Detection, Rectification and Segmentation of Coplanar Repeated Patterns

James Pritts Ondřej Chum Jiří Matas
Center for Machine Perception
Czech Technical University in Prague, Faculty of Electrical Engineering, Department of Cybernetics
[prittjam,chum,matas]@cmp.felk.cvut.cz

Abstract

This paper presents a novel and general method for the detection, rectification and segmentation of imaged coplanar repeated patterns. The only assumption made of the scene geometry is that repeated scene elements are mapped to each other by planar Euclidean transformations. The class of patterns covered is broad and includes nearly all commonly seen, planar, man-made repeated patterns. In addition, novel linear constraints are used to reduce geometric ambiguity between the rectified imaged pattern and the scene pattern. Rectification to within a similarity of the scene plane is achieved from one rotated repeat, or to within a similarity with a scale ambiguity along the axis of symmetry from one reflected repeat. A stratum of constraints is derived that gives the necessary configuration of repeats for each successive level of rectification. A generative model for the imaged pattern is inferred and used to segment the pattern with pixel accuracy. Qualitative results are shown on a broad range of image types on which state-of-the-art methods fail.

1. Introduction

This paper presents a method for the detection, segmentation and rectification of a broad class of coplanar repeated patterns. A pattern is defined to be a coplanar arrangement of a *motif* (see Fig. 3). This definition is very general: no restrictive constraints are imposed on motif placement, *e.g.*, constrained to a lattice, a circle, or a row. The method does not require the presence of parallel scene lines to work. The proposed detection and rectification method applies to nearly all man-made structures with repetitions or symmetries, which include building facades, mosaics, decorative prints and text.

The problem of rectification is closely coupled with the detection of a pattern and its motif in a chicken-and-egg fashion: rectification is easier if the motif is known, and estimation of the motif is easier from a rectified image. The proposed approach begins by establishing tentative intra-image correspondences of local affine frames (LAFs) [15] that could belong to a pattern. Rectification is estimated exclusively from the ratio of scales and lengths of extents of the imaged repeated scene structure.

Figure 1. Two different repeated patterns (left column) detected and used to rectify the images (right column) by the proposed algorithm. Top: similarity rectification with scale ambiguity in the vertical axis from a pattern with reflected elements; bottom: similarity rectification from an irregular pattern with rotated elements—note the right angle of the corner in the rectified image.

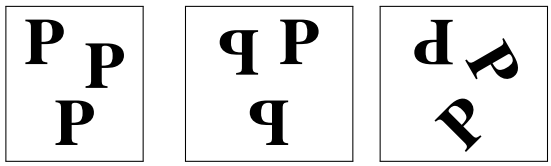
Three types of repeated patterns are considered with the goal being to bring each imaged pattern type as geometrically close as possible to its scene counterpart to improve scene understanding. The three pattern types differ by the class of transformations between motif instances in the scene plane. Each type provides different geometric constraints and allows for different levels of rectification. A set of linear single-view geometric constraints are used that, depending on the pattern arrangement, enable estimation of rectification to within a similarity of the scene pattern. A generative model for the imaging process of the repeated pattern and its corresponding non-linear estimator is given. An explicit model of the pattern is constructed and projected

[1]James Pritts and Ondřej Chum were supported by the GACR P103/12/2310 and MSMT LL1303 ERC-CZ grants, and Jiří Matas by the Czech Science Foundation Project GACR P103/12/G084.

1063-6919/14 $31.00 © 2014 IEEE
DOI 10.1109/CVPR.2014.380

Original image | Vanishing points | TILT [22] | Aiger *et al*. [1] | Proposed

Figure 2. Comparison with other rectification approaches. Publicly available implementation by authors were used for [22] and [1].



(a) (c) (b)

Figure 3. Examples of different configurations of a repetitive pattern: (a) pure translation, (b) composition of an axial symmetry and translation, (c) general rigid transform.

into the image. Reprojection error is minimized by refining estimates of scene and camera geometry and localization of extracted features, and a pixel-accurate segmentation is estimated using the re-projected pattern.

In general, methods for the detection and analysis of coplanar repeated patterns benefit or require rectification of the input image to detect and model repeated patterns. However, prior approaches make restrictive assumptions on scene content, *e.g*., requiring parallel scene lines, specific symmetries, or that repeated pattern elements form a lattice [1, 8, 11, 17, 19, 20, 21]. Some methods need the region of interest to be manually annotated if there is background clutter, one commonly used example is TILT [22], and a nearly universal requirement is that radial distortion is (manually) removed from the inputted image prior to any processing. The proposed method does not need manual annotation or camera calibration, and works for low-texture patterns with significant perspective warping acquired by lens-distorted cameras. Pattern construction and lens distortion estimation is an integral part of the rectification process.

