Planar Affine Rectification from Change of Scale

Ondřej Chum and Jiří Matas

CMP, Dept. of Cybernetics, Faculty of EE, CTU in Prague

Abstract. A method for affine rectification of a plane exploiting knowledge of relative scale changes is presented. The rectifying transformation is fully specified by the relative scale change at three non-collinear points or by two pairs of points where the relative scale change is known; the relative scale change between the pairs is not required. The method also allows homography estimation between two views of a planar scene from three point-with-scale correspondences.
The proposed method is simple to implement and without parameters; linear and thus supporting (algebraic) least squares solutions; and general, without restrictions on either the shape of the corresponding features or their mutual position.
The wide applicability of the method is demonstrated on text rectification, detection of repetitive patterns, texture normalization and estimation of homography from three point-with-scale correspondences.

1 Introduction

The problem[1] of affine rectification of a plane, *i.e.* the problem of transforming an image by a homography so that the vanishing line of the plane becomes the line at infinity, arises in many applications, *e.g.* in document processing [1, 2], detection of repetitive structures [3] and texture analysis [4, 5]. The plane of interest appears in the rectified images as if viewed by an affine camera, *i.e.* projected by a set of parallel rays and scaled. The restoration of affine properties like parallelism and global scale simplifies subsequent application-dependent processing steps like geometric normalization, detection and recognition.

In the paper, a general yet simple method for affine rectification of a plane is introduced. The algorithm exploits knowledge of relative scale changes in the local neighbourhood of image points lying in the plane. The rectifying transformation is fully specified by the relative scale change at three non-collinear points. Another minimal case covered by the method applies in the situation were for two pairs of points the relative scale change is known; the relative scale change between the pairs is not required.

A situation in which the relative scale change is known at different points arises often in practice. Consider, *e.g.* the problem of affine rectification of a repeated pattern on a planar surface, say a facade, Fig. 1. In a perspective image of the facade, the features detected on the windows in general vary in size (area). In reality, it is common that (at least some of) the windows are of the same size. The task addressed in the paper is to find a planar homography H that transforms the image of the facade so that all
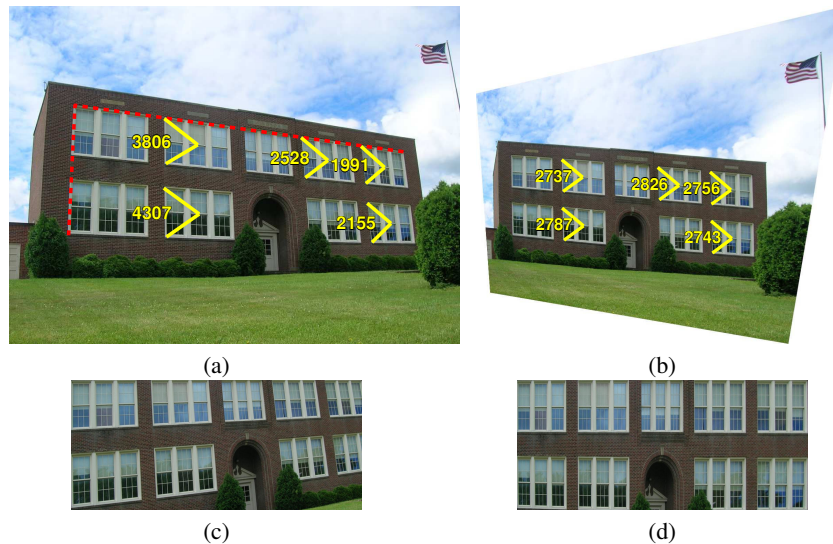

Fig. 1. Affine rectification. Original image (a) - the area of the triangular patches differs from 1991 to 4307 pixels due to the perspective projection. Rectified image (b) - the areas are approximately the same, as in reality. Parallel lines on the facade are not parallel in the original image (c), and are parallel after the normalization (d). The cut-outs (c) and (d) are parallelograms defined by two red line segments in (a).

the window features cover the same area. We show in the experimental section that the method is applicable in many situations.

The proposed method has the following advantages: (generality) no assumptions are made about either the shape of the features or their mutual position; features need not lie on a regular grid nor on lines and may be arbitrarily rotated; (stability) the rectification is computed from ratios of areas, a very stable property insensitive to many image degradations such as discretization; (simplicity) the rectification algorithm is simple, easy to implement and without parameters; (linearity) the constraints on the scale change are expressed as linear constraints on the entries of the homography matrix H that represents the transformation. Linear constraints are very convenient as they can be used with minimal sets (in RANSAC-like [6] robust estimators) as well as in (algebraic) least squares solutions from all available data.

The derivation of the algorithm assumes that the features are sufficiently small so that their scale change reasonably approximates the scale change (of an infinitesimal patch) at corresponding points. Such an assumption is made by wide-baseline matching approaches using affine covariant feature points and/or affine invariant feature descriptors. We show experimentally in Sec. 3 that the assumption holds in practice.

