Fitting a Homography to Observations of Its Differential

Max Dunitz, Ismail Bencheikh, Marie d'Autume, Enric Meinhardt-Llopis, Marc Pic, Gabriele Facciolo, Pablo Musé

Fitting a Homography to Observations of Its Differential

Ismail BENCHEIKH[1,2,*] Max DUNITZ[1,2,*] Marie D'AUTUME[1] Enric MEINHARDT-LLOPIS[1] Marc PIC[2]
Gabriele FACCIOLO[1] Pablo MUSÉ[1,3]

[1]Centre Borelli, Université Paris-Saclay, Université Paris Cité, ENS Paris-Saclay, CNRS, SSA, INSERM
4, avenue des sciences, 91190 Gif-sur-Yvette, France

[2]Advanced Track and Trace, 75 avenue Victor Hugo, 92500 Rueil-Malmaison, France

[3]IIE, Facultad de Ingeniería, Universidad de la República, Av. Julio Herrera y Reissig 565, 11300 Montevideo, Uruguay

[*]These authors contributed equally to this work.

Résumé – Cet article propose une méthode d'estimation d'homographie basée sur des textures auto-rectifiantes, exploitant l'autocorrélation pour se passer de points de repère explicites. Les pics d'autocorrélation locaux permettent d'estimer les transformations affines, puis de reconstruire l'homographie globale. La solution, robuste aux occultations et discrète, s'adapte à des usages comme la stéganographie ou la vision par ordinateur.

Abstract – This paper proposes a homography estimation method using self-rectifying textures, leveraging autocorrelation to avoid explicit key points. Local autocorrelation peaks are used to estimate affine transformations, then reconstruct the global homography. The approach, robust to occlusions and unobtrusive, suits applications such as steganography or computer vision.

1 Introduction

Traditional fiducial markers rectify images using pixel landmarks. For instance, QR code readers often use the pixel coordinates of the code's four corners to identify the rectifying homography, using the three "eyes" to break quadrilateral symmetry [12]. Marker boards (checkerboard patterns of high-contrast fiducial markers) supply many corners whose coordinates can be identified using simple low-level algorithms, enabling the rectification of more complex deformations [6].

Marker boards are robust to tampering and occlusion. But they are conspicuous and unsuited for many applications with aesthetic constraints. For this reason, stealth self-rectifying textures were introduced [9]. These textures encode pose information with landmarks not in the image itself but rather as peaks in its autocorrelation (or power spectrum) image–permitting their inconspicuousness. Their robustness is assured by the fact that each patch contains the landmarks.

Such textures can be rectified patchwise, as described in Section 2. Complex deformations of a texture can be modeled as affine in small patches. As each affine transformation linearly deforms the fiducial markers (peak locations), its linear part can be recovered from the defomed landmarks–but not its phase. Extracting the phase requires access to the template texture (rather than simply the pixel locations of autocorrelation landmarks in the rectified image) and one costly phase correlation operation per patch.

These phase correlations are, in many situations, unnecessary. Rectification (modulo a translation) can be accomplished without access to the template image, only this limited knowledge of its statistical properties–the rectified peak locations. One might work nonparametrically, by integrating the differentials directly or viewing the ensemble of differential observations as an integrable polyvector field [2, 5, 8].

In this article, we fit a parametric model–namely, a rectifying homography–to observations of the local differentials of the deformation (revealed by autocorrelation landmarks in patches). In this way, a self-rectifying texture to which we need not have direct access can be rectified modulo translation. A known texture acquired with homographic deformation can be fully inteiinterinrectified with just one costly phase correlation.

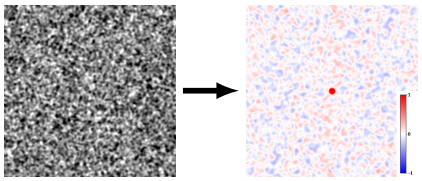
2 Texture Generation and Patchwise Deformation Estimation

We create inconspicuous textures with autocorrelation landmarks by superimposing three shifted copies of a base texture, creating a "fundamental hexagon" of autocorrelation peaks. Other sets of peaks may be used for deformation differential estimation and other methods of generating a texture exhibiting peaks at the desired locations can be envisioned. Indeed, the ill-posed phase retrieval inverse problem from homometric point sets [10] and the stability of peak positions under common operations offer extensive design freedom.