1.1. Related work

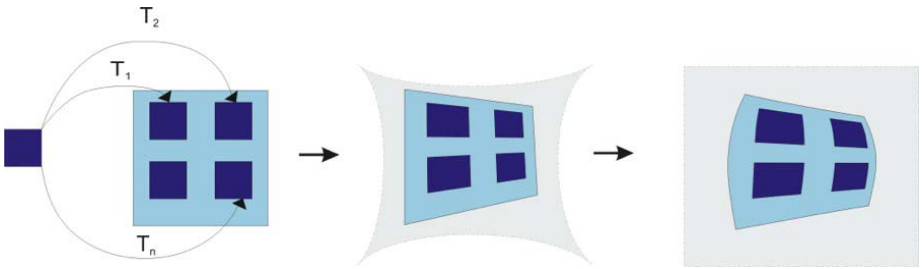
Leung and Malik [10] proposed an early solution that grows a 2-D lattice from initial patches of high texture in a manner analogous to SSD tracking. Schaffalitzky and Zisserman [17] also used local patches to grow a grid, and parameterized imaged planar repetitions as conjugate translations and rotations of local patches. Tuytelaars *et al*. [19] grouped fixed points and lines from symmetries consistent with homologies estimated by Hough transform. Liu *et al*. [12] modeled the lattice structure of textures using a minimal set of tiles. Park *et al*. [16] modeled lattice topology with a Markov random field and allowed for some non-planar deformation. Doubek *et al*. [4] developed a shift-invariant descriptor to detect patterns with translated motifs.

Hong *et al*. [8] developed constraints based on several classes of symmetries to estimate camera pose and pattern structure, but assumed that the symmetry type is given. Wu *et al*. [20, 21] introduced a repetition constraint and a corresponding optimization method to densely reconstruct repetitive structures, but they assumed the presence of parallel scene lines and that repeats occur on a grid. Rectification of repeated structures by rank minimization of the intensity matrix was proposed in [22]. The method is sensitive to background clutter and occlusions, and the region of interest must be manually annotated in problem cases. Liebowitz and Zisserman [11] estimated affine rectification from vanishing lines, then, similar to our approach, upgraded the rectification to within a similarity by enforcing the congruence of the length extents of repeated scene elements. But only rotated repetitions were considered. Recently, Aiger *et al*. [1] also proposed estimating rectification by maximizing congruence of the length extents of repeats. However, optimization was performed jointly for affine and metric rectification, which has no convergence guarantee since congruence is not preserved by either projective or affine transforms.

Köser *et al*. [9] proposed the estimation of conjugate rotation from a planar image correspondence, which they use as a local first-order Taylor approximation to the inter-image homography. Using the same linearization, an algebraic constraint on the scale change of a planar scene feature transformed by a homography was introduced by Chum *et al*. in [3]. The method attempts to find a rectifying homography $\mathtt{H}_\infty$ so that the matching patches of the repeated structure cover identical areas. The proposed method extends the method in [3] by further reducing the ambiguity between imaged pattern and scene pattern for certain types of repeated patterns and significantly increases the precision of rectification by a non-linear optimization that explicitly reconstruct the pattern and accounts for radial lens distortion.

2. Problem formulation

The method searches the input image for a coplanar repeated pattern that can be generated by the assumed model illustrated in Fig. 4. The *motif* is repeatedly stamped on the scene plane (a priori unknown) at locations specified by Eu-



(a) (b) (c) (d)

Figure 4. Generative model of the image of a repetitive pattern: (a) a *motif* is 'stamped' by transformations $\mathtt{T}_j$ onto (b) a plane in 3D. The image of the plane (c) is captured by a projective camera with lens distortion (d). Only the final image (d) is the input of the method; the previous stages are inferred from intra-image correspondences. The only assumptions are made on the type of transformations $\mathtt{T}_j$.

clidean transformations $\mathtt{T}_j$, which constructs the pattern in a rectified space. The construction is imaged by a radially distorted projective camera. Since the construction lies on a scene plane, the imaging process is modeled with a homography. The projective part of the homography is captured by the image of the line at infinity. Finally, radial lens distortion is applied to the image plane creating the input image. Note that the image is the sole input to the algorithm, and all parameters of the generative model are estimated. An assumption of the generative model is that the positioning transformations $\mathtt{T}_j$ are one of three types typically encountered: pure translation, composition of an axial symmetry with translation, or a general Euclidean (rigid) transformation (Fig. 3). It is of little interest to assume that $\mathtt{T}_j$ is as general as a homography because (i) such repeated elements rarely appear on real-world scene planes, and (ii) it is impossible to decouple the projective placement of the scene element from the projective effect of the camera . There are no other assumptions about the positioning transformations: repeated elements do not have to create a lattice or assume any other form.

