Noname manuscript No.
(will be inserted by the editor)

A dyadic solution of relative pose problems

Patrick Erik Bradley

Received: date / Accepted: date

Abstract A hierarchical interval subdivision is shown to lead to a p-adic encoding of image data. This allows in the case of the relative pose problem in computer vision and photogrammetry to derive equations having 2-adic numbers as coeffients, and to use the Hensel lifting method to their solution. This method is applied to the linear and non-linear equations coming from eight, seven or five point correspondences. An inherent property of the method is its robustness.

Keywords Relative pose · p-adic numbers · Essential matrix · n-point method

1 Introduction

The issue of estimating camera motion from two views is classical by now, and methods from projective and algebraic geometry towards its solution were employed at an early stage (e.g. [15]). The beginning of this present century witnesses the application of sophisticated methods from computational commutative algebra in order to rephrase the equations into a form from which solutions can be obtained with relative ease. The relationship between the views is established by finding correspondences between point pairs taken from both images. The fundamental matrix faithfully encodes the geometric relationship between the two images. For normalised cameras, the fundamental matrix coincides with the essential matrix. In general, the two matrices are related through the camera calibration. Hence, if the calibration is known, it is sufficient to estimate the essential matrix in order to solve the relative pose problem. From a conceptional as well as a computational point of view, it makes sense to use only few correspondences of image points in order to estimate the essential matrix E. And different samples of n correspondences lead to a set of candidate essential matrices from which an optimal choice can be made.

Since $E = (e_{ij})$ is a projective 3×3 matrix, i.e. only determined up to a scalar factor, the number of parameters to be found is at most 8. As each pair of corresponding image points leads to a linear equation in the e_{ij}, it suffices to take $n = 8$, as has been showed in [11]. However, this linear method does not take into account further constraints on E. So, a 7-point algorithm was developped [6,9] (cf. also [17] for an overview). The 5-point algorithm of [13] uses the minimal number of point correspondences necessary for determining E. The non-linear constraints lead to homogeneous equations of degree three in four unknowns which are transformed through an elimination process into a univariate equation of degree 10. Hence the number of complex solutions is not more than 10. The elimination process can be simplified [10], or replaced by Gröbner basis methods [14]. There is a lot of existing work towards optimising the performance of this method in the computational algebraic geometry community. Stewenius et al. compare the performance of elimination and Gröbner basis methods [16].

Correspondences often suffer a geometric perturbation, i.e. the correct point u' in image I' corresponding to point u in image I is mostly found up to a small error vector ϵ such that $u' + \epsilon$ instead of u' is matched to u. In this article, a hierarchical method based on interval subdivision is developped to handle this problem. A natural way of encoding hierarchies is provided by the p-adic numbers, where p is a fixed prime number. In

P.E. Bradley
Institut für Photogrammetrie und Fernerkundung, Universität Karlsruhe (TH), Englerstr. 7, 76131 Karlsruhe, Germany
E-mail: bradley@ipf.uka.de

our case, we can use $p = 2$ and represent image points by pairs of binary expansions

$$a = \sum_{n=0}^{\infty} a_i 2^i$$

with coefficients a_i equal to 0 or 1. These expansions can be infinite, theoretically. Practically, the finiteness of resolution means approximation through truncation. The framework for this method is p-adic geometry which has been applied in video segmentation and data analysis [2,12,3].

In the context of relative pose problems, we will use the so-called *lifting method* for solving the equations. This is provided by Hensel's lemma which says that under certain conditions a solution of a given equation modulo p can be expanded to a p-adic solution. Its proof is constructive. In fact, this is the p-adic analogue of a Newton iteration. Applied to the equations of the n-point problems for $n = 5, 7, 8$, we obtain the result that for many choices of point correspondences, a 2-adic solution can be constructed. The 2-adic essential matrix then allows to hierarchically approximate the rigid motion by truncation. A side effect of the lifting method is its high robustness to geometric perturbations. The encoding method ensures further that the number of Newton iterations required is proportional to the order of resolution desired.

