DETECTION OF REPEATED STRUCTURES IN FACADE IMAGES

S. Wenzel, M. Drauschke, W. Förstner

Department of Photogrammetry, Institute of Geodesy and Geoinformation, University of Bonn, Nussallee 15, 53115 Bonn, Germany, {susanne.wenzel, martin.drauschke}@uni-bonn.de, wf@ipb.uni-bonn.de

We present a method for detecting repeated structures, which is applied on facade images for describing the regularity of their windows. Our approach finds and explicitly represents repetitive structures and thus gives initial representation of facades. No explicit notion of a window is used, thus the method also appears to be able to identify other man made structures, e.g. paths with regular tiles.

A method for detection of dominant symmetries is adapted for detection of multiply repeated structures. A compact description of the repetitions is derived from the detected translations in the image by a heuristic search method and the criterion of the minimum description length.

Introduction

Symmetric and repeated structures are typical properties of man made objects. Thus, finding such features in a scene may be indicative of the presence of man made objects. Additionally, a compact description of the found regularities can be suitable as mid-level feature for model-based learning.

Therefore, the intention of this work is, firstly, to check if there are any regularities, and secondly, to infer a compact description of the repeated structure. The description consists of a hierarchy of translations and their appropriate numbers of repetitions.

A typical facade is characterized by perpendicular regularities in horizontal and vertical directions. Therefore, we work on rectified images of facades and, hence, we may limit our description on horizontal and vertical directions. The more general approach in [9] can be used to overcome this limitation.

A lot of work on the detection of repetitive structures in images has been published within the last years. Leung and Malik [5] grouped repeated elements in the context of texture processing, allowing a similar transformation between the items. Another texture-based approach has been proposed by Hays et al. [4], where they map repetitive structures of texels within an iterative procedure.

Schaffalitzky and Zisserman [9] presented a grouping strategy for repetitive elements which are connected by an affine transformation. Tuytelaars et al. [10] can detect regular repetitions under perspective skew. All mentioned works are limited on the constraint that the repetition by the elements can be described by a single 2-dimensional transformation.

Our work is based upon the approach of Loy and Eklundh [7], who proposed a method to detect dominant symmetries in images. The symmetry detection is based on the analysis of feature matches by their location including orientation and scale properties. We adapted this work to find repeated structures. The analysis of feature matches remains and only slightly had to be adjusted on the new problem.

Our method shall be applied for recognizing and outlining of building facades, where we have to face with competing structures of different sizes. Newest techniques for describing facades and their parts are developed by Brenner and Ripperda [1] or Čech and Šára [2] who use formal grammars. So far, these grammars are too general for our problem, but in future, these approaches might be helpful to combine the description of symmetries and repeated structures.

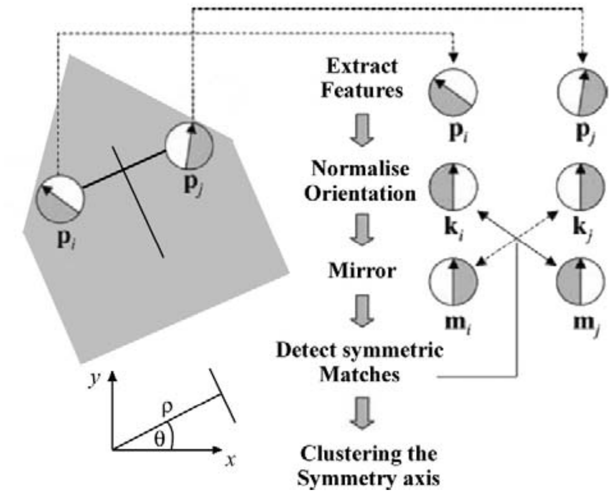

Figure 1: Principal functionality of detecting dominant symmetries, partly taken from [7]. The Hessian normal form of the symmetry axis *(θ,ρ)* is derived from the feature pair p_i and p_j.

Detection of Dominant Symmetries

Loy and Eklundh [7] proposed a method for finding dominant symmetries in images. We give a brief summary on this method. Additionally, its principal functionality is sketched in fig. 1.

Firstly, they detect prominent features by the SIFT operator [6]. So every feature is described by its location (row, column, scale and orientation) and by the descriptor, which encodes the gradient content in the local image patch, normalised with respect to the feature's orientation.