The *motif* is a spatially localized collection of features that form the repeated element. Due to noise or occlusions, we assume that each instance of the motif in the pattern consists of a subset of spatially consistent motif features; *i.e.*, not all motif features have to appear in all instances. To contribute to a solution, an instance of a motif must contain at least two features. For a solution to be found, at least two motif instances must be detected.

3. The Method

This section describes the step-by-step estimation of the parameters of the generative model from a single image of a repeated pattern. An overview of the approach is given in Alg. 1.

Feature appearance matching. To obtain sets of features potentially arising from different instances of the repeated motif, an intra-image matching procedure is performed. Affine invariant local features are first extracted in the image. We use MSER features [14] with the local affine frame (LAF) extension [15] since it provides a local coordinate system for each feature in the form of a triple of points. SIFT descriptors [13] are extracted from the distinguished regions defined by the LAFs. Features used for construction of the reflected instances are detected in a reflected image. Spectral descriptor matching is employed to robustly extract clusters of features with similar appearance. Examples of matching LAFs can be found in Figs. 5 and 6.

Algorithm 1 Overview of the method

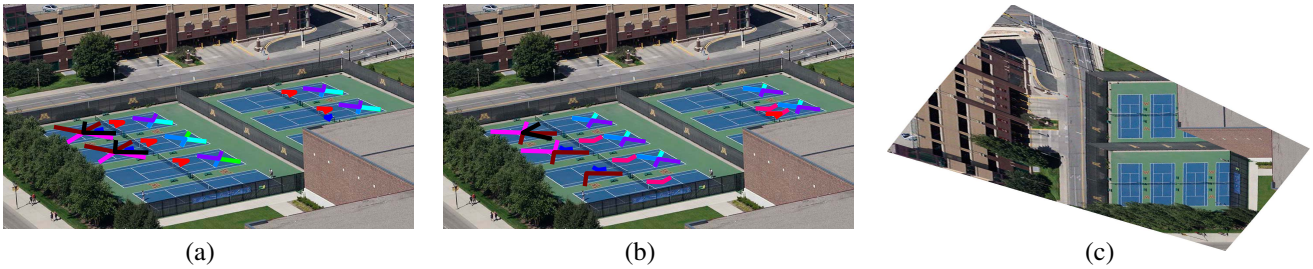
in: an image
out: motif, positioning transforms $\mathtt{T}_j$, line at infinity, radial distortion

1. Feature appearance matching
2. Projective distortion removal
3. Motif construction
4. Affine distortion correction
5. Non-linear optimization

Projective distortion removal. The line at infinity is estimated from the change of scale of matching local features following [3]. In a projective image, patches that are farther from the camera appear smaller. A rectifying homography is estimated so that all matching features detected in different instances of the motif have equal area after rectification. The constraints on the homography are derived from the scale change at a point, or, equivalently, at an infinitesimally small patch around the point. This scale change is given by the determinant of an affine transformation locally approximating (first order Taylor expansion) the homography at the point. The infinitesimal scale changes are estimated from the areas of local, affine covariant features. This approach gives a set of linear equations

