

Autocorrelation-based Fiducial Markers for Traceability

Ismail Bencheikh^{1,2,*} Max Dunitz^{1,2,*} Marie d'Autume¹ Enric Meinhardt-Llopis¹

Marc Pic² Gabriele Facciolo¹ Pablo Musé^{1,3}

¹Université Paris-Saclay, ENS Paris-Saclay, CNRS, Centre Borelli, France

²Advanced Track & Trace, France * Equal contributions.

³IIE, Facultad de Ingeniería, Universidad de la República, Uruguay

Abstract

Classical approaches to the rectification of a single image of a product, without stereo correspondences, require spatial landmarks. These landmarks, constructed from high-contrast elementary shapes that can be detected with simple algorithms, are highly conspicuous. To rectify complex deformations, one can use chessboard patterns of markers with elements that break quadrilateral symmetry, such as the three eyes of a QR code. However, these marker boards are even more conspicuous than a single marker. In traceability applications, only one site of marking is used, limiting the complexity of the surface on which it can be read, and exposing the mark to deidentification attacks for diversion of the product to a grey market. We introduce a method for constructing stealth and robust fiducial markers that can be displayed across a surface, limiting exposure to marker tampering for product deidentification. These markers, which we refer to as self-rectifying textures, can be used to rectify complex deformations by solving an inverse problem rather than relying on pixel correspondences of conspicuous landmarks. These stealth textures place fiducial markers in the autocorrelation of the image. In this way, crops of the deformed texture can be rectified using only these spatially invariant statistical properties. Affine transformations of an image correspond to linear transformations of the autocorrelation, without phase component. Exploiting this fact, self-rectifying textures enable the local estimation of the differential of a planar deformation by identifying landmarks in the autocorrelation image, such as peaks, whose locations in the fronto-parallel view of the texture are known. The translation component can be recovered independently via phase correlation. A rectifying map, modulo translations, can also be fit directly to local observations of the differential of the deformation, without access to the rectified texture or need for phase correlation. Self-rectifying textures can be used for communication, watermarking, authentication, surface identification, calibration,

and geometry processing.



Figure 1. A self-rectifying texture and a QR code are printed on a piece of paper folded across a table edge. This deformation prevents the QR code from being read. By contrast, the real-time self-rectifying texture reader (left) can acquire the deformation differential (left lower inset) from the small patch being processed (left upper inset); from this single patch, after affine rectification, the traceability information is easily decoded. A photograph of the QR code spread across two surface faces (center) cannot be read. However, the geometric information exposed by the self-rectifying texture easily allows rectification of a homography for each planar surface, allowing the QR code to be read.

1. Introduction

Trademark holders, inventory managers, consumer advocates, safety and environmental regulators, customs officials, central banks, and tax authorities seek better technologies and standards to ensure product traceability and supply chain transparency [14, 18, 37]. As physical goods come in many forms, markings containing product information must be easily readable, even on complex surfaces. Traceability markers are common sites of tampering as product deidentification is often necessary for diversion to a grey market. As a defense against tampering, robust traceability marking requires redundancy—and thus, due to aesthetic constraints, stealth.

Fiducial markers (see Fig. 2) generally consist of high-contrast images of large, elementary objects like squares, enabling detection by simple algorithms [5, 17, 21, 43, 46].

They reveal pose information via pixel locations of landmarks—often the corners of a square marker’s bounding box—along with elements that disrupt symmetries to distinguish the landmarks (such as the three “eyes” of QR code). Only the repetition of markers across a surface can endow landmark-based rectification methods with robustness to intense, localized occlusion due to deidentification. Such repetition often grants the ability to rectify planar deformations more complex than homographies. But repeated markers generally assume the form of chessboard-like arrays of high-contrast markers, known as marker boards [19]. We seek a configurable family of *inconspicuous* fiducial markers that, like marker boards, are geometrically informative and robust to deidentification occlusion.

Objectives:

- We introduce a family of fiducial markers that is
- *informative*: exposing local deformation differentials;
 - *inconspicuous*: covering product surfaces with markings, granting robustness to tampering and warp;
 - *template-free*: rectifying complex deformations up to translation without phase correlation; and
 - *configurable*: accomodating a variety of payloads.

1.1. Existing markers are unsuitable for traceability

Our area of focus—anti-diversion traceability—involves smaller payloads than typical 2D communication schemes (on the order of 8 message bytes, or 10^{19} messages), but much larger payloads than typical marker ID dictionaries. Even the standard “low payload” 25×25 module QR code version 2 carries 47 bytes, of which 19-37 are message bytes, depending on the error-correction-level setting [26]. This is far more than we require. On the other hand, marker systems for augmented reality applications, such as AR-ToolKit [10], AprilTags [38], and ArUco markers [19], contain IDs in a pre-defined dictionary, typically with on the order of 10^1 to 10^3 IDs per dictionary—far less than an eight-byte message (10^{19}). While CALTags could offer such combinatorial possibilities, they do not incorporate forward error correction (FEC); ARTags, which do, accommodate only 10^3 marker IDs.

Most fiducial markers have sufficient landmarks and asymmetry to rectify homographies. (With information to estimate normal translations, such as camera intrinsics, these markers—especially AprilTags and ArUco markers—are used by the robotics community to estimate the six degree-of-freedom pose of a camera relative to a planar surface [22, 27].) In the literature, however, only marker boards can rectify complex planar deformations induced by the acquisition of a 2D code printed on a complex surface or by camera lens warp. But these conspicuous marker boards, while useful for instrument calibration, are ill-suited for use

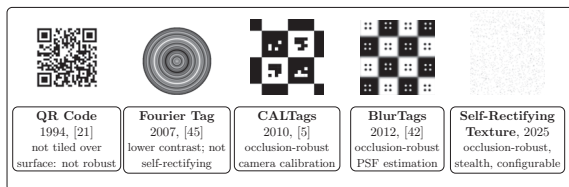


Figure 2. Example fiducial markers. The self-rectifying texture (right) is more discreet than the others in the literature. Most tags with sufficient asymmetry (all but Fourier Tag here) possess enough information to perform landmark-based rectification of a homography, but only conspicuous chessboard patterns repeated across a surface or the self-rectifying texture can reveal the more complex planar deformations encountered in product traceability applications and exhibit the necessary robustness to tampering. Like the QR code, it can be configured to carry arbitrary messages (rather than a fixed set of motifs from a small marker dictionary); its payload size, density, and robustness are highly adaptable, akin to QR code version and error-correction-level selection.

in commercial packaging.