Previous work. Affine rectification algorithms proposed in the literature differ by the assumptions about the structures present in the image that are exploited in the process. The most straightforward approaches detect two distinct vanishing points [7].

The problem of vanishing line detection has been addressed for elements repeated by translation on a plane. The geometric relation of the elements after projective transformation is called *elation* [7]. A comprehensive study of vanishing line (and points) detection based on the elation assumption is given in [3]. Another approach exploiting elations for detection of vanishing line in a projective image of a texture was proposed in [8]. Other approaches, specially in the text analysis, assume, that parallel lines with equal spacing can be detected in the image. The normalization (vanishing line) is then estimated from the intersection of the parallel lines and a cross-ratio of collinear set of points on those lines [1].

Publications on affine rectification have appeared in the field of shape-from-texture [9]. In general, assuming homogeneity of the texture, more complex structure than orientation of a plane can be estimated [10]. However, a fairly complex optimization approach is necessary in this case. There are many approaches to vanishing point and/or line detection from the texture. Voting schemes based on dominant direction of the texture can be used to determine a vanishing point [11]. In [12], another voting scheme based on distortion of the power spectrum under projective transformation is used detect the vanishing line.

Similar idea to ours has appeared in Ohta's 1981 paper [13] on shape from texture. Despite the different derivations the results are closely related. In fact the formulation in [13] is a special case of ours. Our formulation allows to extend the applicability of the idea beyond a planar rectification, for example to multi-view geometry. Our derivation yields a single linear constraint per feature while Ohta's approach produces one linear constraint per a pair of textured regions. Finally, we show significantly higher applicability than [13] or its extension [14], for example the features of interest (or texture) does not have to cover the whole image.

The rest of the paper is organized as follows. First, the method is derived in sections 2 and 2.1. Extension to multiple independent feature sets is introduced in section 2.2. Experiments and applications of the proposed method to various tasks are presented in section 3: simple examples of the minimal cases 3.1, text rectification 3.2, non-linear repeated structures 3.3, segmentation of multiple planes with repeated pattern 3.4, texture rectification 3.5, and experiments on synthetic data 3.6 and 3.7. The applicability of the approach to image to image homography estimation from point-with-scale correspondences is discussed in section 3.8. Conclusions are drawn in section 4. A proof of degenerate case of collinear points can be found in appendix A.

2   The Method

First, the concept of local scale change under planar homography is introduced and its properties are discussed. Next, a decomposition of a homography simplifying the algebra is presented. Finally, it is shown that constraints on the local scale change under planar homography (*i.e.* perspective transformation of a plane) lead to linear constraints on the entries of the homography matrix.

Homography is a mapping from a projective plane P^2 to P^2 and it is commonly represented by a (homogeneous) matrix H, or equivalently, by inhomogeneous pair of

functions (h_x, h_y) [7]. In this section, we restrict the homographies to be in the following form

$$
\mathtt{H} = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{pmatrix} \qquad \text{or} \qquad
\begin{aligned}
h_x(x, y) &= \frac{h_1 x + h_2 y + h_3}{h_7 x + h_8 y + 1}, \\
h_y(x, y) &= \frac{h_4 x + h_5 y + h_6}{h_7 x + h_8 y + 1}.
\end{aligned}
\tag{1}
$$

The sufficiency of the $\mathtt{H}_{3,3} = 1$ parametrization is discussed and justified in section 2.1. The first order Taylor expansion at point (x, y) and the Jacobian $\mathtt{J}_\mathtt{H}$ locally approximating the homography

$$
h(x + \delta_x, y + \delta_y) \approx \begin{pmatrix} h_x(x, y) \\ h_y(x, y) \end{pmatrix} + \mathtt{J}_\mathtt{H}(x, y) \begin{pmatrix} \delta_x \\ \delta_y \end{pmatrix}
\tag{2}
$$

is an affine transformation for which the concept of scale change is well defined. The local scale change at point (x, y) under the perspective transformation is thus defined as the scale change of the first order, *i.e.* affine, approximation at point (x, y)

$$
s(\mathtt{H}, x, y) = \det(\mathtt{J}_\mathtt{H}(x, y)).
\tag{3}
$$

Any homography $\mathtt{H}$ in the form of (1) can be decomposed into a product $\mathtt{A}\hat{\mathtt{H}}$ of an affine transformation $\mathtt{A}$ and a homography $\hat{\mathtt{H}}$ as follows

$$
\begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{pmatrix} = \begin{pmatrix} h_1 - h_3 h_7 & h_2 - h_3 h_8 & h_3 \\ h_4 - h_6 h_7 & h_5 - h_6 h_8 & h_6 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ h_7 & h_8 & 1 \end{pmatrix}
\tag{4}
$$

It can be shown that the scale change of homography $\mathtt{H}$ expressed in terms of $\mathtt{A}$ and $\hat{\mathtt{H}}$ is

$$
s(\mathtt{A}\hat{\mathtt{H}}, x, y) = \det(\mathtt{A}) s(\hat{\mathtt{H}}, x, y).
$$