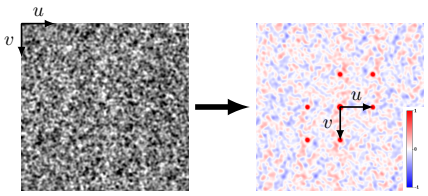
Figure 1 demonstrates how to generate a simple self-rectifying texture whose rectified patches all have prominent autocorrelation peaks at fixed locations due to the known offsets with which the texture was generated. Let $u, v \in \mathbb{R}^2$ be offsets, not necessarily orthogonal, though not collinear. A texture [1] $T : \mathbb{R}^2 \to \mathbb{R}$ can be made self-rectifying by summing shifted copies $T_{\text{sr}} = T + T(\cdot - u) + T(\cdot - v)$. These two offsets generate peaks in the autocorrelation image at the origin and at $H_{\text{rect}} = \{\pm u, \pm v, \pm(u - v)\}$. The only constraint on the base texture T is that its autocorrelation peak at 0 be prominent.

This "fundamental hexagon" H_{rect} is deformed linearly by affine maps, exposing the linear part of the affine map with which the texture is acquired. Let $\mathrm{A} \in \mathrm{Aff}(2, \mathbb{R})$ be an affine

1. To simplify the discussion, we view textures as infinite and real-valued, though this analysis can be adapted to rectangular domains with cyclic shifts and quantized images with little difference.



(a) A texture with prominent autocorrelation peak at 0.



(b) The sum of three images at pairwise lags $\pm u, \pm v$, and $\pm(u-v)$ gives rise to a fundamental hexagon of peaks, which can be used to rectify affine deformations up to translation.

Figure 1 – Texture overlays yield peaks in the autocorrelation image. Three superimposed copies of the texture of **(a)**, with relative lags u and v, produce the fundamental hexagon associated with the six nonzero pairwise lags, visible in the autocorrelation of the texture in **(b)**.

transformation. Thus, for some $\mathbf{A} \in \mathrm{GL}(2,\mathbb{R})$ and $b \in \mathbb{R}^2$, A is the map $x \mapsto \mathbf{A}x + b$. Denoting by R the autocorrelation operator, we observe that $\mathrm{R}(\mathrm{T} \circ \mathrm{A}) \propto \mathrm{R}(\mathrm{T}) \circ \mathbf{A}$. The locations of the fundamental hexagon in a patch in the deformed texture are therefore situated at $H_{\text{def}} = \{\pm\mathbf{A}u, \pm\mathbf{A}v, \pm\mathbf{A}(u-v)\}$. After matching the six peak locations H_{rect} to their deformed positions in H_{def}, we can take $\mathbf{A}$ to be the least-squares estimate of the overdetermined linear system $H_{\text{def}} = \mathbf{A}H_{\text{rect}}$.

With knowledge of the texture, the assignment problem is readily solved using phase correlation. Without access to the texture, one can distinguish peaks by weighting the shifted copies so that $T = T + \alpha_u \cdot T(\cdot - u) + \alpha_v \cdot T(\cdot - v)$ for some $\alpha_u, \alpha_v \in \mathbb{R}$. Often, the six peaks can also be assigned simply by considering their sums and differences. Let $\{p_1, p_2\} \subseteq H_{\text{def}}$ be noncollinear. Then $\{\pm p_1 \pm p_2\} \subseteq \{\pm\mathbf{A}u \pm \mathbf{A}v\}$ only when $\{p_1, p_2\} \subseteq \{\pm\mathbf{A}u, \pm\mathbf{A}v\}$. Prior constraints on $\mathbf{A}$ or asymmetric texture motifs can complete the assignment.