The article is structured as follows. The next two sections are a brief introduction to p-adic numbers, and to Hensel's lemma in a multivariate formulation. This is followed by a section on 2-adic encoding of image pixels for a p-adic camera model. The last section applies Hensel's lemma to the 2-adically defined equations for the eight-, seven-, and five-point problems.

2 p-adic numbers

The p-adic numbers were first described by Hensel in [8]. They are expansions of the form

$$a = \sum a_i p^i \tag{1}$$

into powers of a fixed prime number p and coefficients $a_i \in \{0, \dots, p-1\}$. If there are only finitely many terms in (1), this defines a rational number. Any natural number has a finite p-adic expansion (1) with $i \geq 0$. The important observation is that expansions with infinitely many positive powers of p have a meaning. Namely, by defining

$$|a|_p = p^{-\nu_a},$$

where $\nu_a \in \mathbb{Z}$ is the smallest exponent occurring in expansion (1), one obtains a metric for which the partial sums

$$a_N = \sum_{i \leq N} a_i p^i$$

converge to a:

$$|a - a_N|_p = \frac{1}{p^{N+1}} \longrightarrow 0 \quad \text{for} \quad N \longrightarrow \infty.$$

The domain of all p-adic numbers is denoted by $\mathbb{Q}_p$ and densly contains the rational numbers $\mathbb{Q}$ with respect to this metric. Those p-adic numbers a with $\nu_a \geq 0$ are the p-adic integers, denoted as $\mathbb{Z}_p$. This domain contains densely the usual integers $\mathbb{Z}$. An equivalent description of p-adic integers is given by

$$\mathbb{Z}_p = \left\{ a \in \mathbb{Q}_p \mid |a|_p \leq 1 \right\}.$$

Approximation of p-adic integers by their partial sums, simply termed "truncation", has an algebraic formulation. In its simplest form, a p-adic integer a can be given by its coefficient $a_0 \in \{0, \dots, p-1\}$, and another p-adic integer b having the same coefficient $b_0 = a_0$ approximates a up to that order of magnitude. This is the case if and only if $a - b$ is divisible by p. Hence, we arrive at

$$\mathbb{Z}_p / p\mathbb{Z}_p \cong \mathbb{Z}/p\mathbb{Z} \cong \mathbb{F}_p,$$

where the latter is the finite field with p elements. Similarly, the algebraic formulation of approximation up to terms of higher order is given by

$$\mathbb{Z}_p / p^N \mathbb{Z}_p \cong \mathbb{Z}/p^N\mathbb{Z} =: \mathcal{Z}_{p^N}.$$

In other words, congruences modulo p^N yield finite approximations of p-adic numbers by their partial sums up to terms of order N. It is this property which makes p-adic numbers very suitable for hierarchically organised data. In later sections, we will see how this algebraic formulation can be used in stereo vision. A standard reference for p-adic numbers is [5].

3 Hensel's lemma

An important method in p-adic analysis is the so-called "lifting" of zeros of polynomials from $\mathbb{F}_p$ to $\mathbb{Z}_p$. This uses the fact that

$$\mathbb{Z}_p / p\mathbb{Z}_p \cong \mathbb{Z}/p\mathbb{Z} \cong \mathbb{F}_p$$

by reducing a given equation $f(X) = 0$ over $\mathbb{Z}_p$ modulo p to an equation $f \mod p$ over $\mathbb{F}_p$. Hensel's lemma then gives a criterion when a solution x_1 of $f \mod p$ leads to a solution of f. This solution ξ, if it exists, is constructed by an iterated process. Namely, first a solution x_2 of $f \mod p^2$ is constructed from x_1, and from this a solution x_3 of $f \mod p^3$ etc. Each step of this iteration yields an approximation to the true solution in $\mathbb{Z}_p$ in the sense that

$$f(x_i) \equiv 0 \mod p^i,$$

which in terms of the p-adic norm translates to

$$|f(x_i)| < \frac{1}{p^i}.$$

In other words, the sequence $f(x_i)$ converges p-adically to the value $f(\xi) = 0$. The construction process guarantees further that x_i also converges to the p-adic solution ξ.