The advantage of the decomposition (4) is that the influence of parameters $h_1 \ldots h_6$ on the local scale change is reduced to a single global (*i.e.* position-independent) parameter $\det \mathtt{A}$ in the expression

$$
s(\mathtt{H}, x, y) = \det(\mathtt{A}) s(\hat{\mathtt{H}}, x, y) = \det(\mathtt{A}) \det(\mathtt{J}_{\hat{\mathtt{H}}}(x, y)).
\tag{5}
$$

The determinant of the Jacobian of the matrix $\hat{\mathtt{H}}$ at (x, y) is

$$
\det(\mathtt{J}_{\hat{\mathtt{H}}}(x, y)) = \det \left((h_7 x + h_8 y + 1)^{-2} \begin{pmatrix} h_8 y + 1 & -x h_8 \\ -y h_7 & h_7 x + 1 \end{pmatrix} \right) = (h_7 x + h_8 y + 1)^{-3}.
$$

Setting $\det(\mathtt{A}) = \alpha^3$ and substituting into equation (5), we get

$$
s(\mathtt{H}, x, y) = \alpha^3 (h_7 x + h_8 y + 1)^{-3}.
\tag{6}
$$

After re-arranging the equation, a constraint linear in h_7, h_8, and α is obtained:

$$
\begin{pmatrix} x & y & -s(\mathtt{H}, x, y)^{-1/3} \end{pmatrix} \begin{pmatrix} h_7 & h_8 & \alpha \end{pmatrix}^\top = -1.
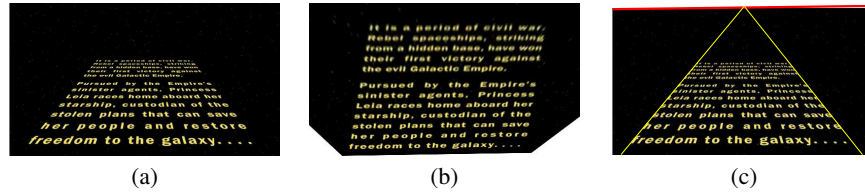\tag{7}
$$

Fig. 2. Star Wars credits: (a) original image, (b) estimated (up to an affine transformation) normalized image, (c) original image with the estimated vanishing line (red) and manually drawn the parallel margin lines (yellow).

Three point locations (x_i, y_i) and the corresponding local scale changes $s(\mathtt{H}, x_i, y_i)$ are required to estimate the homography $\hat{\mathtt{H}}$. Any composition of affine transformation $\mathtt{A}$, $\det(\mathtt{A}) = \alpha^3$ and the homography $\hat{\mathtt{H}}$, *i.e.* $\mathtt{H} = \mathtt{A}\hat{\mathtt{H}}$ will satisfy the constraints on the local scale change. The vanishing line l in the source image is the pre-image of the line at infinity $(0\ 0\ 1)^\top$

$$l = \mathtt{H}^\top (0\ 0\ 1)^\top = \hat{\mathtt{H}}^\top \mathtt{A}^\top (0\ 0\ 1)^\top = \hat{\mathtt{H}}^\top (0\ 0\ 1)^\top = (h_7\ h_8\ 1)^\top. \tag{8}$$

If $p, p > 3$, points with the local scale change are available, the least squares method is applicable. The data matrix $\mathtt{Z} \in R^{p \times 3}$ is composed of rows

$$\mathtt{Z} = \begin{pmatrix} x_i & y_i & -s(\mathtt{H}, x_i, y_i)^{-1/3} \\ \vdots & \vdots & \vdots \end{pmatrix}, \tag{9}$$

one per each point (x_i, y_i). The solution is then obtained as

$$(h_7 \quad h_8 \quad \alpha)^\top = -\mathtt{Z}^\dagger \mathbf{1}^{1 \times p}, \tag{10}$$

where $\mathtt{Z}^\dagger$ is pseudo-inverse of $\mathtt{Z}$ and $\mathbf{1}^{1 \times p}$ is a column vector of p ones.

In many applications, the scale change is not interesting or not known and only relative scale changes at different points are known. Here, the estimated parameter α can be simply ignored. This is *e.g.* the case for the facade example Fig. 1, where the windows are assumed to have the same, but unknown, real size. In such cases, the $s(\mathtt{H}, x, y)$ is multiplied by an unknown scalar.

2.1   The choice of parametrization

The chosen parametrization of matrix $\mathtt{H}$ in section 2 does not cover all possible homographies. Namely, it does not include the set of homographies $\mathcal{H}_0$ with $\mathtt{H}(3, 3) = 0$, *i.e.* homographies that map the origin of the image coordinate system $(0\ 0\ 1)^\top$ to a point at infinity. Hence, if it is possible to choose the origin so that it is guaranteed that the required solution does not map the origin to infinity, the $\mathtt{H}(3, 3) = 1$ parametrization is correct.