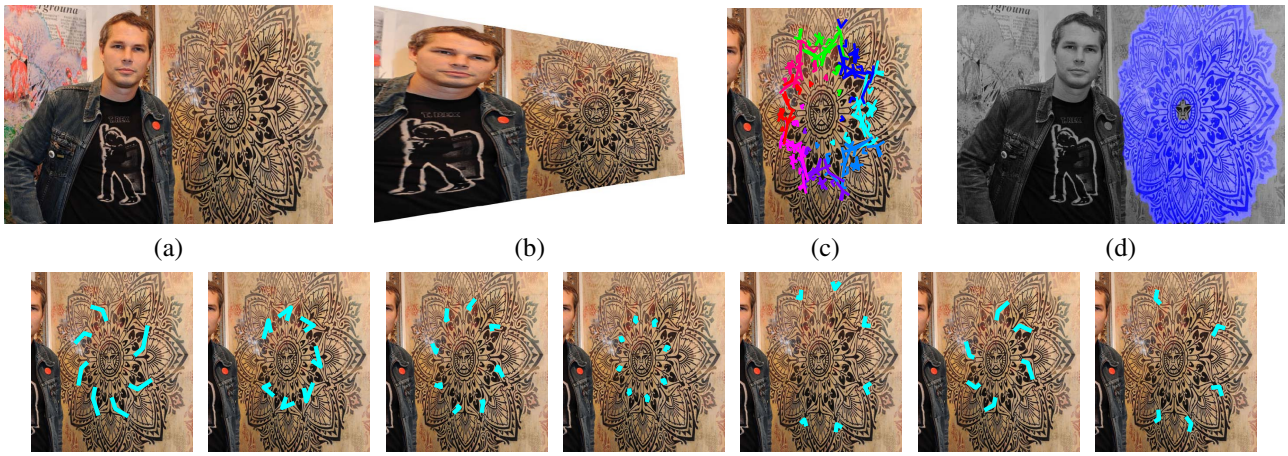
$$x_i h_7 + y_i h_8 - s_i^{1/3} \alpha_j = -1, \qquad (1)$$

where (x_i, y_i) are the coordinates of i-th feature point, s_i is the local scale of that feature, $(h_7, h_8, 1)$ are the coordinates



(a) (b) (c)

Figure 5. Tennis courts - rectification from instances of a reflected motif: (a) LAFs from the non-reflected motif instances, (b) LAFs from the reflected instances, matching LAFs color-coded, (c) the rectified image.



(a) (b) (c) (d)

Figure 6. Rectification from rotated instances of the motif. Top row: (a) the original image, (b) the image with the plane of the pattern rectified (cropped for visualization), (c) LAFs from eight motif instances superimposed over the original image, different instances color-coded, (d) repeated pattern segmented; bottom row: a sample of different sets of corresponding LAFs.

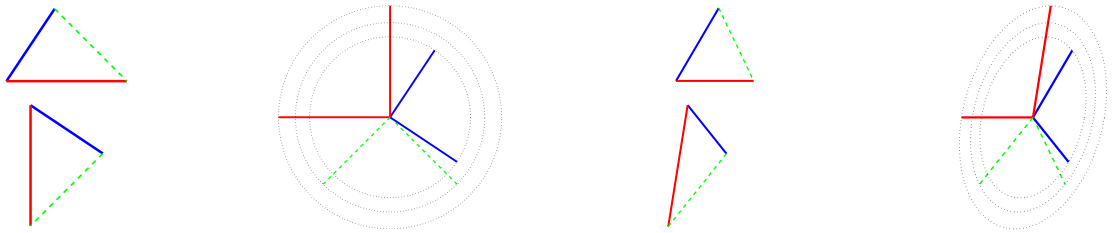
of the line at infinity, and α_j is related to the relative scale of different feature sets (indexed by j, if more than one set of matching features are used to estimate the line at infinity) after rectification. Three corresponding features in general position or three sets of two matching features are sufficient to estimate the line at infinity. For more details see [3].

Motif construction. Each SIFT cluster represents a set of intra-pattern tentative correspondences between motif elements that were detected as MSERs and represented as local affine frames (LAFs). The clustering implicitly assumes that local neighborhoods covered by LAFs have a large overlap with the motif and contain large subsets of the pattern. SIFTs of the largest clusters are used to define sets of overlapping neighborhoods in the pattern. Using the SIFTs' corresponding LAFs, which define local coordinate systems at the SIFT's origin, SIFT neighborhoods are transformed to the standard orthonormal basis, where surrounding features are aggregated in a common space. In the standard basis frame, spatially overlapping features are identified as repeated elements and added to the motif representation. This process is repeated for a fixed number of the top SIFT clusters. Mismatches are discarded by spatial verification in the canonical frame. Instances of a motif are shown in Fig. 6(c).

Affine distortion correction. The method of [3] provides planar rectification up to an unknown affine transformation. If the motif instances are reflected or rotated (which is easily detected even under an affine ambiguity), then it is possible to further reduce the ambiguity. The methods use constraints from the lengths of corresponding line segments and lead to systems of linear equations. The contribution is detailed in 3.1.

Non-linear optimization. The parameters of the generative model for the imaged pattern (Fig. 4) are initialized from the linear estimates of rectification and from the motif construction in rectified space. Subsequently, re-projection error in the image is minimized by a non-linear least squares optimizer (Levenberg–Marquardt). This step is detailed in 3.2.