We now state a multivariate Hensel's lemma, but not in the most general form. For this, $\mathbb{Z}[X_1, \dots, X_n]$ denotes the polynomial ring in n variables with integer coefficients.

Theorem 1 (Multivariate Hensel's lemma) *Let*

$$f = (f_1, \dots, f_m) \in \mathbb{Z}[X_1, \dots, X_n]^m$$

with $m \leq n$, and let $k \geq 2$. Suppose that the vector $x = (x_1, \dots, x_n) \in \mathbb{Z}^n$ is a solution of the congruences

$$\begin{aligned} f_1(X_1, \dots, X_n) &\equiv 0 \mod p^{k-1} \\ &\vdots \\ f_m(X_1, \dots, X_n) &\equiv 0 \mod p^{k-1} \end{aligned}$$

and that the matrix

$$D_f(x) = \left(\frac{\partial f_i}{\partial X_j} \right)$$

is modulo p of rank m. Then there exist $t = (t_1, \dots, t_n) \in \{0, \dots, p-1\}^n$ such that

$$f(x + p^{k-1}t) \equiv 0 \mod p^k.$$

In particular, the equations $f(X) = 0$ have a solution in $\mathbb{Z}_p$.

Proof Consider the linear part in the Taylor expansion of f in x:

$$f(X) = f(x) + D_f(x) \cdot (X - x) + \text{terms of higher order},$$

where $X = (X_1, \dots, X_n)$ denotes the vector of variables. By assumption, it holds true that

$$f(x) = p^{k-1} \cdot a$$

for some vector $a \in \mathbb{Z}^m$. Due to the rank condition, the system of congruences

$$a + D_f(x) \cdot t \equiv 0 \mod p \tag{2}$$

has a solution vector $t \in \{0, \dots, p-1\}^n$. Hence,

$$\begin{aligned} f(x + p^{k-1}t) &\equiv f(x) + D_f(x) \cdot p^{k-1}t \\ &\equiv p^{k-1}(a + D_f(x) \cdot t) \equiv 0 \mod p^k. \end{aligned}$$

Iteration proves the last assertion.

Usually, Hensel's lemma is stated in the case of a single univariate polynomial f having modulo p^{k-1} a zero x. The rank condition translates to $f'(x) \not\equiv 0 \mod p$, i.e. x is a simple zero of f modulo p. The proof is a p-adic analogue of a Newton iteration. From (2) it follows for $m = n$ that the rank condition implies uniqueness of the lift. The univariate case can be found e.g. in [5].

4 p-adic projective cameras

4.1 p-adic encoding of image pixels

Assume that a rectangular 2D image I is given by $M \times N$ pixels and that m, n are minimal such that $M \leq 2^m$, $N \leq 2^n$. We will use a binary encoding of a given image point with pixel coordinates

$$(x, y) \in \{0, \dots, M-1\} \times \{0, \dots, N-1\}.$$

This means that x and y will be represented by a pair of binary numbers obtained in a hierarchical manner. Namley, consider one coordinate at a time, say x. It can be arrived at by a sequence of iterated subdivisions of the interval $0 < \dots < 2^m - 1$ into intervals of equal length. After m iterations each interval obtained contains precisely one pixel x-coordinate. This means that a given value of x lies in a uniquely determined nested sequence of intervals produced by this subdivision process. The intervals form a rooted binary tree whose leaves correspond uniquely to the x-coordinates of pixels. By assigning for a given interval its left half the value 0, and 1 for its right half, we obtain the binary representation

$$r(x) = \sum_{\nu=0}^{m-1} \alpha_\nu 2^\nu, \quad \alpha_\nu \in \{0, 1\}$$

by traversing the path from root down to x and picking up the zeros and ones along the way.