A frequent choice of the origin of the image coordinate system – the (top left) corner of the image – does not always guarantee the above described property. In particular, in
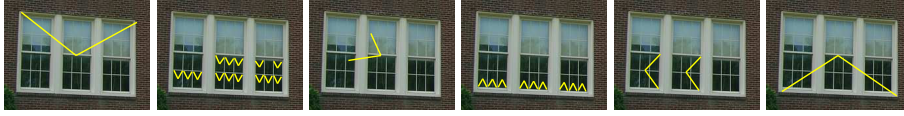
Fig. 3. An example of multiple features on an element of repeated pattern.

Fig. 2, the top left corner lies on the vanishing line mapped to a line at infinity by the affine rectifying homography.

The origin must not lie on the vanishing line, as the estimated transformation sends the vanishing line to infinity. Since the algorithm is used for affine rectification which is equivalent to detection of the vanishing line, based on scale change of measured features, good candidates for the origin are the measured points. This stems from the fact, that the point and its relative finite scale change could not have been measured at the line at infinity.

More generally, since the traditional (directional) camera sees only points in front of the camera [15], the vanishing line cannot 'cut through' the observed points. Therefore, any point inside the convex hull of the observed points will serve well as the origin of the coordinate system. The centre of gravity of the observed points was used in our implementation.

Note on the data normalization. In the least squares problem, some algebraic error (with no direct geometric meaning) is minimized. It has been shown that in such problems, it is advantageous to normalize the data points so that the elements of the measurement matrix Z have similar magnitudes [16]. Choosing the origin at the centroid of the data, re-scaling the data and suitable selection of the relative scale change prior to evaluation eqn. (10) can be used to stabilize the least squares solution.

2.2   Extension to multiple independent sets

As mentioned above, often only the relative scale change between a set of points is known. This section addresses the situation where multiple such sets are available. The relative scale change is known within each set, the relations between different sets is unknown.

As an example, let us have a look at the repetitive structures again. In general, the features detected in the image do not correspond one to one to the repeated elements. Typically, each element is covered by a number of features, as in Fig. 3. This number is also varying, as the repeatability of the features (as well as the stability of the descriptors) is not perfect. For each individual set of matching features, one can hypothesize that these are of the same size in reality, since else it is unlikely for the appearance of two patches to match. However, the area ratio of different features is not known in general.

For the sake of clarity, the derivation is demonstrated for two sets only. The extension to a general number of such sets is straightforward. There are two unknowns h_7 and h_8 shared between all the sets. Each set introduces an additional variable α_k. The variable represents the relative scale change of the whole set with respect to other sets.

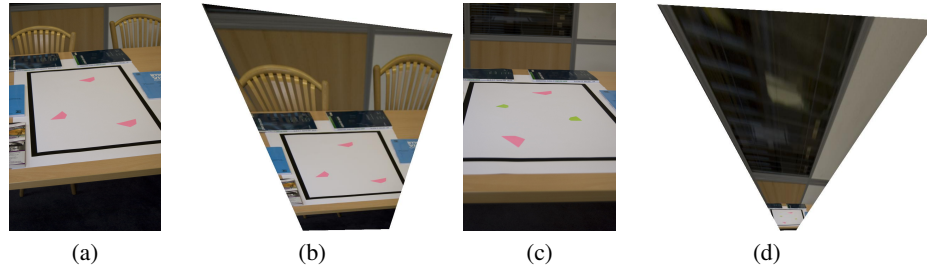(a)   (b)   (c)   (d)

Fig. 4. Toy examples for the minimal cases of three points (a-b) and two plus two points (c-d).

The equations are then arranged in the same way as in equation (7):

$$\begin{pmatrix} x_i \ y_i \ -s(\mathtt{H}, x_i, y_i)^{-1/3} & 0 \\ x_j \ y_j & 0 & -s(\mathtt{H}, x_j, y_j)^{-1/3} \end{pmatrix} \begin{pmatrix} h_7 \ h_8 \ \alpha_1 \ \alpha_2 \end{pmatrix}^\top = \begin{pmatrix} -1 \\ -1 \end{pmatrix}. \qquad (11)$$

Each feature in a set adds one constraint, at least two features have to be available for each set to add more constraints than unknowns. In general, if there are p points in q sets, there are $2 + q$ unknowns and p constraints. For two sets, two points per set are sufficient to estimate the rectifying transformation. For an example, see section 3.1.

3   Experiments

In this section a variety of experiments with different settings are presented.

3.1   Toy example

Two images of coplanar patches – Fig. 4(a) and (c) – are used to demonstrate the minimal cases described in sections 2 and 2.2. Very simple colour segmentation was used to locate the pink and green patches. The patches were represented by their location (the centre of gravity) and the scale (the number of pixels occupied by the patches). To simulate the two cases of minimal sets, the experiments was designed as follows: 1. the pink patches are of the same size, 2. the green patches are of the same size, and 3. the relative sizes of the pink and green patches are unknown.

The rectified images – Fig. 4(b) and (d) – show that the part of the scene that has been reduced by the projective transformation (further away from the camera) is expanded by the normalization. Also note that after the normalization, the parallel lines on the sheet of paper are again (very close to) parallel.