Robustness. Since the intra-image matching is prone to false matches, estimation of rectification needs to be robust to outliers. RANSAC [6] is used to estimate the line at infinity and to estimate an further affine distortion reduction if reflected or rotated repeats are detected. The motif construction stage filters out sets of features that do not have sufficient support in their neighborhood, which is similar to the approach of [18].



(a) (b) (c) (d)

Figure 7. (a) repeated element is rotated and translated, (b) corresponding (color coded) vectors form circles, (c) repeated element transformed by an affine transformation, (d) corresponding vectors form ellipses that differ only by a diameter.

Segmentation of the pattern. Methods from image co-segmentation [5, 2] are used to estimate the pixel-accurate spatial mask of the repeated pattern. The constructed pattern is reprojected to determine intra-image mappings from corresponding locations of the motif instances. A small correlation window is moved over one instance of the pattern. The intensity function inside the window is correlated with the intensity in the window transformed into others instances of the motif. Locations with high correlation are considered repeated and thus belonging to the repeated structure. Regions without texture are not considered for correlation. For examples of repeated pattern localization, see Figs. 6 (d), 8(c) – (d), 9(c) and 13 (c).

3.1. Length constraints

The following paragraphs describe the process of upgrading an affine-rectified pattern obtained from the method of [3]. Unlike the area ratios of repeated elements, their length ratios are not invariant to an affine transformation; the length of vectors in different directions are affected differently by an affine transformation. The constraints provided by the length of corresponding line segments will be examined to estimate an affine transformation $\mathtt{A}$ that reduces the remaining affine ambiguity. Depending on the type of the repeated pattern construction, the ambiguity is reduced to a similarity, which corresponds to an arbitrary choice of an orthogonal isotropic coordinate system.

Since transforming a line segment by a translation has no effect on its length, lengths of free vectors will be studied. These vectors are obtained from the line segment AB by translating A to the origin as $\mathbf{x} = (x, y)^\top = B - A$. This means that the translation part of the unknown affine transformation $\mathtt{A}$ need not be considered; let $\bar{\mathtt{A}}$ be the 2×2 upper left sub-matrix of an affine transformation $\mathtt{A}$. Let the relation between the undistorted vectors $\hat{\mathbf{x}}$ and the observed vectors $\mathbf{x}$ be

$$\hat{\mathbf{x}} = \bar{\mathtt{A}}\mathbf{x}. \tag{2}$$

The derived constraints are linear, so they are simple to implement and efficient to use in a RANSAC-like [6] robust estimator. Fast estimation can be achieved from minimal sampling, or a more accurate least-squares solution can be obtained from many correspondences.

Symmetry. This paragraph examines the configuration of repeated elements that is reflected about a line of symmetry. Such a configuration frequently occurs on man-made objects, especially on building facades [20]. Let the two corresponding vectors from the reflected instances of the motif be denoted $\mathbf{x}$ and $\mathbf{x}'$ respectively. We will assume, without loss of generality, that the line of symmetry is a vertical line on the scene plane. Thus, we are looking for an affine transformation $\bar{\mathtt{A}}$ with rows $\mathbf{a_1}^\top$ and $\mathbf{a_2}^\top$, so that

$$\mathrm{diag}(-1, 1)\bar{\mathtt{A}}\mathbf{x} = \bar{\mathtt{A}}\mathbf{x}'. \tag{3}$$

This leads to a set of two homogeneous equations

$$\mathbf{a_1}^\top(\mathbf{x} + \mathbf{x}') = 0, \quad \text{and} \tag{4}$$

$$\mathbf{a_2}^\top(\mathbf{x} - \mathbf{x}') = 0. \tag{5}$$

A single reflected vector correspondence is enough to compute $\mathbf{a_1}$ and $\mathbf{a_2}$ up to a scalar factor. Any matrix with rows $\lambda_1\mathbf{a_1}^\top$ and $\lambda_2\mathbf{a_2}^\top$ for non-zero scalars $\lambda_{1,2}$ will rectify the vectors to a reflected position. Therefore, the rectification has an ambiguity of an anisotropic scaling along the direction of the axis of symmetry, the overall scale, and the rotation, which was fixed by aligning the axis of symmetry with a vertical line. This is equivalent to detecting two perpendicular vanishing points, one in the direction of the axis of symmetry.