From observations $\{\mathbf{A}_i\}_{i=1}^n$ of the linear part of the local affine approximation to an arbitrary deformation H, computed on patches $\{P_i\}_{i=1}^n$ centered at $\{y_i\}_{i=1}^n$ in the deformed image, we can rectify the texture patchwise. The phase correlation reveals the translation parameter b_i of the local affine map: $b_i = \operatorname{argmax}\Big((\widehat{T}_{\text{sr}} \circ \widehat{P}_i^*)/|\widehat{T}_{\text{sr}} \circ \widehat{P}_i^*|\Big)^\vee$. Thus, the first-order Taylor approximation of the deformation, valid near $x_i = \mathrm{H}^{-1}(y_i) = \mathbf{A}_i^{-1}(y_i - b_i)$, is $h \mapsto y_i + \mathbf{A}_i(h - x_i) = \mathbf{A}_i h + b_i$; its inverse $h \mapsto x_i + \mathbf{A}_i^{-1}(h - y_i)$ rectifies the patch.

When the deformation H is a homography, we can use these Taylor approximations to compute the global rectifying homography H^{-1}. The translation parameters b_i give us coordinate pairs $(\mathbf{A}^{-1}(y_i - b_i), y_i)$ in the rectified and deformed images, respectively, facilitating traditional correspondence-based homography estimation. This approach nevertheless requires n phase correlations and access to the texture. Instead, we seek a fully correspondence-free estimation of the rectifying homography, modulo the translation parameters, without access to the texture, only knowledge of its fundamental shifts u and v.

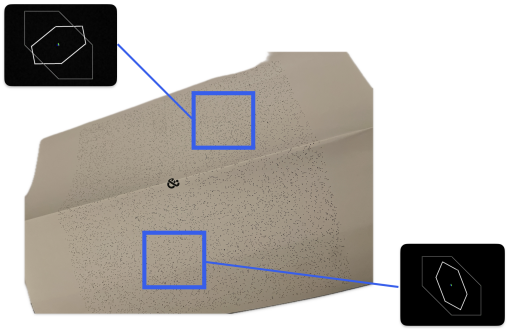


Figure 2 – A self-rectifying texture covering a rectangular prism and the fundamental hexagons of two patches.

3 Jacobian-based Homography Fitting

An isomorphism of the real projective plane $\mathbb{P}_2$ (the set of lines through the origin in $\mathbb{R}^3$) is called a homography. Homographies form the projective linear group $\mathrm{H} \in \mathrm{PGL}(3,\mathbb{R}) \cong \mathrm{GL}(3,\mathbb{R})/\mathbb{R}^\times$. These nonlinear planar deformations $\mathrm{H} : \mathbb{R}^2 \to \mathbb{R}^2$ can be represented by an equivalence class of 3×3 invertible matrices $\{\eta\mathbf{M}_{\mathrm{H}} \mid \eta \neq 0\}$ that operate on homogeneous coordinates. We can always choose as class representer the unique matrix $\mathbf{M}_{\mathrm{H}}$ with $\det \mathbf{M}_{\mathrm{H}} = 1$. Here

$$\mathbf{M}_{\mathrm{H}} = \begin{pmatrix} h_{1,1} & h_{1,2} & h_{1,3} \\ h_{2,1} & h_{2,2} & h_{2,3} \\ h_{3,1} & h_{3,2} & h_{3,3} \end{pmatrix} = \begin{pmatrix} — & h_1 & — \\ — & h_2 & — \\ — & h_3 & — \end{pmatrix} \tag{1}$$

corresponds to the planar deformation $\mathrm{H} : \mathbb{R}^2 \to \mathbb{R}^2$:

$$\begin{pmatrix} x & y \end{pmatrix}^T \mapsto \frac{1}{\langle h_3, \begin{pmatrix} x & y & 1 \end{pmatrix}^T \rangle_{\mathbb{R}^3}} \begin{pmatrix} \langle h_1, \begin{pmatrix} x & y & 1 \end{pmatrix}^T \rangle_{\mathbb{R}^3} \\ \langle h_2, \begin{pmatrix} x & y & 1 \end{pmatrix}^T \rangle_{\mathbb{R}^3} \end{pmatrix}. \tag{2}$$