3.2   Text Rectification

In text localization and recognition in photographs taken in unconstrained conditions, geometric rectification is performed before classification of characters. The algorithm proposed in the paper is significantly simpler and more general than the approaches
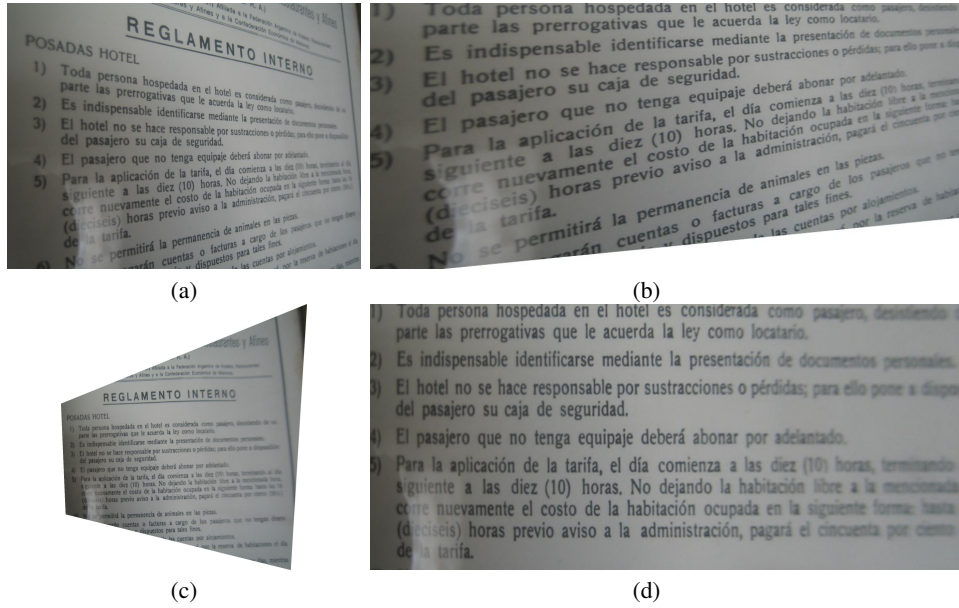
Fig. 5. Text rectification. (a) original image, (b) manual rectification using an affine transformation, (c) automatic affine rectification (d) manual rectification using an affine transformation after removing the perspective.

commonly used in document processing for affine rectification, *e.g.* [2] who requires a reliable procedure for fitting a baseline and topline of the text.

Applicability of the proposed procedure to the text rectification problem is demonstrated in Fig. 5. The top-left image (b) shows that affine normalization, is insufficient, non-parallel lines in the original image (a) say non-parallel. Fig. 5(c) shows the results of the proposed algorithm. The correspondences necessary for estimation of the relative scale are obtained fully automatically on identical characters by the MSER+LAF method [17]. Outliers and out-of-plane pairs are removed by RANSAC. The rectification based on tens of scale ratios is quite precise, see Fig. 5(d) which is an affine transformation of the rectified of image (c). The final affine rectification was done manually as it is not the topic of the paper - the proposed algorithm has no concept of a line of text or left margin; an example of an automatic method is in [2].

3.3   Darts

The "Darts" image, Fig. 6, is an example where direct detection of vanishing points and hence the vanishing line is difficult. The dominant linear features on the board that intersect in the bull's eye have different orientations and intersect the vanishing line in different ideal points.

The proposed method estimates the rectifying transformation from multiple sets of corresponding features, Fig. 6(e). Notice that the correspondences are between features with different orientations and not lying on straight lines. After (manually) mapping the
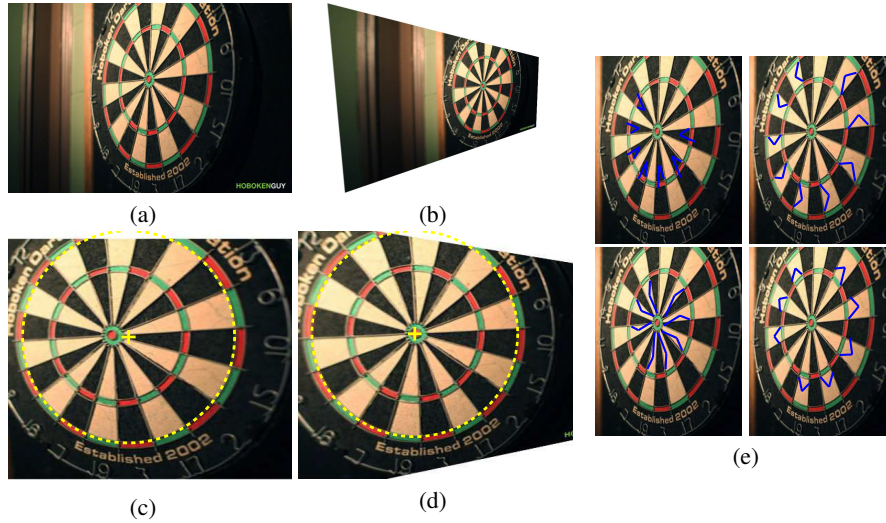
Fig. 6. Darts: (a) the original image; (b) automatic rectification; (c) manual rectification to a circle (dashed, centre labelled with '+') by an affine transformation from the original image, (d) the rectified image; (e) some of the matching feature groups superimposed.

the ellipse corresponding to the inner rim of the double scoring area, a fronto-parallel view of the board is obtained. The centre of the dotted yellow circle is very close to the centre of the bulls eye.