While many markers offer some protection against occlusion, only a furtive marker that can cover a large surface is robust to deidentification attacks.

Traceability applications have varying two-dimensional code design needs. For codes with fine-grained detail, often examined with specialized instruments at high levels of magnification, small contiguous regions must be decodable. For codes masked by branding materials and product information and acquired with a smartphone, a code block can span a large surface, with the parity-check matrix linking spatially distant bits. Another approach suitable for commercial packaging consists in printing random linear combinations of small blocks—small enough to fit between branded product markings—so that the complete message can be decoded with enough acquired linear combinations.

1.2. Direct rectification of planar markers

The perspective projection of a 3D scene onto an image plane induces geometric distortion. Information displayed on a planar surface—such as a QR code on the face of a cuboid package—that is acquired with perspective can be placed in a standard view (fronto-parallel, typically) via a homography. A homography $H : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a nonlinear planar deformation determined by an isomorphism of the real projective plane \mathbb{P}_2 (the set of lines through the origin in \mathbb{R}^3). Thus, each homography maps distinct lines to distinct lines, preserving incidence. While homographies preserve collinearity, they break parallelism and distort angle and distance ratios. This distorts encoded data motifs, rendering their accurate recovery difficult. The rectifying homography is found using the acquired pixel coordinates of

at least four spatial landmarks [22], whose locations in a rectified code are generally known. Applying the rectifying homography to the acquired image places the data in a standard form accessible to the decoder.

But spatial landmarks impede stealth: subtle landmark motifs are difficult to detect reliably. One can compensate for poor keypoint detection reliability with quantity. With knowledge of the entire fiducial marker that was acquired in standard form (which is rare in traceability applications), one can use algorithms such as SIFT to find available keypoint correspondences between the known and acquired patterns, from which the rectifying homography for the acquired image can be calculated [22, 33]. But such an approach requires explicit knowledge of the entire pattern (rather than a small set of reliably detected, high-contrast landmarks, like QR code “eyes,” which do not contain variable coded data and thus can be known without a complete template). Even with template access, such approaches may still suffer from matched descriptor points that are degenerate, unreliable, or unavailable without tweaks to the keypoint isolation algorithm. This especially holds for patterns made from faint, repetitive motifs, such as the patterns we seek to create (the ideal traceability markers are imperceptible). For traceability markers, one often has access to the *sites* at which data might be encoded but not the data itself. In such cases, cross-correlation-based methods are more practical than correspondence-based methods. Such methods are computationally expensive and ill-suited for the recovery of many rectifying homography parameters. An overview of existing inverse, template-free rectification methods is provided in the Supplementary Materials, Sec. A.

1.3. Self-rectifiable textures: textures that pose a natural inverse problem for rectification

Fiducial markers are examples of *self-rectifiable patterns*. Structures embedded within the marker—typically the pixel indices of landmarks—reveal information about the pose of the acquisition camera. These patterns may possess data payloads (channel coded, for robustness to occlusion [21]) or auxiliary calibration tools (such as dots for the identification of spatially varying point-spread functions [43]). They announce themselves to the acquisition camera with large, elementary geometric forms and high-contrast, typically binary, signals. Notable exceptions are the ARToolKit [39] and Fourier tag [46] (see Fig. 2). The former contains a large square image that is the weighted sum of a payload consisting of a Discrete Cosine Transform (DCT-II) basis function (or its negative) and a fixed DCT-II basis function used to compensate for the bilateral symmetry of even basis functions. Marker identification is accomplished using phase correlation. The Fourier tag embeds in a large circle a smooth radial signal, fixed across a sector. This

signal is the Discrete Fourier Transform (DFT) of a symmetrized binary signal that permits partial recovery (of its low-frequency component) in the presence of blur. In both cases, the fiducial marker is grayscale and smooth.

Motivated by traceability applications, we seek to design self-rectifying patterns in the form of inconspicuous textures that can be printed across a curved packaging surface. Deformed, cropped texture portions can be rectified modulo translation using a single image without template, based on known statistical properties of the texture.

To achieve inconspicuousness, we place landmarks not in the texture itself but rather in its autocorrelation image. Since the positions of peaks in the autocorrelation image are mapped linearly by the deformation differential, the autocorrelation of patches of a deformed texture expose the local behavior of the deformation. Moreover, to rectify the deformation, the original texture need not be known, only the sites of its most prominent autocorrelation peaks. Template-free rectification (modulo translations) can thus be performed, avoiding expensive phase correlations.¹ With additional knowledge of texture (such as the sites at which data appear in the binary texture in the rectified view), the translation can be acquired with a single phase correlation, thanks to the separability property described in Sec. 3. Thus, the complete affine approximation of the deformation of the patch can be acquired (and rectified) with a single phase correlation, in sharp contrast with the computationally expensive correlation-based acquisition of up to six affine parameters in the template-matching literature. Such methods require a phase correlation in each step of the optimization [28].

While a single square marker, such as a QR code, can only be used to rectify homographies and chessboard arrays of square markers are ill-suited for the aesthetic demands of packaging, *self-rectifiable textures*² can be inconspicuously placed across a curved surface, making its rectification possible and its information decodable. As these textures are repeated and contain no landmarks in the spatial domain, there is no unique site of vulnerability. Data payloads encoded in such textures can be made robust to occlusion and tampering via traditional channel coding over the whole texture, over blocks, or via random linear combinations of blocks. In this way, no particular part of the texture

¹These have cost $O(mn \log mn)$, where m and n are the maximum dimension of the query patch and template. Thus, they are “as expensive” as an autocorrelation computation. If the template has 100 times the area of the query patch, one can compute over 100 patch autocorrelations for the cost of a phase correlation.

²This usage is unrelated to that of [54]. In that work, rectification refers to improvements in constrained synthetic texture generation; “self-rectification” hints at the self-attention in the training process, not the texture itself. In our work, we use “self-rectifying” or “self-rectifiable” to indicate that the texture itself contains sufficiently rich geometric information to enable template-free rectification of planar deformations modulo translation.