The Jacobian matrix of H, parameterized by $\mathbf{M}_{\mathrm{H}}$ in (1), evaluated at some $y_i \in \mathbb{R}^2$, then, is as follows:

$$\nabla_{x_i} \mathrm{H} = \frac{\begin{pmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \end{pmatrix} - (\mathrm{H}x_i) \otimes \begin{pmatrix} h_{3,1} & h_{3,2} \end{pmatrix}}{\left\langle h_3, \begin{pmatrix} x_i[1] & x_i[2] & 1 \end{pmatrix}^T \right\rangle_{\mathbb{R}^3}}. \tag{3}$$

In each patch P_i, centered at y_i, we can observe the deformed version H_{def}^i of the fundamental hexagon of peaks H_{rect}. The pixel locations of these fiducial markers in the autocorrelation of P_i exposes the differential of the deforming homography H, evaluated at $x_i = \mathrm{H}^{-1}(y_i) = \mathbf{A}_i^{-1}(y_i - b_i)$. This matrix $\mathbf{A}_i$ can be recovered by solving $H_{\text{def}}^i = \mathbf{A}_i H_{\text{rect}}$. We wish to estimate H from the ensemble of observations of $\nabla_{\mathrm{H}^{-1}(y_i)}\mathrm{H}$.

Problem 1: Homography Recovery (up to Translation)

Given observations $\{\mathbf{A}_i\}_{i=1}^n$ of the Jacobian matrices $\mathbf{A}_i = \nabla_{\mathrm{H}^{-1}(y_i)}\mathrm{H}$ of a planar deformation H computed on patches of deformed texture with centers $\{y_i\}_{i=1}^n \subseteq \mathbb{R}^2$, search for a homography behaving as H:

$$\mathrm{H}^* = \operatorname*{argmin}_{\mathrm{H} \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^n \operatorname{loss}\left(\nabla_{\mathrm{H}^{-1}(y_i)}\mathrm{H}, \mathbf{A}_i\right). \tag{4}$$

Or seek the rectifying homography $\mathrm{G} = \mathrm{H}^{-1}$:

$$\mathrm{G}^* = \operatorname*{argmin}_{\mathrm{G} \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^n \operatorname{loss}\left(\nabla_{y_i}\mathrm{G}, \mathbf{A}_i^{-1}\right). \tag{5}$$

With hypothesis space $\mathcal{H} = \mathrm{PGL}(3, \mathbb{R})$, the problems (4)-(5) are ill-posed. Indeed, the differentials are (and thus the loss is) invariant to a translation τ since $\nabla_x \tau = \mathbf{I}$ for all $x \in \mathbb{R}^2$:

$$\begin{aligned}\nabla_{(\mathrm{H}\circ\tau)^{-1}(y_i)} (\mathrm{H} \circ \tau) &= \nabla_{(\tau\circ\tau^{-1}\circ\mathrm{H}^{-1})(y_i)}\mathrm{H} \cdot \nabla_{(\mathrm{H}\circ\tau)^{-1}(y_i)}\tau \\ &= \nabla_{\mathrm{H}^{-1}(y_i)}\mathrm{H} \cdot \mathbf{I} = \nabla_{\mathrm{H}^{-1}(y_i)}\mathrm{H}.\end{aligned}$$

We choose a loss that punishes equally the unfaithfulness of each entry of the fit Jacobian matrices to the observations. (Metrics of average rectification error in pixels over a domain, as in [7], can sidestep the unbalanced impact of each entry.)

In our experiments, we solve problem 5 with Frobenius-norm loss: $\mathrm{loss}(a, b) = ||a - b||_F$. Problem (5), without the inverse homography in the objective, assumes a simpler form than (4). (Indeed, if we square each patchwise loss term and multiply through the squared denominators, we obtain a quartic polynomial objective–though nonconvex, in eight variables. But this renders the penalty more sensitive to outliers without computational benefit.) The loss is robust and its dependence on the variables is not mediated through an inverse.

Our inverse problem with non-convex loss requires an iterative approach. In these experiments, we optimize over the Euclidean space $\mathbb{R}^6$ to learn six of the eight homography parameters–an approach we justify in the following section. Due to our small datasets, we considered nonstochastic gradient descent with various settings of line search; however, convergence was too slow for our application. We found acceptable performance using the trust region method as implemented in `pymanopt` [1, 11] using the manifold $\mathbb{R}^6$.