A direct affine mapping of ellipse corresponding to the inner rim in (a) to a circle results in image (c), but this view does not correspond to a fronto-parallel view of the board.

3.4   Segmentation of multiple planes with repeated pattern

The proposed method is not restricted to a single planar rectification. With RANSAC, a robust estimator, it is possible to separate features on a single plane from outliers. In the presence of multiple models (in our case multiple planes), consecutive execution of RANSAC with removal of features consistent with detected model [18] provides an efficient strategy.

In the simplest case, the two (or multiple) planes would not share features (different buildings, etc.). The example in Fig. 7 is more challenging, it shows that even if the planes share a common repetitive pattern and therefore the MSER+LAF method establishes correspondences between the two planes, the geometric constraints on the relative scale change are sufficiently discriminative to segment the planes.

3.5   Textures

The proposed method is also applicable to irregular statistical textures. In Fig. 8, an example of affine texture rectification is shown. For statistical textures, the MSER and

Fig. 7. Two planar surfaces with a repetitive pattern segmented by RANSAC. Left: one group of matching features, inliers to one model in red, the other in green, outliers to both models in yellow. Right: the convex hull of consistent features.
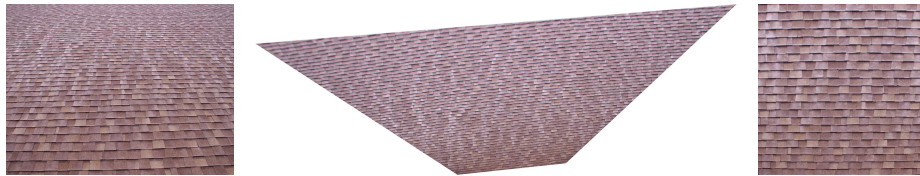

Fig. 8. Texture rectification: the original image (left), affine rectification (middle), and the rectified texture (right).

LAF method is not suitable since it requires that corresponding regions are geometrically close to identical. In the example, an affine covariant elliptical region detector [19] together with the SIFT descriptor [20] was used.

3.6   Scale change from local patches

One of the inputs of the proposed method is a scale change of an infinitesimal patch. However, it is typically only possible to measure the scale (change) at image patches that are of area of tens of pixels. In the first experiment, we measure how the estimate of the scale change affects the results.

First, A pattern of 5×5 local affine patches is generated. It is transformed by a homography with varying values of h_7 and h_8. The pattern is then resized and translated to fit a 800×600 image. Examples of four patterns are shown in Fig. 9. All situations in the experiment from the 'convex hull' of these four examples.

Each synthetic image was processed as follows. Each local affine patch was represented by the centre of gravity of the triangle and by the scale (area) of the triangle. A normalizing homography that transforms all patches to equal scale was estimated using the proposed method. In an ideal case, when the infinitesimal scale change is estimated exactly, all transformed patches would have exactly the same scale. The ratio of maximal resulting scale to the minimal resulting scale was recorded for each parameter setting. The results are visualized in Fig. 10.
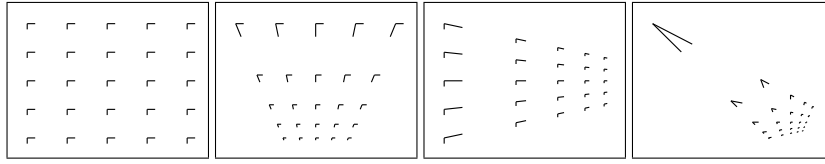
Fig. 9. Four examples of different levels of perspective deformations used in the synthetic experiments. All images are 800×600 pixels.
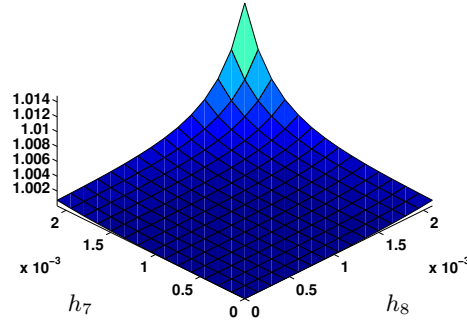

Fig. 10. Scale error after affine rectification. Ratio of the largest and the smallest feature after rectification to equal size.

It can be seen that even for extreme perspective deformations, the local scale is estimates sufficiently precisely and the ratio of areas of the largest and the smallest normalized patches is close to one. If necessary, the procedure can be iterated to eliminate the effect of the inaccurate estimation of the scale change. In the above experiment, after first iteration, the scale ratio of the areas of the largest to the smallest normalized patches was one up to numerical precision.