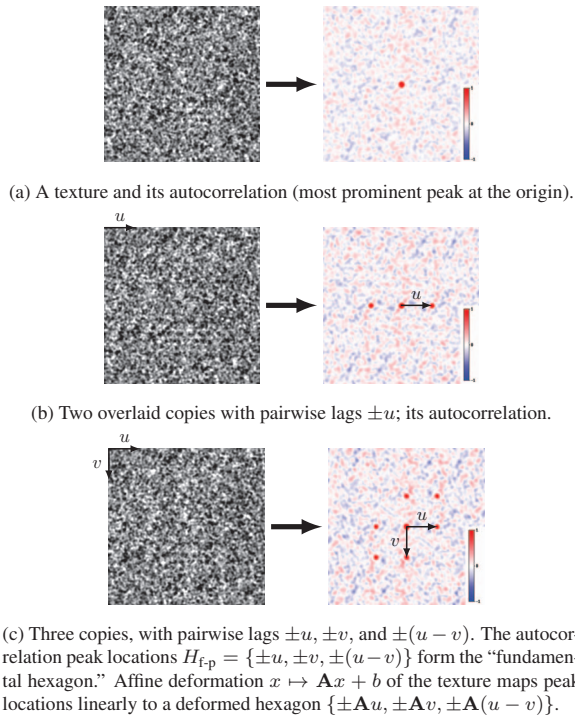


Figure 3. Three textures in fronto-parallel view (left) and their autocorrelation images (right): a base texture f in (a); $f + f(\cdot - u)$ in (b); and $f + f(\cdot - u) + f(\cdot - v)$ in (c). The texture of (b) is used in [42] to rectify scale and rotation without reflection. The texture in (c) has a “fundamental hexagon” of autocorrelation peaks; their locations in deformed image patches are mapped linearly by the deformation differential, permitting template-free rectification of more general deformations modulo translation.

need be acquired to decode a message, only *enough* of the texture.

1.4. Outline

In this article, we generate simple autocorrelation-based self-rectifying textures by summing three copies of a texture with noncollinear offsets (see Algorithm 1), generating a hexagon of known peaks in the autocorrelation image of the texture in the fronto-parallel view (see Fig. 3). When patches of the texture are acquired with perspective on a planar or curved surface, the hexagon in the patch autocorrelation is mapped linearly by differential of the planar deformation. The relative offsets of autocorrelation peaks in a patch of the query image expose the linear part of the deformation between the patch in fronto-parallel view and in the query view; this linear part is found in Algorithm 2. The translation component can be computed using phase correlation (with knowledge of the texture or just locations where data motifs in the texture might be present), enabling

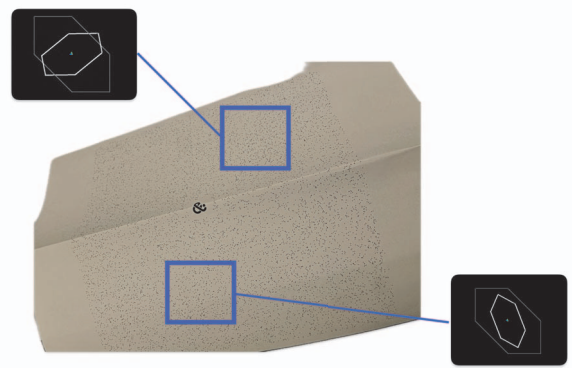


Figure 4. A self-rectifiable texture covering a rectangular prism and the fundamental hexagons of two patches. The orientation of the autocorrelation peaks of each patch reveals the differential of the texture deformation at the patch.

patchwise rectification of complex deformations, as in Algorithm 3. Textures are generated in Sec. 2 and affine maps are rectified in Sec. 3. The robustness of patch-based local affine map acquisition is assessed in Sec. 4. Fully template-free rectification of deformations modulo translations is considered in Sec. 5 and the Supplementary Materials, Sec. C.

2. Generation of self-rectifiable textures

Fig. 3 illustrates the transformation of a sample texture (in fronto-parallel view) into one that is self-rectifying. We require a base texture, such as that of Fig. 3a, whose most prominent autocorrelation peak is at the origin. The autocorrelation of two overlaid copies of the texture in Fig. 3b, at relative shifts 0 and u , possesses three prominent peaks, at 0 and $\pm u$. These peak locations are used to rectify rotation-dilations (orientation-preserving similarities with at least one fixed point) in [42].

A self-rectifying texture is presented in Fig. 3c. Created using Algorithm 1, this texture consists of three summed copies of the base texture with relative shifts of 0, u and v . This shifted overlay results in a “fundamental hexagon” of its most prominent autocorrelation peaks at six nonzero lags: $H = \{\pm(u-0), \pm(v-0), \text{ and } \pm(v-u)\}$. Under mild assumptions, the fundamental hexagon and origin form the seven most intense peaks in the autocorrelation image; our ability to isolate them depends on the prominence of the peak at the origin in the autocorrelation image (and, in practice, the patch size). In the ideal case, the self-rectifying texture consists of three summed, shifted copies of white noise, with autocorrelation equal to $3\delta_0 + \sum_{p \in H} \delta_p$ (where δ_p is the shifted Dirac delta $\delta(\cdot - p)$). But this works far beyond the ideal case: three summed, shifted copies of even substantially self-correlated textures have recoverable

fundamental hexagons.

From these six peak locations—all samples of the oblique Bravais lattice [24] with primitive vectors u and v —we can extract the transformed shifts u and v . For instance, taking all sums and differences of two distinct vectors and reflected copies $\pm a$ and $\pm b$, we recover the third peak and the vector sum of two peaks only when $\{\pm a, \pm b\} = \{\pm u, \pm v\}$. From the offsets alone we cannot uniquely identify u and v in general without additional information, such as their relative positions in the rectified texture and constraints on the affine transformation; oriented motifs in the base texture; additional shifted copies; or distinct weights in the sums of the shifted copies of the base texture. Of course, if we know the form of the rectified texture, we can execute the assignment between hexagon peaks and shifts using the phase correlation.

As Sec. 3 details, affine transformations $x \mapsto \mathbf{A}x + b$ of the texture image effect linear transformations of its autocorrelation, and of the positions of the autocorrelation peaks. The fundamental hexagon is transformed linearly by \mathbf{A} . This result extends in practice to patches of the texture, as Fig. 4 demonstrates. Thus, the peaks in the autocorrelation image of a patch of a texture expose the linear part of the local affine approximation to the deformation we wish to rectify. This approach is novel but familiar: autocorrelation peaks are widely used in signal processing, from channel estimation (*e.g.*, using m -sequences) to watermark recovery [29].