We will interpret the natural number $r(x)$ as a 2-adic integer: $r(x) \in \mathbb{Z}_2$. Each partial sum

$$r_\ell = \sum_{\nu=0}^{\ell} \alpha_\nu 2^\nu$$

is a 2-adic approximation of $r(x)$. Its error is bounded by the 2-adic distance

$$|r_\ell - r(x)|_2 .$$

In fact, $r(x)$ itself is a 2-adic approximation to some imaginary "pixel" ξ of infinite precision. The number m is given by the degree of resolution with which ξ is viewed on the given image. And an infinitely precise pixel would have a 2-adic expansion with possibly infinitely many coefficients equal to 1.

We proceed similarly for the y-coordinate, and obtain an encoding

$$c_2\colon I \to \mathbb{Z}_2 \times \mathbb{Z}_2, \quad (x,y) \mapsto (r(x), s(y)),$$

where s is the binary encoding of the y-coordinate.

Any p-adic encoding of a number x bears the problem that its p-adic approximations r_ℓ are determined by arithmetic properties of x, ignoring their (euclidean) geometric properties. This means that although x_ℓ could be p-adically very close to x, it could be at large distance for the euclidean norm. Hence, arbitrary p-adic encodings of pixel data would be quite unsuitable for many applications in computer vision. However, the binary encoding through interval splitting does not suffer this disadvantage. Namely, the map c_2 respects the image geometry in the following sense:

Proposition 1 *The map*

$$\iota_2\colon \sum_{\nu=0}^{\infty} \alpha_\nu 2^\nu \mapsto \sum_{\nu=0}^{\infty} \alpha_\nu 2^{-\nu}$$

yields an inclusion of $\mathbb{Z}_2$ into the real interval $[0,2)$. It has the property

$$|r-s|_2 < 2^{-\ell} \Rightarrow |\iota_2(r) - \iota_2(s)|_\infty < 2^{-\ell}$$

for all $\ell \in \mathbb{N}$.

Proof Any sum of negative powers of 2 can be majorised by a (possibly shifted) geometric sum. Hence, such a sum converges for the real metric. This implies that the map ι_2 is well-defined for arbitrary 2-adic numbers.

The inclusion property is clear, at least for finite 2-adic expansions. However, the map ι_2 is injective on all of $\mathbb{Z}_2$. Otherwise, assume that

$$\iota_2(a) = \iota_2(b)$$

for some $a \neq b$. Then $a = \sum a_i 2^i$ and $b = \sum b_i 2^i$ must differ in some coefficient:

$$a_n \neq b_n.$$

Let n be minimal such that this occurs. Then it follows that

$$|\iota_2(a) - \iota_2(b)|_\infty \geq \frac{1}{2^N},$$

because the inequality holds true for all partial sums with at least N terms. This is a contradiction. Hence, ι_2 is injective.

ssume $|r-s|_2 < 2^\ell$. This means that the 2-adic expansions of r and s have the first $\ell+1$ terms in common. The last assertion follows from this.

As a consequence, we obtain an embedding of the image I into $[0,2)^2$ via composition of c_2 and $\iota_2 \times \iota_2$.

4.2 p-adic camera model

Recall that a projective camera is a map between projective spaces

$$\kappa\colon \mathbb{P}^3 \to \mathbb{P}^2$$

given by a 3×4-matrix of rank 3. Usually, cameras are modelled as being defined over the real numbers, i.e. they are given by real 3×4-matrices. However, when dealing with data coming from 3D-objects from the real world, cameras are usually approximated by rational matrices. In this way, we arrive at rational cameras:

$$\kappa\colon \mathbb{P}^3(\mathbb{Q}) \to \mathbb{P}^2(\mathbb{Q})$$

as maps between the rational points of projective spaces. For such cameras, the real completion is only one choice of many. Hence, p-adic cameras

$$\kappa\colon \mathbb{P}^3(\mathbb{Q}_p) \to \mathbb{P}^2(\mathbb{Q}_p)$$

could also be considered in order to use methods from p-adic geometry for camera computations.