3.7 Sensitivity to noise

This experiment also uses the settings from Fig. 9. Here, the transformed patches (the coordinates of the triangle corner points) were corrupted by additive Gaussian noise with $\sigma = 1.5$ pixels. Robust rectifying homography estimation via RANSAC was applied and three quantities were measured. First, how well the estimated homography rectifies the noise-less patches. The number of correctly rectified noise-less patches (the scale change error below 1.1) is shown in Fig. 11 (a). Second, the number of RANSAC inliers is shown in Fig. 11 (b). The number of inliers is well correlated with the number of correctly rectified noiseless patches. Third, the average scale error on all noiseless patches (not only inliers) is depicted in Fig. 11 (c). All plots are averages over 50 executions of RANSAC.
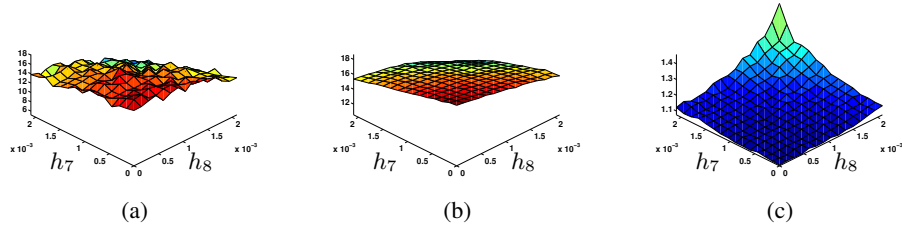
Fig. 11. Affine rectification estimated from regions corrupted by noise. Each plot shows 225 different settings of the parameters projective deformation h_7 and h_8: (a) noiseless features with scale error below 1.1, (b) number of RANSAC inliers, (c) average scale error on noiseless features; results averaged over 50 executions.

3.8 Image to image homography

Another straightforward application of the proposed method is the estimation of image to image planar homography form scale-covariant features, such as the DoG [20]. Only three correspondences are required to estimate the full projective homography. First, the projective part Ĥ is estimated from the scale change between the tree corresponding features. The affine part A is then given by the coordinates of the corresponding features in the two images. Sampling three instead of four points in RANSAC speeds up the robust estimation process, if scale information is available which is the case for scale and affine covariant features.

Furthermore, one non-linear constraint is available. It is not used in the estimation and can be used to verify that a homography matching the three point-with-scale correspondences induces the correct scale change. This constraint is the scale of the affine transformation A. The scale of the affine part, given by $\det(\mathtt{A})$, is obtained during the estimation of the projective part as α^3 in eqn. (7). Using the constraint in RANSAC, a number of contaminated samples can be rejected without the necessity of calculating consensus set size.

Images used in the experiment are a subset of a standard dataset [19], see Fig. 12. A combination of DoG features with the SIFT descriptor [20] was used, followed by a RANSAC with a local optimization step [21]. The comparison of the number of RANSAC samples in the homography estimation is shown in Table 1; results were averaged over 50 executions. The results show that using three point-with-scale correspondences allows to estimate the homography in significantly lower number of samples. If the scale consistency check is applied (the threshold was set to 1.1 in the experiments), the consensus is computed for only a small fraction of the samples – the last column of Table 1. On the other hand, the three point-with-scale samples provide a little less stable performance (slightly lower average of estimated inlier ratios) than four point correspondences.

4   Conclusions

A simple yet generally applicable method for affine rectification of a plane exploiting knowledge of relative scale changes was presented. The method also allows estimat-
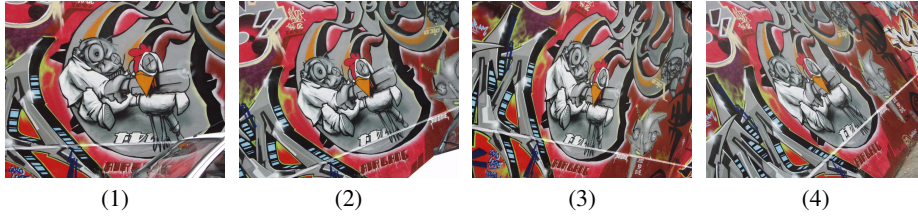
(1)   (2)   (3)   (4)

Fig. 12. Graffiti images.

image pair	tentative	% 4-inliers	% 3-inliers	4-samples	3-samples	3-valid
1–2	877	61.09	61.13	31.4	18.3	6.3
1–3	694	33.81	32.55	356.7	151.3	19.8
1–4	493	12.48	10.99	20861.0	4207.8	148.9
2–3	988	52.98	52.42	57.0	30.5	8.8
2–4	732	30.17	28.57	565.0	209.9	22.2
3–4	1043	61.74	61.27	30.1	18.2	5.2

Table 1. Comparison of the image to image homography estimation from samples of four point correspondences and three point-with-scale correspondences. The number of tentative correspondences ('tentative'), percentage of inliers detected by the methods ('% 4-inliers' and '% 3-inliers'), the number of samples required in RANSAC ('4-samples' and '3-samples'), and the number of scale consistent samples in the point-with-scale method ('3-valid'). Results averaged over 50 executions.

ing the homography between two views of a planar scene from three point-with-scale correspondences. A significant speed-up was achieved w.r.t. the standard four point procedure.