3.1 Euclidean optimization on $\mathbb{R}^6$

By the quotient manifold theorem [3], $\mathrm{PGL}(3, \mathbb{R})$ is a quotient manifold. It is an easy matter to show that $\mathrm{PGL}(3, \mathbb{R})$ is isomorphic to the special linear group $\mathrm{SL}(3, \mathbb{R})$ as a Lie group and manifold: indeed, for n odd, we have that $\mathrm{GL}(n, \mathbb{R}) = \mathbb{R}^\times \times \mathrm{SL}(n, \mathbb{R})$. Consequently, the Lie algebra of $\mathrm{PGL}(3, \mathbb{R})$ is $\mathfrak{sl}(3, \mathbb{R})$, the traceless 3×3 matrices. Optimizing over $\mathrm{GL}(3, \mathbb{R})$ makes little sense as it is dense in Euclidean space $\mathbb{R}^{3\times3}$. In practice, we find that optimizing over $\mathrm{SL}(3, \mathbb{R})$ does not speed up first-order methods for our loss; gradient descent over $\mathbb{R}^9$ has no trouble converging to an invertible matrix, interpretable as a homography. In hopes of speeding up convergence by passing from the usual Euclidean optimization over $\mathbb{R}^9$ to $\mathbb{R}^6$, we consider the action of translations on $\mathrm{SL}(3, \mathbb{R})$.

Let $\mathrm{T}(2, \mathbb{R})$ be the group of translations $\begin{pmatrix} x & y \end{pmatrix}^T \mapsto \begin{pmatrix} x + a & y + b \end{pmatrix}^T$ parameterized by $\begin{pmatrix} a & b \end{pmatrix}^T \in \mathbb{R}^2$. As a homography matrix applied to homogeneous coordinates, each element of $\mathrm{T}(2, \mathbb{R})$ is in $\mathrm{SL}(3, \mathbb{R})$:

$$\mathrm{T}(2, \mathbb{R}) = \left\{ \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \;\middle|\; \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 \right\} \cong \mathbb{R}^2. \qquad (6)$$

A closed subgroup of $\mathrm{SL}(3, \mathbb{R})$, $\mathrm{T}(2, \mathbb{R})$ is a Lie group with Lie algebra spanned by two matrices $\mathbf{P}_1$ and $\mathbf{P}_2$, with the usual commutator bracket. Each matrix has only one nonzero element, in the top-right and middle-right entries (where a and b are), respectively. Since $\mathrm{T}(2, \mathbb{R})$ is not a normal subgroup of $\mathrm{SL}(3, \mathbb{R})$, the quotient $\mathrm{SL}(3, \mathbb{R})/\mathrm{T}(2, \mathbb{R})$ is not a Lie group. However, by the quotient manifold theorem, this homogeneous space is a smooth manifold. It is covered by the union of three manifolds, one of which is the Lie group $\mathrm{A}_2(2, \mathbb{R})$ that consists of the transposes of matrices in the affine group $\mathrm{Aff}(2, \mathbb{R})$. The Lie algebra of $\mathrm{A}_2(2, \mathbb{R})$ is $\mathfrak{a}_2(2, \mathbb{R})$, a maximal Lie subalgebra of $\mathfrak{sl}(3, \mathbb{R})$, and spanned by the traceless matrices $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3, \mathbf{D}, \mathbf{R}_1$, and $\mathbf{R}_2$ of [4], Section 4. These matrices, combined with $\mathbf{P}_1 = \mathbf{R}_1^T$ and $\mathbf{P}_2 = \mathbf{R}_2^T$, form an orthonormal basis for $\mathfrak{sl}(3, \mathbb{R})$. (Swapping $\mathbf{P}_1$ and $\mathbf{P}_2$ for $\mathbf{R}_1$ and $\mathbf{R}_2$, we recover the other maximal subalgebra of $\mathfrak{sl}(3, \mathbb{R})$, that of $\mathrm{Aff}(2, \mathbb{R})$; see [4], Lemma 18.)