Algorithm 1: One way to create self-rectifying textures: superimpose three shifted copies of a texture.

```
// Key property: if u and v are not collinear
// and f has its most prominent autocorrelation
// peak at 0, the most prominent autocorrelation
// peaks of r form the ``fundamental hexagon.``
```

- 1 **Input:** texture $f : \mathbb{R}^2 \rightarrow \mathbb{R}$; shifts u and v in \mathbb{R}^2
 - 2 **Result:** $r = f + f(\cdot - u) + f(\cdot - v)$
-

Algorithm 1 is but one of many ways to create self-rectifiable textures. It is the simplest method—used here to streamline the exposition and keep focus on the stealth nature of self-rectifying textures, which consequently require inverse rather than correspondence-based rectification—but it is not the one we use in practice. One can place fixed constellations of points around repulsively sampled points. More shifts can be employed. One can choose instead to place peaks in the power spectrum instead (which slightly changes the rectification problem; see Fig. 5a). More generally, one can search for textures with desired autocorrelation (or, equivalently, power spectrum) motifs. The problem of inferring a function from its autocorrelation is widely studied under the name “phase retrieval” as the autocorrelation of a signal determines its Fourier transform modulus but

not phase. Classically, it arises when employing diffracted electromagnetic radiation to determine the structure of an object (*e.g.*, in X-ray crystallography, transmission electron microscopy, etc.) [45]: in such cases, one seeks to image an object with access only to (the Fourier transform of) its autocorrelation. Today, the problem is found in many other domains [3]. Homometry makes this inverse problem ill-posed (a rich space of measures, stochastic and deterministic, can possess the same Fourier transform modulus); solutions return one of many functions consistent with the known autocorrelation (see, *e.g.*, Wrinch’s algorithm [47]). Unique recovery is possible only in certain constrained settings—most notably, binary images of compact convex planar shapes [6]. In our application, unlike in computational imaging, this ill-posedness is an asset, offering design flexibility to satisfy other constraints (due to printing technology, aesthetic imperatives, the need for stochasticity [20]) or to avoid repetition. Moreover, as rectification depends only on pixel locations of autocorrelation structures such as peaks, not the autocorrelation itself, the flexibility is still greater than that found in the classical phase retrieval problem.

3. Efficient extraction of local affine maps thanks to the autocorrelation’s separation of translation from the linear part

For simplicity, we consider images to be integrable functions $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. We are interested in the behavior of images of self-rectifying textures under deformations of the plane. Let $\mathbf{A} \in \text{Aff}(2, \mathbb{R})$ be an affine transformation $x \mapsto \mathbf{A}x + b$ parameterized by $\mathbf{A} \in \mathbb{R}^{2 \times 2}$ (invertible) and $b \in \mathbb{R}^2$. We observe an image g that is f deformed by an affine map: $g = f \circ \mathbf{A}$, so $g(x) = f(\mathbf{A}x + b)$. We wish to recover \mathbf{A} knowing f (or its properties). We denote as \mathbb{T} the map from an image f to its deterministic autocorrelation $R_{f,f} \stackrel{\text{def}}{=} f \star \overleftarrow{f}$ (*i.e.*, $(\mathbb{T}f)(\tau) = \int_{\mathbb{R}^2} f(x + \tau)f(x) dx$). Translations preserve the autocorrelation; thus, by a regular variable substitution, $\mathbb{T}(f \circ \mathbf{A}) \propto \mathbb{T}(f) \circ \mathbf{A}$ (here \propto reads “is proportional to”):

$$\begin{aligned} \mathbb{T}(g)(\tau) &= \int_{\mathbb{R}^2} f(\mathbf{A}(x + \tau) + b)f(\mathbf{A}x + b) dx & (1) \\ &= \frac{1}{|\det \mathbf{A}|} \int_{\mathbb{R}^2} f(t + \mathbf{A}\tau)f(t) dt \propto \mathbb{T}(f)(\mathbf{A}\tau). \end{aligned}$$

Our forward model can be stated simply: if p is the location of a peak in f , then $\mathbf{A}p$ is the location of a peak in $\mathbb{T}f$. The inverse problem for the recovery of the affine map’s linear part \mathbf{A} from the autocorrelation of $g = f \circ \mathbf{A}$ is solved in Algorithm 2. For patchwise rectification, we can recover the phase b using phase correlation. Crucially, this approach, unlike typical template-matching approaches (discussed in

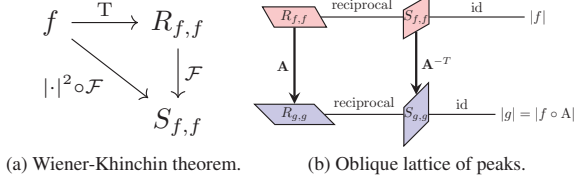


Figure 5. By the commutative diagram (a), an oblique lattice of peaks in a texture f and its power spectrum $S_{f,f}$ is reciprocal to that of its autocorrelation $R_{f,f}$. In (b), these signal peak locations are related to those of $g = f \circ A$, deformed by affine map A with linear part \mathbf{A} . Observations of autocorrelation peaks of a deformed texture expose the linear part \mathbf{A} of an affine deformation; observations of power spectrum peaks, \mathbf{A}^{-T} .

A), separates recovery of \mathbf{A} and b so that only these two parameters need to be recovered via phase correlation.

Define the operator $N_g(f)$ giving the phase correlation with an image g of an image f as follows:

$$N_g f = \mathcal{F}^{-1} \left(\frac{\widehat{f}(\omega) \overline{\widehat{g}(\omega)}}{|\widehat{f}(\omega) \widehat{g}(\omega)|} \right).$$

By a regular variable substitution,

$$\widehat{f \circ A}(\xi) = \frac{1}{|\det \mathbf{A}|} e^{ib^T \mathbf{A}^{-T} \xi} \widehat{f}(\mathbf{A}^{-T} \xi). \quad (2)$$

From (2), we can establish the translation-covariance of the operator N_f : any pure translation $A_t : x \mapsto \mathbf{I}x + b$ satisfies

$$N_g(f \circ A_t) = N_g(f) \circ A_t.$$

As translations do not affect the autocorrelation, we can rectify \mathbf{A} and b separately, as in Algorithm 3. Due to the robustness of the phase correlation, patchwise rectification is often possible when one lacks access to a complete template consisting of the fronto-parallel view of the self-rectifiable texture, but does have the *sites* where coded data may appear (say, in a binary image where only pixels that communicate data may be illuminated).