Flipped versions of the features are obtained after resorting the descriptor elements, see [11]. They subsequently match the sets of original features and their flipped versions to get pairs of potential symmetric features. Every pair is represented by the Hessian normal form of their symmetry axis with their normal and distance from origin. These coordinates are clustered over these parameters to find dominant symmetries among the found features.

The quality of symmetry M of each feature pair p_i und p_j is measured by

$$M_{ij} = \Phi_{ij} S_{ij} \tag{1}$$

where Φ_{ij} and S_{ij} are two weights defined as

$$\Phi_{ij} = -\cos(\varphi_i + \varphi_j - 2\theta) = -\cos(\alpha + \beta) \tag{2}$$

and

$$S_{ij} = \left[\exp\left(\frac{-|s_i - s_j|}{\sigma_s (s_i - s_j)} \right) \right]^2 \tag{3}$$

The angle-weight $\Phi_{ij} \in [-1,1]$ returns a high value for those feature pairs whose orientations are symmetrical with respect to the proposed symmetry axis, cf. fig. 2. The angles α and β add up to 180°, if the orientations are exactly symmetrical in respect to the proposed symmetry axis.

The scale-weight $S_{ij} \in [0,1]$ is used for limiting the differences between both features with respect to their scales s_i and s_j. Larger differences can be tolerated by increasing the parameter σ_s. Loy and Eklundh [7] introduced another weight with respect to the distance between both features, but this is only advantigous, if one would like to insert prior knowledge of the observed object. Since we want to look for all kind of symmetries within a facade images, we do not use this weight.

The Hessian normal form of the symmetry axes *(θ,ρ)* of all found potential symmetric feature pairs are accumulated with respect to their weightings in a two dimensional array. The result is a two dimensional histogram of the sum of symmetry measures over the parameters θ and ρ of the symmetry axes. Dominant symmetries of an image appear as relative maxima of this histogram. In contrast to [7] where the goals was to find only the major symmetry, we search for all significant symmetries by investigating all peaks of the histogram which are supported by at least t feature pairs.

Fig.3 shows the histogram in respect to the image of fig.4. This facade is described solely by horizontal symmetries. Therefore the histogram has its global maximum at

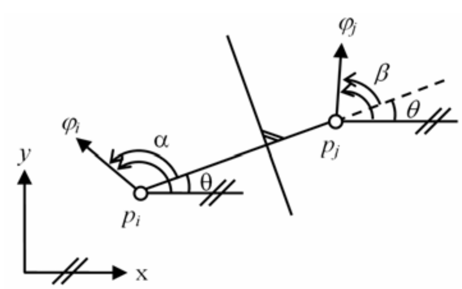

Figure 2: Illustration of the functionality of the angle-weight, according (2).

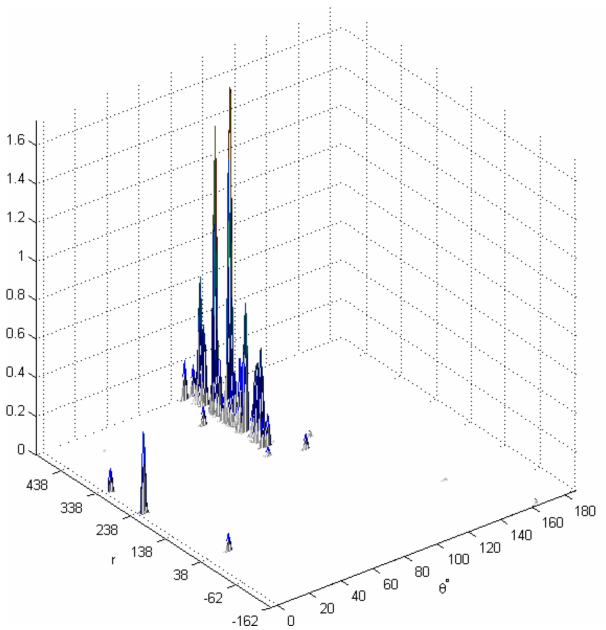

Figure 3: 2D-histogram over the polar coordinates of the symmetry axes for the example from fig.4

(θ = 90°, ρ = 391pix) and additional local maxima along the 90° grid line.