$\mathrm{T}(2, \mathbb{R})$ is a normal subgroup of $\mathrm{Aff}(2, \mathbb{R})$. Similarly, the Lie group $\{\mathbf{M}^T \mid \mathbf{M} \in \mathrm{T}(2, \mathbb{R})\}$ of matrix transposes of $\mathrm{T}(2, \mathbb{R})$ is a normal subgroup of $\mathrm{A}_2(2, \mathbb{R})$. Identifying translations with $\mathbb{R}^2$, it is easy to see that $\mathrm{A}_2(2, \mathbb{R}) \cong \mathbb{R}^2 \rtimes \mathrm{GL}(2, \mathbb{R})$–and is dense in $\mathbb{R}^6$.

We want to represent elements of $\mathrm{SL}(3, \mathbb{R})$ modulo translations by an element in $\mathrm{A}_2(2, \mathbb{R})$. The action $(\mathbf{M}, \mathbf{T}) \mapsto \mathbf{MT}$ of multiplying a matrix $\mathbf{M} \in \mathrm{SL}(3, \mathbb{R})$ on the right by a translation matrix $\mathbf{T} \in \mathrm{T}(2, \mathbb{R})$ partitions $\mathrm{SL}(3, \mathbb{R})$ into orbits of dimension 2 consisting of matrices related by translations: $\mathbf{M}_1 \sim \mathbf{M}_2 \iff \mathbf{M}_2 = \mathbf{M}_1\mathbf{T}_{12}$ for some $\mathbf{T}_{12} \in \mathrm{T}(2, \mathbb{R})$.

Let $\mathbf{M} \in \mathrm{SL}(3, \mathbb{R})$ and consider its orbit $\{\mathbf{MT} \,|\, \mathbf{T} \in \mathrm{T}(2, \mathbb{R})\}$ under the action of $\mathrm{T}(2, \mathbb{R})$. If the first 2×2 principal submatrix of $\mathbf{M}$ happens to be invertible, there is a unique choice of $\mathbf{T} \in \mathrm{T}(2, \mathbb{R})$ of the form (6) such that $\mathbf{M} = \mathbf{AT}$, where $\mathbf{A} \in \mathrm{A}_2(2, \mathbb{R})$. We find the orbit representative of $\mathbf{M} = (m_{i,j})_{i,j=1}^3$ by "translating to zero" its top-right and middle-right entries: $[\mathbf{M}] = \mathbf{MT}^{-1} \in \mathrm{A}_2(2, \mathbb{R})$, where

$$\mathbf{T}^{-1} = \begin{pmatrix} 1 & 0 & b_1 \\ 0 & 1 & b_2 \\ 0 & 0 & 1 \end{pmatrix} \text{ with } \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} m_{1,1} & m_{1,2} \\ m_{2,1} & m_{2,2} \end{pmatrix}^{-1} \begin{pmatrix} m_{1,3} \\ m_{2,3} \end{pmatrix}.$$

While $\mathbf{M}[1:2, 1:2]$ could be singular, at least one of the three 2×2 principal minors must be nonzero as $\mathbf{M} \in \mathrm{SL}(3, \mathbb{R})$. Thus, we can always permute two rows, multiply a row by -1 to preserve the unit determinant, and then "translate to zero." Any $\mathbf{M} \in \mathrm{SL}(3, \mathbb{R})$ thereby admits the decomposition $\mathbf{M} = \mathbf{AN}$, where $\mathbf{A} \in \mathrm{A}_2(2, \mathbb{R})$ and $\mathbf{N}$ takes one of three forms:

$$\mathbf{N}_1 = \begin{pmatrix} 1 & 0 & b_1 \\ 0 & 1 & b_2 \\ 0 & 0 & 1 \end{pmatrix}; \mathbf{N}_2 = \begin{pmatrix} -1 & c_1 & 0 \\ 0 & c_2 & 1 \\ 0 & 1 & 0 \end{pmatrix}; \mathbf{N}_3 = \begin{pmatrix} c_1 & 0 & -1 \\ c_2 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

The first matrix $\mathbf{N}_1$ is a translation; the row swaps of $\mathbf{N}_2$ and $\mathbf{N}_3$, in homogeneous coordinates, send the origin to infinity.