The utility of the method was demonstrated on text rectification, detection of repetitive patterns, texture normalization and estimation of homography from three points-with-scale correspondences.

References

1. Clark, P., Mirmehdi, M.: Rectifying perspective views of text in 3d scenes using vanishing points. Pattern Recognition **36** (2003) 2673–2686
2. Myers, G.K., Bolles, R.C., Luong, Q.T., Herson, J.A., Aradhye, H.: Rectification and recognition of text in 3-d scenes. IJDAR **7** (2005) 147–158
3. Schaffalitzky, F., Zisserman, A.: Planar grouping for automatic detection of vanishing lines and points. Image and Vision Computing **18** (2000) 647–658
4. Ribeiro, E., Hancock, E.R.: 3-d planar orientation from texture: Estimating vanishing point from local spectral analysis. In Carter, J.N., Nixon, M.S., eds.: BMVC, British Machine Vision Association (1998)
5. Lelandais, S., Boutté, L., Plantier, J.: Shape from texture: Local scales and vanishing line computation to improve results for macrotextures. Int. J. Image Graphics **5** (2005) 329–350
6. Fischler, M., Bolles, R.: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. CACM **24** (1981) 381–395

7. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Second edn. Cambridge University Press, ISBN: 0521540518 (2004)
8. Criminisi, A., Zisserman, A.: Shape from texture: homogeneity revisited. In: Proc. BMVC., UK (2000) 82–91
9. Witkin, A.: Recovering surface shape and orientation from texture. Artificial Intelligence **17** (1981) 17–45
10. Forsyth, D., Ponce, J.: Computer Vision: A Modern Approach. Prentice Hall (2003)
11. Rasmussen, C.: Texture-based vanishing point voting for road shape estimation. In: Proc. BMVC. (2004)
12. Ribeiro, E., Hancock, E.: Estimating the perspective pose of texture planes using spectral analysis on the unit sphere. Pattern Recognition **35** (2002) 2141–2163
13. Ohta, Y., Maenobu, K., Sakai, T.: Obtaining surface orientation from texels under perspective projection. In: IJCAL, Vancouver, Canada (1981) 746–751
14. Aloimonos, Y.: Shape from texture. Biological Cybernetics **58** (1988) 345–360
15. Hartley, R.: Chirality. IJCV **26** (1998) 41–61
16. Hartley, R.: In defence of the 8-point algorithm. In: ICCV95. (1995) 1064–1070
17. Obdržálek, Š., Matas, J.: Object recognition using local affine frames on distinguished regions. In: Proc. BMVC. (2002) 113–122
18. Torr, P.H.S.: Outlier Detection and Motion Segmentation. PhD thesis, Dept. of Engineering Science, University of Oxford (1995)
19. Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., Van Gool, L.: A comparison of affine region detectors. IJCV **65** (2005) 43–72
20. Lowe, D.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60** (2004) 91–110
21. Chum, O., Matas, J., Obdržálek, Š.: Enhancing RANSAC by generalized model optimization. In: Proc. of the ACCV. Volume 2. (2004) 812–817

A Degenerate case

Assume three collinear points (x, y), $(x + \alpha d_x, x + \alpha d_y)$, and $(x + \beta d_x, y + \beta d_y)$. Let the h_7 and h_8 be the parameters of the decomposition of the normalizing homography by eqn. 1. Then, the 3×3 data matrix Z from eqn. (9) has the following form

$$
\mathsf{Z} = \begin{pmatrix} x & y & h_7 x + h_8 y + 1 \\ x + \alpha d_x & y + \alpha d_y & h_7(x + \alpha d_x) + h_8(y + \alpha d_y) + 1 \\ x + \beta d_x & y + \beta d_y & h_7(x + \beta d_x) + h_8(y + \beta d_y) + 1 \end{pmatrix}
$$

The matrix Z is singular with vector $\mathbf{n}$

$$
\mathbf{n} = (-h_7 x d_y + h_7 d_x y - d_y, h_8 y d_x - x h_8 d_y + d_x, x d_y - d_x y)^\top
$$

spanning the null space of Z. The vector $\mathbf{h} = (h_7, h_8, 1)^\top$ solves the equation $Z\mathbf{h} = -\mathbf{1}$. Hence, there is a one-dimensional family of solutions $\mathbf{h} + \lambda \mathbf{n}$. It corresponds to a pencil of lines $\mathbf{h} + \lambda \mathbf{n_0}$, where

$$
\mathbf{n_0} = (-h_7 x d_y + h_7 d_x y - d_y, h_8 y d_x - x h_8 d_y + d_x, 0)^\top.
$$

All lines in the pencil pass through a point $\mathbf{h} \times \mathbf{n_0}$, which is the vanishing point lying on a line given by the collinear points.