Eq. (2) allows us to complete Fig. 5b, which shows how peak locations in the Fourier transform magnitude, power spectrum, or autocorrelation of a Bravais lattice are transformed linearly when the lattice image undergoes an affine deformation A . Our example self-rectifying texture, generated using Algorithm 1, possesses not a lattice of peaks in the autocorrelation domain, but seven samples thereon (the fundamental hexagon). This windowing renders the dual-lattice peaks in the power spectrum too difficult to detect to permit an estimation of \mathbf{A} . However, the power spectrum peaks of other self-rectifying textures may be used to estimate the linear part of the affine deformation \mathbf{A} .

Algorithm 2: Estimate linear component of affine transformation from fundamental hexagon.

- 1 **Input:** fundamental shifts u and v in \mathbb{R}^2 ; patch P of deformed texture
- 2 **Result:** linear part \mathbf{A} of affine deformation
- // Compute periodic component [32] autocorrelation
- 3 $R = \text{idft2}(|\text{dft2}(\text{per_component}(P))|^2)$
- // if most prominent peaks exhibit the correct form, fit a Gaussian to each peak's neighborhood, give subpixel maxima locations
- 4 $H_{\text{def}} = \text{get_fundamental_hexagon}(R)$
- // solve the assignment problem between peaks in $H_{\text{f-p}} = \{\pm u, \pm v, \pm(u-v)\}$ and their deformed twins in H_{def} using oriented texture motifs, weights in the overlay sum, constraints on \mathbf{A} and identification of lattice primitive vectors, or phase correlation
- 5 $\text{match} \leftarrow \text{solve_assignment}(\mathbf{A}H_{\text{f-p}}, H_{\text{def}})$
- // \mathbf{A} is least-squares fit to an overdetermined system, well-posed whenever u, v noncollinear
- 6

$$\mathbf{A} = \underset{\mathbf{A} \in \mathbb{R}^{2 \times 2}}{\text{argmin}} \sum_{i=1}^6 \|\mathbf{A}H_{\text{f-p}}[i] - H_{\text{def}}[\text{match}(i)]\|_{\mathbb{R}^2}^2 \quad (3)$$

4. Robustness of patchwise rectification

Complex deformations are often rectified patchwise, by fitting affine maps or homographies to each deformed patch. These maps are widely used in remote sensing, image stitching, and registration applications, as well as—in reverse—in image projection on nonplanar surfaces [4, 31, 55]. Algorithm 3 details our patchwise rectification process. Fig. 6 shows patchwise rectification of a homography.

The change of variables used in Eq. (1) assumes an infinite texture. Preprocessing can enhance peak prominence in finite textures, particularly for small or axis-aligned shifts. In Algorithm 2, we retain patch periodic components [32]. Texture motif and density choices also impact hexagon peak acquisition. In this study, we consider two base textures: a random binary texture and a binary commercial marker of 3×3 square motifs placed randomly on a regular grid of a tunable density, using a (128, 64) low-density parity check (LDPC) code. The regular grid of the latter sacrifices peak prominence for ease of decoding.

Patch sampling strategies require great care when deformations are not affine. Patches barely larger than the fundamental shifts u and v lack signal correlation runtime; the echoes of shifted copies are barely recognizable, and the corresponding autocorrelation peaks are hard to dis-

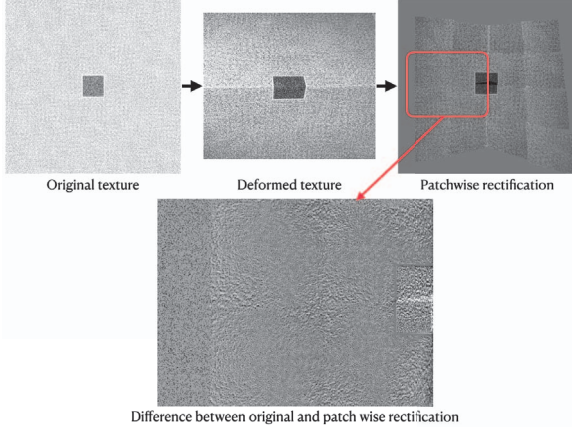


Figure 6. Patchwise rectification of a self-rectifying texture printed at 300 DPI and affixed across a table border. Patchwise rectification is accomplished using Algorithm 3, given knowledge of the *sites* at which data may be encoded in a commercial traceability marker but not the template code. As the difference image attests, redressed points are sufficiently well-placed to enable decoding.

tinguish from the background. Moreover, due to the use of the DFT to compute the autocorrelation, their locations can be aliased (see Supplementary Materials, Fig. 7). With patch lengths a bit over twice the fundamental shifts, peaks are prominent. Beyond that, our linearized forward model can break down: peaks of large patches appear at locations inconsistent with a deformed fundamental hexagon $\mathbf{A}H_{f-p}$ due to approximation error. Affine deformations map peaks situated on an oblique Bravais lattice in H_{f-p} to peaks on another oblique lattice $\mathbf{A}H_{f-p} = H_{def}$; non-affine maps warp the lattice. While the difference between primitive vectors is in the fundamental hexagon $u - v \in H_{f-p}$, this does not hold for their deformed pairs u' and v' : $(u' - v') \notin H_{def}$. We consider the hexagon H_{def} undetectable when we cannot isolate the primitive vectors of H_{def} in this way, within tolerance.

In the Supplementary Materials, Sec. D, we show that the recovery of the linear part \mathbf{A} of an affine map from the fundamental hexagon is possible even with poor conditioning. It is standard [22] to write the singular value decomposition of $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = (\mathbf{U}\mathbf{V}^T)\mathbf{V}\mathbf{\Sigma}\mathbf{V}^T$ in the orientation-preserving subgroup of $GL(2, \mathbb{R})$ as a rotation \mathbf{V}^T by a “tilt angle” ϕ , a non-isotropic scaling by a diagonal matrix $\mathbf{D} = \text{diag}(1, t)$ in the rotated coordinates, a rotation \mathbf{V} by $-\phi$, a final rotation $\mathbf{U}\mathbf{V}^T$ by θ , and an isotropic scaling (zoom) λ . When $t < 1$, the tilt t is the inverse of the condition number of the linear part \mathbf{A} of an affine map.