As the homographies we seek to rectify have invertible first 2×2 principal submatrix, $\mathrm{A}_2(2, \mathbb{R})$ is the relevant portion of $\mathrm{SL}(3, \mathbb{R})/\mathrm{T}(2, \mathbb{R})$. We choose the matrix scale so $\mathrm{A}_2(2, \mathbb{R})$ is dense in a six-dimensional affine subspace of $\mathbb{R}^9$:

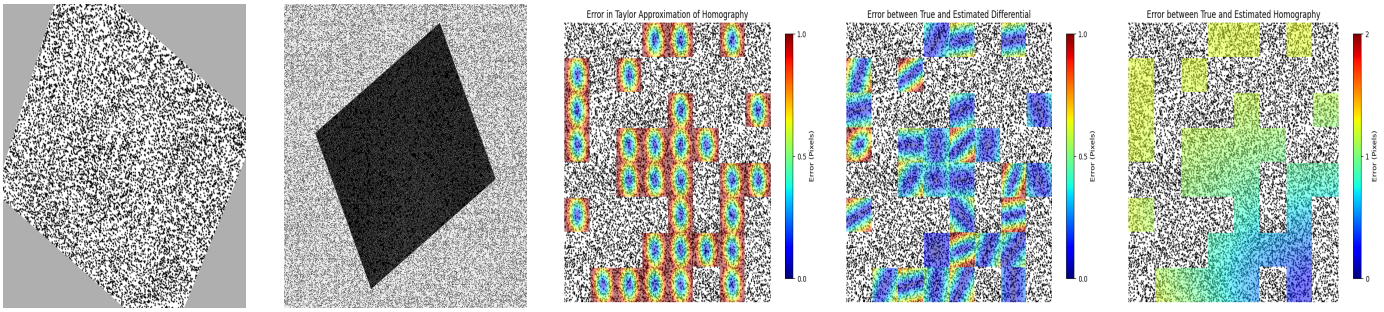
$$\mathrm{A}_2(2, \mathbb{R}) \cong \mathrm{A}_2'(2, \mathbb{R}) = \left\{ \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ x & y & 1 \end{pmatrix} \in \mathbb{R}^{3\times3} \;\middle|\; ad \neq bc \right\}.$$

4 Experiments

We deform a (cropped) self-rectifying texture

$$T_{\mathrm{sr}} = T + T\left(\cdot - \begin{pmatrix} 50 & 0 \end{pmatrix}^T\right) + T\left(\cdot - \begin{pmatrix} 0 & 50 \end{pmatrix}^T\right),$$

where T is composed of a steganographic message (with 1s encoded 5×5 squares and 0s blank) using a known homography H, representative of a handheld camera view of the



(a) $T_{\text{sr}} \circ \text{H}$. (b) Patchwise rectified. (c) Taylor error. (d) Patchwise error. (e) Global error.

Figure 3 – **(a)**: example self-rectified texture T_{sr}, deformed by H. **(b)**: patchwise rectification of (a) using peaks extracted from the autocorrelation of each of the uniformly spaced 118×118 patches in the deformed image to estimate $\mathbf{A}_i$ and a phase correlation to extract the translation component of the first-order approximation of the rectifying homography $\text{G} = \text{H}^{-1}$ about y_i. **(c)-(e)** present a crop of the deformed texture (a), in each case with a different error shading overlaid on a selection of the 118×118 patches. **(c)**: the patchwise rectification error of the first-order Taylor approximation overlaid. (With known homography, patchwise rectification error stems only from Taylor-series truncation, not estimation of $\mathbf{A}_i$.) For each pixel in a selected patch with offset p relative to patch center y_i, $||\text{G}(y_i + p) - \text{G}y_i - (\nabla_{y_i}\text{G})p||_{\mathbb{R}^2}$ is shaded, clipped at 1 pixel. **(d)**: the rectification error $||\mathbf{A}_i^{-1}p - (\nabla_{y_i}\text{G})p||_{\mathbb{R}^2}$ when using the linear part of the rectifying map $\mathbf{A}_i^{-1}$, determined by the $\mathbf{A}_i$ estimated from fundamental hexaon movement, and the linear part of the first-order rectifying map calculated at the patch center $(\nabla_{\text{H}^{-1}y_i}\text{H})^{-1} = \nabla_{y_i}\text{G}$. Also clipped at 1 pixel. **(e)**: the error of the global homography fit to the patchwise differentials.