Rotation. We will assume that on the scene plane, there are sets (indexed by j) of corresponding rotated line segments of length r_j, each set having at least two elements (indexed by i). The length constraint can be written as

$$\hat{\mathbf{x}}_{ij}^\top\hat{\mathbf{x}}_{ij} = \hat{r}_j^2. \tag{6}$$

The situation is depicted in Fig. 7(b). Corresponding vectors are color coded and Equation (6) is depicted as circles with diameter r_j. The length constraint (6) in the image transformed by $\bar{\mathtt{A}}^{-1}$ according to (2) changes to

$$\mathbf{x}_{ij}^\top\Sigma\mathbf{x}_{ij} = \hat{r}_j^2, \quad \text{where} \quad \Sigma = \bar{\mathtt{A}}^\top\bar{\mathtt{A}}. \tag{7}$$

In equation 7, Σ represents an ellipse (visualized in Fig. 7c), where

$$\Sigma = \begin{pmatrix} a & b \\ b & c \end{pmatrix}. \tag{8}$$

Equation (7) can be rewritten as

$$(x_{ij}^2 \quad 2x_{ij}y_{ij} \quad y_{ij}^2 \quad -1)(a \quad b \quad c \quad r_j^2)^\top = 0, \qquad (9)$$

which gives a system of homogeneous linear equations. There are three unknowns for Σ, and each set of matching line segments adds one unknown r_j. Each participating line segment in general position (rotation) adds one constraint. For two pairs of line segments, there are $3 + 2 = 5$ unknowns and four linear equations, yielding a one-dimensional linear space of solutions. An alternative minimal solution is given by one triplet of reflected line segments, having $3 + 1 = 4$ unknowns and 3 linear equations.

The affine transformation can be derived from the solution of the system of linear equations (9) up to a scale factor and a rotation. The unknown scale comes from the homogeneous nature of the system–both Σ and r_i^2s can be multiplied by a positive scalar. The rotational ambiguity comes from the ambiguity of the decomposition $\Sigma = \bar{\mathtt{A}}^\top \bar{\mathtt{A}}$.

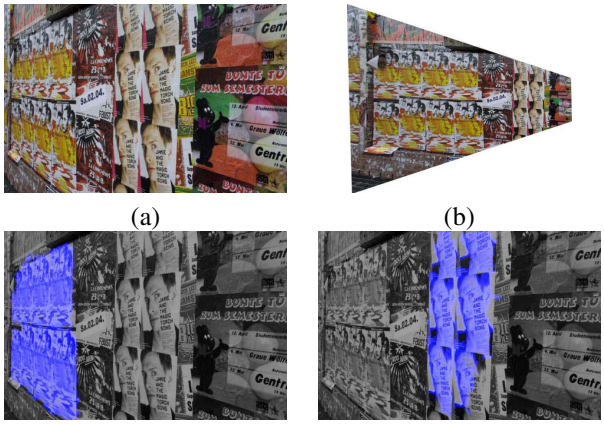
A rotation by 180 degrees (or an integer multiple) creates a special case: if the pattern is only rotated by integer multiplications of 180 degrees, then the matching vectors lie on parallel lines. Since affine transformations affect the lengths of vectors on parallel lines equally, the situation is similar to the pure translation case with full affine ambiguity.

3.2. Non-linear estimation

The parameters of the generative model depicted in (Fig. 4) are estimated by a non-linear minimization of pattern re-projection error. Scene plane rectification, position transformations and localization of detected features are jointly refined, and the camera model is extended to include radial distortion. Radial distortion is modeled by the division model [7]. The minimization is accomplished by Levenberg–Marquardt. The parameters of the optimization include the position of the feature points in the motif $\hat{\mathbf{v}}_i$, the parameters of the rectifying homography $\hat{\mathtt{H}}$, which is a composition of the non-projective upgrade (if there is one) and $\hat{\mathtt{H}}_\infty$, the parameters of the positioning transformations $\hat{\mathtt{T}}_j$ for each motif instance, and the radial distortion parameter $\hat{\lambda}$. A geometric re-projection error of the reconstructed pattern with the measured feature points in the original image counterpart is given by the following cost function

$$\min_{\hat{\lambda},\hat{\mathtt{H}},\hat{\mathtt{T}}_j,\hat{\mathbf{v}}_i} \sum_{i=1}^{N} \sum_{j:\exists \mathbf{u}_{ij}} d(\mathcal{L}_{\hat{\lambda}}(\hat{\mathtt{H}}^{-1}\hat{\mathtt{T}}_j(\hat{\mathbf{v}}_i)), \mathbf{u}_{ij}). \qquad (10)$$

The outer sum is over all N points of the motif, the inner sum is over instances j where the i-th motif point was observed, $d(\mathbf{x}, \mathbf{y})$ is the Euclidean distance between the images of $\mathbf{x}$ and $\mathbf{y}$, and $\mathcal{L}_{\hat{\lambda}}$ implements the radial distortion. The rectifying homography $\hat{\mathtt{H}}$ is minimally parameterized based on the geometric constraints available from the pattern arrangement (see 3.1).