In fact, since finite resolution prevents stereoscopic vision at arbitrarily large distance, it can become possible to shift coordinate systems in such a way that cameras are described by matrices whose entries are natural numbers, at least approximately. In the following subsection, we will show how this said approximation can be done for the 2-adic norm. Precisely the hierarchical method used in the previous subsection in 2D extends in a natural way to 3D in order to arrive at 2-adic camera model

$$\kappa\colon \mathbb{P}^3(\mathbb{Z}_2) \to \mathbb{P}^2(\mathbb{Z}_2).$$

4.3 How to interpret a 2-adic essential matrix

Assume that through correspondences between some points $u \in I$, $u' \in I'$ in 2-adic camera images $I, I \cong \mathbb{P}^2(\mathbb{Z}_2)$, a 2-adic essential matrix $E \in \mathbb{Z}_2^{3\times 3}$ has been produced. This matrix is the limit of matrices $E_\nu \in \mathcal{Z}_{2^\nu}^{3\times 3}$. Through its factorisation

$$E_\nu = T_\nu R_\nu$$

into a skew-symmetric matrix T_ν and a rotation R_ν it allows to determine a point $U_\nu \in \mathbb{P}^3(\mathcal{Z}_{2^\nu})$ of which the two cameras obtain the image points $u_\nu, u'_\nu \mod 2^\nu$. These are related through E_ν:

$$u_\nu^T \cdot E_\nu \cdot u'_\nu = 0.$$

The U_ν converge 2-adically to a point $U \in \mathbb{P}^3(\mathbb{Z}_2)$, and the cameras

$$\kappa, \kappa'\colon \mathbb{P}^3(\mathbb{Z}_2) \to \mathbb{P}^2(\mathbb{Z}_2)$$

yield $\kappa(U) = u, \kappa'(U) = u'$. In the same way as u_ν is an approximation of pixel u at resolution ν, the 3D-point U_ν is an approximation of voxel U at resolution ν, because we can apply the map ι_2 in precisely the same manner also to three dimensions to obtain the inclusion $\mathbb{P}^3(\mathbb{Z}_2) \to [0,2)^3$.

5 Reconstruction from point correspondences

We will in the following content ourselves with solving the reconstruction problem with two calibrated cameras by finding (possible candidates for) essential matrices in the 2-adic setting.

Given n correspondences, we arrive at the system of linear equations

$$u_i^T E\, u_i', \qquad i = 1, \dots, n. \tag{3}$$

For $n \leq 9$, we can use the Hensel lifting method to solve these linear equations, written as

$$Ax = 0 \tag{4}$$

with coefficient matrix $A \in \mathbb{Z}_2^{n\times 9}$. Due to finite resolution and our coding method, the matrix entries lie in some set $\{0, \dots, 2^\nu - 1\}$. Hence, the matrix does not change its shape modulo 2^ν.

Assume that we are given a basis $B = (b_1, \dots, b_m)$ in normal form of the solution space of (4) modulo 2. We will interpret B as an ordered set, but also as a matrix whose columns are the vectors $b_1, \dots, b_m$.

Theorem 2 (Linear Hensel's lemma) *If the rank m of $A \mod 2$ equals $\operatorname{Rank}(A)$, then B lifts to a set of linearly independent solutions of (4) with natural numbers as entries.*

Proof The linear equation (4) is equivalent to a linear equation in staircase normal form. This normal form can be viewed as a system of $\operatorname{Rank}(A)$ equations in the 9 unknowns. By assumption, $m = \operatorname{Rank}(A)$. Hence, Theorem 1 applies, and B lifts to a set of m solution vectors

$$\tilde{B} = \left\{\tilde{b}_1, \dots, \tilde{b}_m\right\}$$

to (4).