texture, in Figure 3a. From patches of size 118×118 in the deformed image, we compute the error of a first-order Taylor series expansion using the known gradient at the patch centers. A selection of these theoretical best errors are overlaid on a crop of the deformed texture in Figure 3c. Despite the large patch size, the error between the estimated and true differential in Frobenius norm is small (10^{-5} in most patches); measured in pixels, this norm error induces non-negligible patchwise rectification error (see Figs. 3b and 3d). From the differentials estimated on patches, a `pymanopt TrustRegions` optimization over $\mathbb{R}^6$ of the sum of Frobenius-norm patchwise loss finds an element of $\mathbb{R}^6$–and indeed of $\text{A}'_2(2, \mathbb{R})$–with loss on the order of 10^{-3}. Phase correlation is used to find the best translation. As Figures 3d-3e indicate, complete rectification of a homography using the homography fit to the patchwise differentials performs better than rectification patchwise–and requires just one phase correlation, saving far more than the added 50 ms `TrustRegions` optimization.

5 Conclusion

We have demonstrated the ability to learn a homography modulo translation from observations of its differential computed on patches using the properties of self-rectifying textures. This is at once more computationally efficient and more widely applicable than rectification by patch as it requires no phase correlation and thus no access to the texture, only the fiducial marker (the shifts that produce the fundamental hexagon). We leave as future work the rectification of more complex deformations, locally using homographies or using more complex parametric models, such as thin-plate splines.

References

[1] P-A Absil, Christopher G Baker, and Kyle A Gallivan. Trust-region methods on Riemannian manifolds. *Found. Comp. Math.*, 7:303–330, 2007.

[2] Mikhail Bessmeltsev and Justin Solomon. Vectorization of line drawings via polyvector fields. *ACM Trans. Graph. (TOG)*, 38(1):1–12, 2019.

[3] Nicolas Boumal. *An introduction to optimization on smooth manifolds*. Cambridge University Press, 2023.

[4] Yevhenii Yu Chapovskyi, Serhii D Koval, and Olha Zhur. Subalgebras of Lie algebras. Example of $\mathfrak{sl}(3, \mathbb{R})$ revisited. *arXiv preprint. arXiv:2403.02554*, 2024.

[5] Olga Diamanti, Amir Vaxman, Daniele Panozzo, and Olga Sorkine-Hornung. Integrable polyvector fields. *ACM Trans. Graph. (TOG)*, 34(4):1–12, 2015.

[6] Sergio Garrido-Jurado, Rafael Muñoz-Salinas, Francisco José Madrid-Cuevas, and Manuel Jesús Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6):2280–2292, 2014.

[7] Changsoo Je and Hyung-Min Park. Homographic p-norms: Metrics of homographic image transformation. *Signal Process. Image Commun.*, 39:185–201, 2015.

[8] David Palmer, David Bommes, and Justin Solomon. Algebraic representations for volumetric frame fields. *ACM Trans. Graph. (TOG)*, 39(2):1–17, 2020.

[9] Justin Picard, Jean-Pierre Massicot, Alain Foucou, and Zbigniew Sagan. Method and device for securing documents, United States Patent 0220364A1, Sep. 2, 2010.

[10] Marjorie Senechal. A point set puzzle revisited. *Eur. J. Comb.*, 29(8):1933–1944, 2008.

[11] James Townsend, Niklas Koep, and Sebastian Weichwald. Pymanopt: A Python toolbox for optimization on manifolds using automatic differentiation. *J. Mach. Learn. Res.*, 17(137):1–5, 2016.

[12] J Udvaros and L Szabó. Recognize and decode QR codes from images. In *Int. Conf. Control Decis. Inf. Technol. (CoDIT)*, pages 1393–1397. IEEE, 2024.