion from Fuzzy Correspondences . In *IEEE Workshop on Applications of Computer Vision*, Oct. 1998.
2. Z. Zhang *et. al.* A Robust Technique for Matching Two Uncalibrated Images Through the recovery of unknown epipolar geometry. *Artificial Intelligence Journal*, 78:87 – 119, 1995.
3. O. D. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint.* MIT Press, 1993.
4. R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge University Press, ISBN: 0521623049, 2000.
5. K. Jajuga. Linear Fuzzy Regression . *Fuzzy Sets and Systems*, 20:343 – 353, 1986.
6. C. V. Jawahar. Stereo Correspondence Based on Correlation of Fuzzy Texture Measures. In *Indian Conference on Computer Vision, Graphics and Image Processing*, pages 273–278, 1998.
7. C. V. Jawahar, A. M. Namboodiri, and P. J. Naryanan. Integration of Stereo Correspondence Based on Fuzzy Notions. In *International Conference on Advances in Pattern Recognition and Digital Techniques (ICAPRDT)*, December 1999.
8. K. Kanatani. Optimal fundamental matrix computation: algorithm and reliability analysis . In *Proc. of the SSII2000*, pages 14–16, Jun. 2000.
9. R. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323 – 344, 1987.