Let us re-examine the proof of Theorem 1 in order to see that $\tilde{B}$ is linearly independent. First, notice that B is read off the staircase normal form of $A \mod 2$. This means that each $b \in B$ has in some row j_b entry 1, and below it all entries are zero. Further, all other entries of row j_b of matrix B are zero. The sequence $(j_b)_{b\in B}$ is strictly increasing with respect to the order of occurrence of b in B. Now, a lift of b to $b^{(k)}$ modulo 2^k has the property that the entry in row j_b is odd, and all entries below are even. Likewise, all other entries of row j_b in a lift of B to $B^{(k)}$ modulo 2^k are even. This description shows that the rows given by the sequence (j_b) form a submatrix of $B^{(k)}$ having rank m. Since this holds true for all $k > 0$, it follows that $\tilde{B}$ is linearly independent.

The lifting process stops after a finite number of steps, because of our initial assumptions.

5.1 Reconstruction from 8 points

Assume now that the rank of $A \mod 2$ be n, and that a basis $B = \{b_1, \dots, b_{9-n}\}$ of the solution space be given in normal form. Then, by Theorem 2, finding a lift to $\mathcal{Z}_{2^\nu}$ of B yields a basis of the solution space of (4).

As an application, we obtain a 2-adic version of the 8-point algorithm of [11], simply by setting $n = 8$. However, in the same way as its original, this ignores the rank constraint $\det(E) = 0$ for the essential matrix. Hence, we obtain the result:

Theorem 3 *Under the assumptions above for the matrix $A \in \mathbb{Z}_2^{8\times 9}$, the corresponding 8-point problem has a unique solution E, if additionally $\operatorname{Rank}(E) = 2$. If it can be assumed that $A \in \mathcal{Z}_{2^N}$, then E is computed after $N - 1$ iterations from $E \mod 2$.*

Proof Since $\operatorname{Rank}(A \mod 2) = 8$, there is one solution basis vector $b \in \mathcal{Z}^9$. By Theorem 2, it lifts to a non-trivial solution e of $Ax = 0$. If the the matrix E corresponding to e is of rank 2, it is the unique solution to the 8-point problem. The last assertion is an immediate consequence of the proof of Theorem 2.

5.2 Reconstruction from 7 points

The 7-point method by [6,9] yields 7 linear constraints (3) plus the cubic constraint

$$\det(E) = 0. \tag{5}$$

Let us write that system of equations as $f(X) = 0$. If the 7 points are sufficiently in general position, then the rank of $D_f(x) \mod 2$ is 8 for some solution x modulo 2:

$$f(x) \equiv 0 \mod 2, \tag{6}$$

and we can lift to a 2-adic solution. The reason is:

Lemma 1 *For $h := \det(E)$ it holds true that*

$$D_h(e) \not\equiv 0 \mod 2, \tag{7}$$

if $e \in \mathbb{Z}_2^{3\times 3}$ is a sufficiently general instance of E.

Proof The polynomial h is of degree three in the variables given by the entries of E. Let x be such an entry. Then $h_x := \frac{\partial h}{\partial x}$ is a polynomial of degree 2. Modulo 2 only those terms of h_x vanish which are of the form $2xy$ for some other variable $y \neq x$. Since $h = \det(E)$ is a square-free polynomial, this can never happen.

In 2-adic analytic geometry, the conditions (6) and (7) define an open subset in the space of all 3×3-matrices with entries in $\mathbb{Z}_2$. This can be seen easily by translating (6) to the inequality

$$|f(x)|_2 < 1,$$

and dealing similarly with (7).

We can go further and derive explicit conditions for the existence of $\mathbb{Z}_2$-rational solutions. Namely, write down the 1-parameter solution of the linear equations as

$$E = xE_1 + (1-x)E_2,$$

and obtain a polynomial

$$\det(E) = h(x) = ax^3 + bx^2 + cx + d$$

of degree 3. Now consider $h(x) \mod 2$. In case $d \equiv 0 \mod 2$, zero is a simple zero in $\mathbb{F}_2$ if and only if

$$c \equiv 1 \mod 2. \tag{8}$$

In case $d \equiv 1 \mod 2$, one is a simple zero in $\mathbb{F}_2$ if and only if

$$b \equiv c \equiv 0 \mod 2, \tag{9}$$

in case $a \equiv 1 \mod 2$, and

$$b \equiv 0, \quad c \equiv 1 \mod 2 \tag{10}$$

otherwise.