(a) (b) (c) (d)

Figure 8. Poster wall: (a) original image. Two different (c) and (d) coplanar translated repeated pattern are detected. The plane is rectified (b) up to an affine transformation.

4. Experiments

In this section, we present experimental results that validate the broad applicability of the proposed method. The ambiguity of image registration is resolved by registering the rectified image to the original image using a transformation that corresponds to the rectification type: affine for an affine ambiguity, similarity for similarity ambiguity, etc.. This is done solely for the visualization purposes so that the rectified images are not significantly distorted.

4.1. Pure Translation

The case where the motif is only translated in the scene plane offers rectification only up to an arbitrary affine transformation of the scene plane; thus, it is the least interesting. Examples of repeated patterns with translated motifs are shown in Figs. 8 and 9.

4.2. Reflection

The presence of a reflected motif(s) allows for the detection of two perpendicular vanishing points, one of which is aligned with the axis of symmetry. Fig. 10 shows an example of a symmetric repeated pattern: the windows. The figure also provides a comparison to an affine rectification gotten by the method given in [3]. More examples are shown in Figs. 5, 1-top, Figs. 11. The motif in the FEET example in Fig. 11 is not self-symmetric.

4.3. Rotation

Fig. 2 compares the rectification results of different methods on the **target** image from [3], which contains a rotated repeated pattern. A rectification of an image of the famous Escher painting "**Tessellation 85**" is shown in Fig. 12. This experiment demonstrates the robustness of the proposed algorithm to noise. Since the painting is hand-drawn, each of the lizards, bats and fish are, in fact, unique.

(a) (b) (c)

Figure 9. Translated pattern: (a) original image, (b) rectified up to an affine transformation, (c) repeated pattern extent.

(a) (b) (c)

Figure 10. Reflected pattern: (a) original image of a schoolhouse, (b) affine rectification by [3], (c) similarity rectification by the proposed method. Images (a) and (b) courtesy of authors [3].

Figure 11. Footprints - rectification of a reflected pattern up to unknown scale of the x-axis.

Figure 12. Tessellation 85: Similarity rectification from a rotated pattern. The rectified image was cropped for visualization to handle the extreme elongation.

Still, a perfect similarity rectification is achieved. Further examples of similarity rectification from rotated motifs are given in Figs. 1 bottom and 6.

An example of rotated pattern with no geometric structure is shown in Fig. 13. The PAMI issue that is on the same plane as the pattern was rectified to have 88.4 degree angle between its sides and an aspect ratio of 2:2.77, while in reality it is 2:2.73 (the corners of the journal were manually clicked).

4.4. Improving the descriptor accuracy

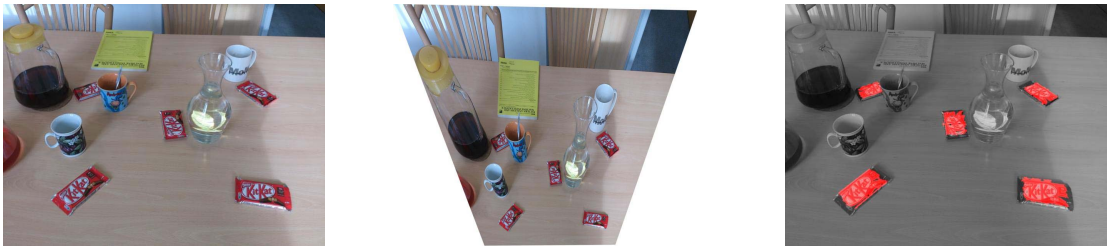
The non-linear minimization not only improves the estimate of the rectification transformation, but, more importantly, it estimates the noise free feature points. In order to quantify the gain, the following statistics were computed both before the rectification (*i.e.* over the detected features), and after the rectification (over the geometrically corrected features). For each feature in the motif, an average SIFT descriptor was computed from its instances. The average distance of feature SIFT descriptors to the corresponding average SIFT descriptor was computed. The rectification reduces the intra-cluster SIFT distance by 35.2% on average for the images in Figs.2, 8 ,9. This is an important result for applications such as image retrieval.