While a continuous image of an infinite texture always permits recovery of nondegenerate \mathbf{A} , per Equation (1), reliable recovery of the transformed hexagon from a finite,

sampled patch depends on the texture, sampling strategy (algorithm to choose patch size(s), locations, and number), and acquisition mode. These parameters must be assessed with experimentation specific to the application. In our experiments, we used the naive strategy of sampling patches of a fixed size uniformly for piecewise rectification and randomly, with variable size determined by Supplementary Materials Algorithm 6, for homographic rectification.

Algorithm 3: Patchwise rectification of a planar deformation. Requires phase correlation and a template (fronto-parallel view of texture or simply the *sites* at which binary data may be encoded).

```

1 Input: A template  $T$  and deformed texture  $I$ 
2 Result: Patchwise rectified image  $R$ 
//  $I$  is divided into patches, overlapping where
// needed, with size and shape chosen so that the
// periodic component of each patch has a
// detectable transformed fundamental hexagon.
3  $P_1, \dots, P_n = \text{patchify}(I)$ 
// Get the local affine approximation of the
// deformation  $\varphi$ . For each patch  $P_i$  centered at
//  $y_i$ , our estimate of the deformation's
// differential  $\nabla_{\varphi^{-1}(y_i)}\varphi$  is  $\mathbf{A}_i$  and with the
// phase  $b_i$  we complete the first-order
// approximation about  $x_i = \varphi^{-1}(y_i) = \mathbf{A}_i^{-1}(y_i - b_i)$ :
//  $h \mapsto y_i + \mathbf{A}_i(h - x_i) = \mathbf{A}_i h + b_i$ .
4 for  $i = 1, \dots, n$  do
5    $\mathbf{A}_i = \text{get\_linear\_component}(P_i)$ 
6    $b_i = \underset{b \in \mathbb{R}^2}{\text{argmax}} N_{\mathbf{A}_i^{-1}P_i}(b)$ 
// Apply the inverse map and interpolate.
7 for  $i = 1, \dots, n$  do
8   Form  $R_i$  by applying the map  $h \mapsto \mathbf{A}_i^{-1}(h - b_i)$ 
// to  $P_i$ , using bicubic interpolation to remain on
// pixel grid
9 Stitch together the overlapping  $P_i$  to form  $R$ .
```

5. Extension to template-free rectification

The patchwise rectification described in Algorithm 3 requires access to a template: an image in the fronto-parallel view of the binary texture being rectified or of sites where data may appear. This template is used to recover the translation component of the local affine map via a phase correlation. In certain applications, rectification of a deformation without access to such a template is desirable.

From the autocorrelation peaks of each deformed patch P_i , centered at y_i , we can estimate the differential

$$\mathbf{A}_i \approx \nabla_{\varphi^{-1}(y_i)}\varphi \quad (4)$$

of the deformation φ using Algorithm 2. For this, we need

no template image, only the shift parameters u, v used to generate the self-rectifiable texture using Algorithm 1 and modest additional information to distinguish symmetries and resolve the assignment problem between peaks. The differential of a homography, for instance, involves all its parameters, including the translation parameters. It is natural, then, to wonder whether φ can be recovered from our indirect observations \mathbf{A}_i of its differential $\nabla_{\varphi^{-1}(y_i)}\varphi$, made using peak displacement.

Not entirely, it turns out. As our observations are made through the translation-invariant autocorrelation, they cannot distinguish between two deformations related by translation, φ and $\varphi \circ \tau$. By the chain rule, the local affine transformations estimated using P_i , centered at y_i , of both φ and $\varphi \circ \tau$ share a linear part, since $\nabla_x \tau = \mathbf{I}$ for all $x \in \mathbb{R}^2$:

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

Thus, our indirect observations alone cannot be used to distinguish φ and $\varphi \circ \tau$; we can only hope to recover φ modulo translation. Nevertheless, we can still rectify without template by assembling our scattered observations \mathbf{A}_i into a global model of φ modulo translation. With access to our deformation modulo translation, we need just a single observation of the translation—say, via a synchronization sequence, or via information encoded in our texture—to resolve the ill-posedness of our inverse problem and recover φ . Moreover, in certain circumstances, recovering φ modulo translations suffices to decode the message.

To assemble our scattered observations \mathbf{A}_i of $\nabla_{\varphi^{-1}(y_i)}\varphi$ into an estimate of φ modulo translations, we have two approaches: by fitting the \mathbf{A}_i to a parametric model of φ or by obtaining a non-parametric representation of φ .

Template-free rectification can be performed nonparametrically, by integrating the observations (Eq. (4)) directly over a triangle mesh into a global model. The two rows of our observations \mathbf{A}_i form an integrable polyvector field. The literature on polyvector fields [9, 41, 51] offers ways to integrate [15] these observations. They also permit denoising: a polyvector field can be represented as a polynomial, determined by two complex coefficients, that exhibits the same symmetries [40]. The Dirichlet energy of the vector field of coefficients across the image or graph of sample sites is widely used as a smoothness penalty.

In practice, parametric methods require fewer observations to characterize the rectifying map for a marker on a packaging surface and thus are more suitable for real-time rectification than direct integration. We can model the *rectifying* deformation parametrically—using, for instance, homographies, polynomials, or thin-plate splines—and fit its parameters to our inverted observations $\mathbf{B}_i = \mathbf{A}_i^{-1}$. In the Supplementary Materials, Sec. C, we show that the

problem of fitting rectifying homographies modulo translation to these observations can be reduced to an optimization over a six-dimensional affine subspace of \mathbb{R}^9 . In short, we can view our homographies—the projective linear group $\text{PGL}(3, \mathbb{R})$ —as isomorphic to the special linear group $\text{SL}(3, \mathbb{R})$ as a Lie group and manifold. The quotient manifold of $\text{SL}(3, \mathbb{R})$ modulo translations is covered by the union of three manifolds. Only one of these manifolds, the Lie group $\text{A}_2(2, \mathbb{R})$ of matrix transposes of the affine group [13], is relevant to the rectification of homographies encountered in practice. The manifold $\text{A}_2(2, \mathbb{R})$ is dense in a six-dimensional affine subspace \mathcal{H} of \mathbb{R}^9 over which we optimize, using `pymanopt`'s trust-region algorithm [1, 50] and Euclidean manifold. As we do not encounter singular optima in practice, we can interpret the result as a member of $\text{A}_2(2, \mathbb{R})$, representing an orbit of $\text{SL}(3, \mathbb{R})$ modulo translations—the matrix that has been “translated to zero.”