Theorem 4 *The 7-point problem has a $\mathbb{Z}_2$-rational solution for many choices of 7 point correspondences. Concretely, this is the case if (8), (9), or (10) hold true in their respective cases, together with the requirement that the rank of E be precisely 2.*

Proof Solve the 7-point problem modulo 2, and lift whenever possible as discussed above.

Remark 1 *Due to non-linearity of the constraints, we cannot expect anymore to be able to lift solutions modulo 2 to solutions which are defined over the natural numbers. In other words, generally, the iteration will never stop unless an order of resolution is specified. Then, a 2-adically approximate solution will be obtained. This is not different from the classical situation over the real numbers.*

5.3 Solving the 5-point non-linear equations

Given five corresponding pairs of image points yields (3) with $n = 5$. If the rank of the corresponding matrix is five, the general solution can be written as

$$E = xE_1 + yE_2 + zE_3 + wE_4. \tag{11}$$

Inserting this into the trace condition [4]

$$2 \cdot EE^TE - \mathrm{Trace}(EE^T) \cdot E = 0 \tag{12}$$

yields 9 cubic equations in four variables. We follow the easily understandable method of elimination via hidden variables used by [10], and obtain the linear equation

$$C(z) \cdot X = 0,$$

after setting $w = 1$. Here,

$$X = (x^3, y^3, x^2y, xy^2, x^2, y^2, xy, x, 1)$$

is the vector of monomials, and the entries of $C(z)$ are polynomials in z. Now, one seeks z such that

$$\det(C(z)) = 0. \tag{13}$$

The left hand side turns out to be a polynomial of degree 10.

We will use the same lax formulation of our theorem as for the 7-point problem, but will be more specific in the proof:

Theorem 5 *There exists a $\mathbb{Z}_2$-rational solution to the 5-point problem for many choices of 5 corresponding pairs of points.*

Proof Let

$$g(z) := \det(C(z)) = \sum_{i=0}^{10} a_i z^i.$$

We consider two cases.

First, assume $a_0 \equiv 0 \mod 2$. In this case, $g(0) \equiv 0 \mod 2$, and $z = 0$ is a simple zero modulo 2 if and only if $a_1 \equiv 1 \mod 2$.

Secondly, if $a_1 \equiv 1 \mod 2$, then $f(1) \equiv 0 \mod 2$ if and only if the number of odd coefficients in $g(z)$ is even. Since

$$f'(z) \equiv a_9z^8 + a_7z^6 + a_5z^4 + a_3z^2 + a_1 \mod 2,$$

1 is a simple zero modulo 2 if and only if in addition the number of odd coefficients in $f'(z)$ is odd.

Remark 2 *Hensel's lemma can be interpreted in this context as a p-adic stability result. Namely, assume that, due to correspondence error, a given choice of n points yields perturbed equations*

$$f(x) + \epsilon(x) = 0,$$

where $f(x)$ contain the "true" coefficients perturbed by some noise coefficients contained in $\epsilon(x)$ with 2-adic maximum norm

$$\|\epsilon\|_2 = \max\{|\epsilon_i|_2 \mid i = 1, \dots, m\} \leq 2^{-N},$$

where m is the number of noise coefficients ϵ_i. This means that the noise coefficients satisfy the congruence

$$\epsilon_i \equiv 0 \mod 2^N.$$

Hence, the first N iterations of Hensel lifting will lead to identical approximations to the solution of the unperturbed equations

$$f(x) = 0.$$

As the 2-adic encoding comes from interval subdivisions, this observation implies a greater stability in comparison with the classical approach over the real numbers. In particular, the existence of a liftable solution modulo 2 is not affected by perturbations ϵ with $\|\epsilon\|_2 \leq \frac{1}{2}$. This is definitively in contrast to the situation over the real numbers, where small perturbations in coefficients can drastically change the number of real solutions. That issue is addressed for the 5-point relative pose problem e.g. in [1].