4.5. Multiple coplanar patterns

It is natural to consider multiple motifs for modeling images that contain multiple patterns. Fig. 8 shows two different groups of posters cosegmented from multiple motif detections. In this case, the instances are on the same plane, and the constraints on the line at infinity and the radial distortion are shared. If the instances of the motifs live on a different planes, then new parameters for the line at infinity for each plane are introduced, but the radial distortion parameter is the same for all the planes.

5. Conclusions

A novel and fully automated method for the detection, rectification and pixel accurate segmentation of coplanar repeated patterns is introduced. Besides coplanarity, the algorithm makes virtually no assumptions about the arrange-



(a) (b) (c)

Figure 13. Rotated pattern randomly repeated: (a) original image, (b) rectified up to similarity, (c) repeated pattern segmentation.

ment of the motif. A set of linear single-view geometry constraints are introduced to estimate rectification up to a similarity of the scene plane from a minimal set of intra-image correspondences—affine distortion can be removed from just one rotated or reflected repeated scene element. These constraints are used in a fast RANSAC framework to provide robust intra-pattern feature matching and pattern rectification.

The imaged pattern is modeled generatively. Model optimization concurrently refines rectification, pattern geometry and lens distortion. The model-based approach enables the reprojection of the constructed pattern, which is used to segment the pattern with pixel-level accuracy.

Experiments verify that the method works when the pattern is a small part of the image, of low texture, with an arbitrarily arranged motif and with a significant perspective warp. Thus the proposed method is applicable to a broad class of patterns, including man-made, planar structures with repetition or symmetry, such as building facades, mosaics, decorative prints and text.

References

[1] D. Aiger, D. Cohen-Or, and N. J. Mitra. Repetition maximization based texture rectification. *Computer Graphics Forum*, 31(2):439–448, 2012. 2

[2] J. Cech, J. Matas, and M. Perdoch. Efficient sequential correspondence selection by cosegmentation. In *CVPR*, 2008. 5

[3] O. Chum and J. Matas. Planar affine rectification from change of scale. In *ACCV*, 2010. 2, 3, 4, 5, 6, 7

[4] P. Doubek, J. Matas, M. Perdoch, and O. Chum. Image matching and retrieval by repetitive patterns. In *ICPR*, 2010. 2

[5] V. Ferrari, T. Tuytelaars, and L. Van Gool. Simultaneous object recognition and segmentation by image exploration. In *ECCV*, 2004. 5

[6] M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *CACM*, 24(6):381–395, June 1981. 4, 5

[7] A. W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *CVPR*, 2001. 6

[8] W. Hong, A. Yang, K. Huang, and Y. Ma. On symmetry and multiple-view geometry: Structure, pose, and calibration from a single image. *IJCV*, 60(3):241–265, December 2004. 2

[9] K. Köser, C. Beder, and R. Koch. Conjugate rotation: Parameterization and estimation from an affine feature correspondence. In *CVPR*, 2008. 2

[10] T. Leung and J. Malik. Detecting, localizing and grouping repeated scene elements from an image. In *ECCV*, 1996. 2

[11] D. Liebowitz and A. Zisserman. Metric rectification for perspective images of planes. In *CVPR*, 1998. 2

[12] Y. Liu, R. Collins, and T. Y. A computational model for periodic pattern perception based on frieze and wallpaper groups. *PAMI*, 26:354–371, 2004. 2

[13] D. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004. 3

[14] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *BMVC*, 2002. 3

[15] Š. Obdržálek and J. Matas. Object recognition using local affine frames on distinguished regions. In *BMVC*, 2002. 1, 3

[16] M. Park, R. Collins, and L. Y. Deformed lattice discovery via efficient mean-shift belief propagation. In *ECCV*, 2008. 2

[17] F. Schaffalitzky and A. Zisserman. Geometric grouping of repeated elements within images. In *BMVC*, 1998. 2

[18] C. Schmid and R. Mohr. Combining greyvalue invariants with local constraints for object recognition. In *CVPR*, 1996. 4

[19] T. Tuytelaars, A. Turina, and L. Van Gool. Noncombinatorial detection of regular repetitions under perspective skew. *PAMI*, 25(4):418–432, April 2003. 2

[20] C. Wu, J.-M. Frahm, and M. Pollefeys. Detecting large repetitive structures with salient boundaries. In *ECCV*, 2010. 2, 5

[21] C. Wu, J. M. Frahm, and M. Pollefeys. Repetition-based dense single-view reconstruction. In *CVPR*, 2011. 2

[22] Z. Zhang, A. Ganesh, X. Liang, and Y. Ma. TILT: transform invariant low-rank textures. *IJCV*, 99(1):1–24, 2012. 2