Let H be the rectifying homography and $G = H^{-1}$ the deformation. The scattered observations \mathbf{A}_i of $\nabla_{G^{-1}(y_i)}G$ computed with Algorithm 2 can be used to estimate H . By the inverse function theorem, $\mathbf{B}_i = \mathbf{A}_i^{-1}$ is an estimate of $\nabla_{y_i}H$. Using the robust (non-squared) Frobenius norm loss, we choose parameters of H by matching the parametric form $\nabla_{y_i}H$ to our scattered data $\mathbf{B}_i = \mathbf{A}_i^{-1}$ in Algorithm 4.

Algorithm 4: Estimation of homography.

- 1 **Input:** Observations $\{\mathbf{A}_i\}_{i=1}^n$ of $\{\nabla_{G^{-1}(y_i)}G\}_{i=1}^n$
 - 2 **Output:** Rectifying homography H
 - 3 Solve $H^* = \underset{H \in \mathcal{H}}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^n \|\nabla_{y_i}H - \mathbf{A}_i^{-1}\|_F$
-

Nonparametric homography recovery is demonstrated in Supplementary Materials, Fig. 9c, Fig. 8, and Sec. E.

6. Conclusion

Self-rectifying textures permit the estimation of local surface geometry and perspective using a single image without keypoint matches. Rectification can be performed patchwise, using correlation with texture templates (or data encoding sites), or via model-fitting. Only the two translation parameters are fit with expensive phase correlations; the texture exposes the other four for affine patchwise rectification or homographic rectification modulo translation. With no spatial landmarks, these stealth textures can be placed across a surface, granting robustness to tampering. Since rectification depends on the autocorrelation, it is robust to a variety of contaminants that preserve locations of prominent peaks, such as blur and noise. Acquisition of the linear component of the local affine approximation to a deformation is robust and accurate in traceability applications.

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. 8
- [2] Dror Aiger, Daniel Cohen-Or, and Niloy J Mitra. Repetition maximization based texture rectification. In *Computer Graphics Forum*, pages 439–448. Wiley Online Library, 2012. 11
- [3] Emmanuel Amiot. Homometry and the phase retrieval problem. *Music Through Fourier Space: Discrete Fourier Transform in Music Theory*, pages 27–49, 2016. 5
- [4] Vicente Arevalo and Javier Gonzalez. Improving piecewise linear registration of high-resolution satellite images through mesh optimization. *IEEE Transactions on Geoscience and Remote Sensing*, 46(11):3792–3803, 2008. 6
- [5] Bradley Atcheson, Felix Heide, and Wolfgang Heidrich. Caltag: High precision fiducial markers for camera calibration. In *Vision, Model., Visualization*, pages 41–48, 2010. 1
- [6] Gennadiy Averkov and Gabriele Bianchi. Confirmation of Matheron’s conjecture on the covariogram of a planar convex body. *Eur. J. Math.*, 11(6):1187, 2009. 5
- [7] Coloma Ballester and Manuel González. Affine invariant texture segmentation and shape from texture by variational methods. *J. Math. Imaging Vis.*, 9:141–171, 1998. 11
- [8] Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape matching and object recognition using shape contexts. *IEEE transactions on pattern analysis and machine intelligence*, 24(4):509–522, 2002. 11
- [9] Mikhail Bessmeltsev and Justin Solomon. Vectorization of line drawings via polyvector fields. *ACM Trans. Gr.*, 38(1): 1–12, 2019. 8
- [10] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. Artoolkit: A computer vision based augmented reality toolkit. *IEEE VR2000, New Jersey*, 2000. 2
- [11] Jacques Brochard, Majdi Khoudeir, and Bertrand Augereau. Invariant feature extraction for 3D texture analysis using the autocorrelation function. *Pattern Recogn. Lett.*, 22(6):759–768, 2001. 11
- [12] Vicent Caselles, Bartomeu Coll, and Jean-Michel Morel. Geometry and color in natural images. *J. Math. Imaging Vis.*, 16:89–105, 2002. 11
- [13] 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. 8, 14
- [14] U.S. Consumer Product Safety Commission. Tracking label business guidance. Technical report, Guidance (updated October, 2022), 2022. 1
- [15] Olga Diamanti, Amir Vaxman, Daniele Panozzo, and Olga Sorkine-Hornung. Integrable polyvector fields. *ACM Trans. Graph. (TOG)*, 34(4):1–12, 2015. 8
- [16] Csaba Domokos, Jozsef Nemeth, and Zoltan Kato. Nonlinear shape registration without correspondences. *IEEE Transactions on pattern analysis and machine intelligence*, 34(5): 943–958, 2011. 11
- [17] Mark Fiala. ARTag, a fiducial marker system using digital techniques. In *CVPR 2005: IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit.*, pages 590–596. IEEE, 2005. 1
- [18] Directorate-General for Environment. *Proposal for Ecode-sign for Sustainable Products Regulation*. The European Commission, 2022. 1
- [19] 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 Recognit.*, 47(6):2280–2292, 2014. 2
- [20] Uwe Grimm and Michael Baake. Homometric point sets and inverse problems. *Z. Kristallogr.*, 223(11-12):777–781, 2008. 5
- [21] Masahiro Hara, Motoaki Watabe, Tadao Nojiri, Takayuki Nagaya, and Yuji Uchiyama. Optically readable two-dimensional code and method and apparatus using the same, 1998. US Patent 5,726,435. 1, 3
- [22] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2003. 2, 3, 7
- [23] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, 2 edition, 2004. 15
- [24] Yuchen He and Sung H Kang. Lattice identification and separation: theory and algorithm. *SIAM J. Imaging Sci.*, 12(4): 2063–2096, 2019. 5
- [25] Berthold KP Horn and Michael J Brooks. The variational approach to shape from shading. *Comput. Gr. Image Process.*, 33(2):174–208, 1986. 11
- [26] International Organization for Standardization and International Electrotechnical Commission. Information technology – automatic identification and data capture techniques – qr code bar code symbology specification, 2015. Standard number: ISO/IEC 18004:2015. 2
- [27] Michail Kalaitzakis, Brennan Cain, Sabrina Carroll, Anand Ambrosi, Camden Whitehead, and Nikolaos Vitzilaios. Fiducial markers for pose estimation: Overview, applications and experimental comparison of the artag, apriltag, aruco and stag markers. *Journal of Intelligent & Robotic Systems*, 101(4):71, 2021. 2
- [28] Shun’ichi Kaneko, Yutaka Satoh, and Satoru Igarashi. Using selective correlation coefficient for robust image registration. *Pattern Recognition*, 36(5):1165–1173, 2003. 3, 11
- [29] Martin Kutter. Watermarking resistance to translation, rotation, and scaling. In *Multimedia Syst. Appl.*, pages 423–431. SPIE, 1999. 5
- [30] Jongwoo Lim and Ming-Hsuan Yang. A direct method for modeling non-rigid motion with thin plate spline. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, pages 1196–1202. IEEE, 2005. 11
- [31] Chung-Ching Lin, Sharathchandra U Pankanti, Karthikeyan Natesan Ramamurthy, and Aleksandr Y Aravkin. Adaptive as-natural-as-possible image stitching. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1155–1163, 2015. 6
- [32] Lionel Moisan. Periodic plus smooth image decomposition. *J. Math. Imaging Vis.*, 39:161–179, 2011. 6