6 Conclusion

An encoding scheme for image pixels through hierarchical interval subdivision is proposed. This allows a 2-adic encoding of pixels driven by geometry. As an application to stereo vision, the 8-, 7- and 5-point equations are formulated with coefficients from the ring $\mathbb{Z}_2$ of 2-adic integers. These polynomial equations are solved using some multivariate forms of Hensel's lemma. The essential matrices are obtained in the form of sequences of matrices modulo powers of 2, corresponding to 2-adic approximations to the exact solutions in $\mathbb{Z}_2$, whenever these exist. Conditions on coefficients of equations modulo 2 implying the existence of $\mathbb{Z}_2$-rational solutions are derived. One feature of the hierarchical encoding is that the number of iterations in solving the linear parts of the equations is logarithmic in the number of pixels. Also the precision in the solution of the non-linear equations is directly related to the desired resolution in 3D. An immediate consequence of Hensel's lemma is that p-adically small perturbations of the equations do not affect the first approximations to their solution. Further, the existence of liftable solutions is not affected by relatively large perturbations. This indicates a greater computational benefit from the 2-adic approach compared to the classical approach using computational complex algebraic geometry before discarding non-real solutions.

Acknowledgements Sven Wursthorn is thanked for the introduction into this fascinating topic, and Boris Jutzi for multiple fruitful discussions.

References

1. D. Batra, B. Nabbe, M. Hebert. An Alternative Formulation for the Five Point Relative Pose Problem. IEEE Workshop on Motion and Video Computing (2007)
2. J. Benois-Pineau, A.Yu. Khrennikov, N.V. Kotovich. Segmentation of images in p-adic and Euclidean metrics. Dokl. Math., 64, 450–455 (2001)
3. P.E. Bradley. Degenerating families of dendrograms. J. Classification, 25, 27–42 (2008)
4. O.D. Faugeras and S. Maybank. Motion from point matches: multiplicity of solutions. IJCV, 4, 225–246 (1990)
5. F.Q. Gouvêa. p-adic numbers: an introduction (2nd ed.). Springer (2000). ISBN 3540629114
6. R. Hartley. Projective reconstruction and invariants from multiple images. T-PAMI, 16, 1036–1040 (1994)
7. R. Hartley and A. Zisserman. Multiple view geometry in computer vision. Cambridge University Press (2000)
8. K. Hensel. Über eine neue Begründung der Theorie der algebraischen Zahlen. Jahresbericht der Deutschen Mathematiker-Vereinigung 6, 83-88 (1897)
9. T.S. Huang and A.N. Netravali. Motion and structure from feature correspondence: a review. Proc. IEEE, 82, 252–268 (1994)
10. H. Li and R. Hartley. Five-point motion estimation made easy. ICPR 2006, 630–633 (2006)
11. H.C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. Nature, 293, 133–135 (1981)
12. F. Murtagh. On ultrametricity, data coding, and computation. J. Classification, 21, 167–184 (2004)
13. D. Nistér. An efficient solution to the five-point relative pose problem. IEEE T-PAMI, 26, 756–770 (2004)
14. H. Stewénius. Gröbner basis methods for minimal problems in computer vision. Dissertation, Lund, Sweden (2005)
15. R. Sturm. Das Problem der Projektivität und seine Anwendung auf die Flächen zweiten Grades. Math. Ann., 1, 533–574 (1869)
16. H. Stewénius, C. Engels, D. Nistér. Recent developments on direct relative orientation. ISPRS J. Photogrammetry and Remote Sensing, 60, 284–294 (2006)
17. Z. Zhang. Determining the epipolar geometry and its uncertainty: a review. IJCV, 27, 161–198 (1998)