- [33] Lionel Moisan, Pierre Moulon, and Pascal Monasse. Fundamental matrix of a stereo pair, with a contrario elimination of outliers. *Image Process. Line*, 6:89–113, 2016. [3](#)
- [34] Mads Nielsen, Peter Johansen, Andrew D Jackson, Benny Lautrup, and Søren Hauberg. Brownian warps for non-rigid registration. *Journal of Mathematical Imaging and Vision*, 31(2):221–231, 2008. [11](#)
- [35] Jingyi Ning, Lei Xie, Yi Li, Yingying Chen, Yanling Bu, Baoliu Ye, and Sanglu Lu. Moirépose: ultra high precision camera-to-screen pose estimation based on moiré pattern. In *Proceedings of the 28th annual international conference on mobile computing and networking*, pages 106–119, 2022. [11](#)
- [36] Richard Nock and Frank Nielsen. Statistical region merging. *IEEE Transactions on pattern analysis and machine intelligence*, 26(11):1452–1458, 2004. [11](#)
- [37] U.S. Department of Homeland Security Office of Strategy, Policy, and Plans. Combating trafficking in counterfeit and pirated goods. Technical report, U.S. Department of Homeland Security, 2020. Published pursuant to the Presidential Memorandum dated April 3, 2019. [1](#)
- [38] Edwin Olson. Apriltag: A robust and flexible visual fiducial system. In *2011 IEEE international conference on robotics and automation*, pages 3400–3407. IEEE, 2011. [2](#)
- [39] Charles B Owen, Fan Xiao, and Paul Middledin. What is the best fiducial? In *IEEE Int. Workshop Augmen. Reality Toolkit*, pages 8–15. IEEE, 2002. [3](#)
- [40] David Palmer, David Bommes, and Justin Solomon. Algebraic representations for volumetric frame fields. *ACM Trans. Gr.*, 39(2):1–17, 2020. [8](#)
- [41] D. Palmer, O. Stein, and J. Solomon. Frame field operators. *Comput. Gr. Forum*, 40(5):231–245, 2021. [8](#)
- [42] Justin Picard, Jean-Pierre Massicot, Alain Foucou, and Zbigniew Sagan. Method and device for securing documents, United States Patent 0220364A1, Sep. 2, 2010. [4](#), [11](#)
- [43] Alexander Reuter, Hans-Peter Seidel, and Ivo Ihrke. BlurTags: spatially varying PSF estimation with out-of-focus patterns. In *WSCG 2012: Int. Conf. Comput. Gr., Visualization, Comput. Vision*, pages 239–247, 2012. [1](#), [3](#)
- [44] Mariano Rodríguez, Gabriele Facciolo, and Jean-Michel Morel. Robust homography estimation from local affine maps. *Image Process. Line*, 13:65–89, 2023. [14](#)
- [45] Joseph Rosenblatt. Phase retrieval. *Commun. Math. Phys.*, 95(3):317–343, 1984. [5](#)
- [46] Junaed Sattar, Eric Bourque, Philippe Giguere, and Gregory Dudek. Fourier tags: Smoothly degradable fiducial markers for use in human-robot interaction. In *CRV 2007: Can. Conf. Comput. Rob. Vision*, pages 165–174. IEEE, 2007. [1](#), [3](#)
- [47] Marjorie Senechal. A point set puzzle revisited. *Eur. J. Combinatorics*, 29(8):1933–1944, 2008. [5](#)
- [48] J.S. Seo, J. Haitzma, T. Kalker, and C.D. Yoo. Affine transform resilient image fingerprinting. In *ICASSP 2003: IEEE Int. Conf. Acoust. Speech Signal Process.*, pages III–61, 2003. [11](#)
- [49] Richard Szeliski. *Computer Vision: Algorithms and Applications*. Springer, London, UK, 2010. [16](#)
- [50] 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. [8](#)
- [51] Amir Vaxman, Marcel Campen, Olga Diamanti, Daniele Panozzo, David Bommes, Klaus Hildebrandt, and Mirela Ben-Chen. Directional field synthesis, design, and processing. *Comput. Gr. Forum*, 35(2):545–572, 2016. [8](#)
- [52] Dor Verbin and Todd Zickler. Toward a universal model for shape from texture. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recogn.*, pages 422–430, 2020. [11](#)
- [53] Andrej Zdešar, Igor Škrjanc, and Gregor Klančar. Homography estimation from circular motion for use in visual control. *Robotics and autonomous systems*, 62(10):1486–1496, 2014. [11](#)
- [54] Yang Zhou, Rongjun Xiao, Dani Lischinski, Daniel Cohen-Or, and Hui Huang. Generating non-stationary textures using self-rectification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7767–7776, 2024. [3](#)
- [55] Bo Zhu, Li-jun Xie, Qi-hui Wang, Ting-jun Yang, and Yao Zheng. An intelligent projection system adapted to arbitrary surfaces. In *2011 First International Conference on Instrumentation, Measurement, Computer, Communication and Control*, pages 293–298. IEEE, 2011. [6](#)