Fig. 4 f) shows the five detected symmetries in one image. In fig. 4 a) - e), we show each of the detected symmetry axis together with the convex hull of its supporting feature points. For this example, we detected 1617 features which form 151 potential symmetrical feature pairs. The major symmetry axis, cf. fig. 4 a), is supported by 34 feature pairs. The other four symmetry axes shown in fig. 4 b) - e) are supported by 21, 13, 13 and 9 feature pairs. All of the detected symmetries lie in the building facade, other objects of the image do not disturb the symmetry detection. Furthermore, the convex hulls of the involved features in all symmetries lead directly to the image region, which is characterized by the symmetrical structures.

Detection of Repeated Structures

We adapted the basic idea of clustering feature pairs within a single image to detect repeated structures. Obviously, the flipping of the feature descriptors can be omitted. Instead, we match the detected features with each other, such that we find pairs of very similar features, similar with respect to orientation and scale. Additionally, the weight according to orientation is adapted to our purpose. Thus, the angle-weight Φ_{ij} is simplified to $\Phi^*_{ij} \in [-1,1]$

$$\Phi^*_{ij} = \cos(\varphi_i - \varphi_j) \quad (4)$$

and it supports mostly those feature pairs with similar orientation.

Thus, the quality of repetition M^* is measured by

$$M^*_{ij} = \Phi^*_{ij} S_{ij} \quad (5)$$

Again clustering over directions and amount of translations yields the dominant translations in the image.

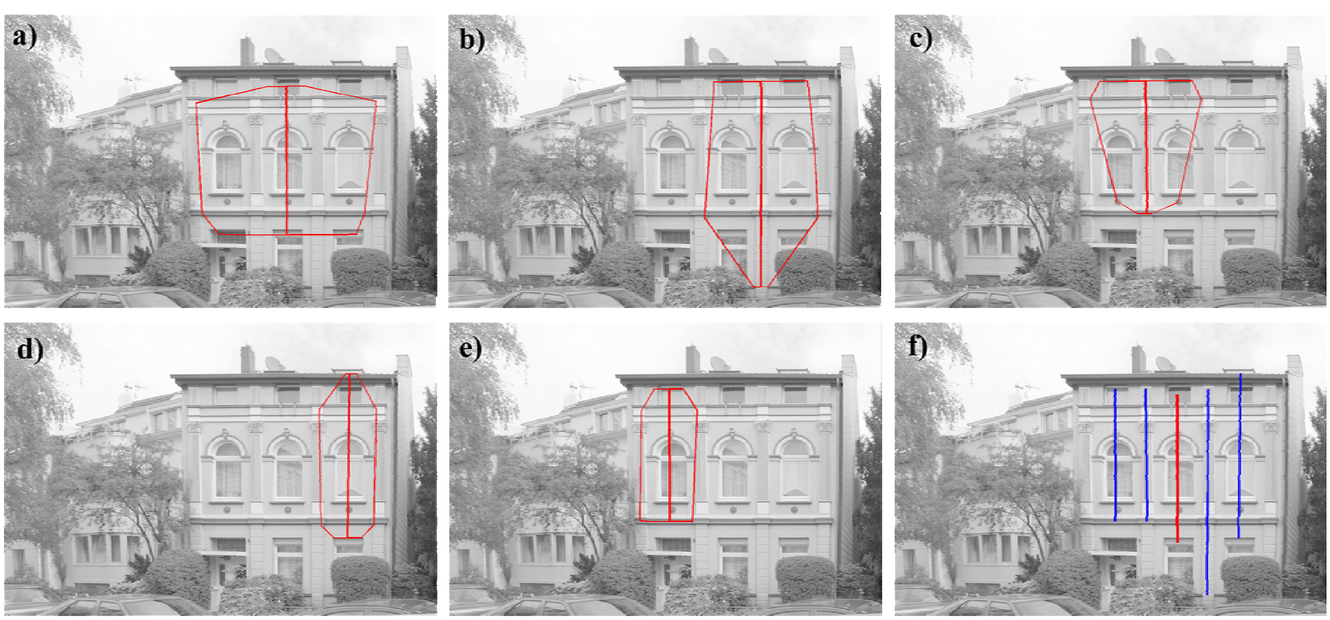

Figure 4: Results for symmetry: Exactly five dominant symmetry axes were found. a) - e) Single results for detected symmetries with convex hulls of involved features. f) Combination of all found symmetries.

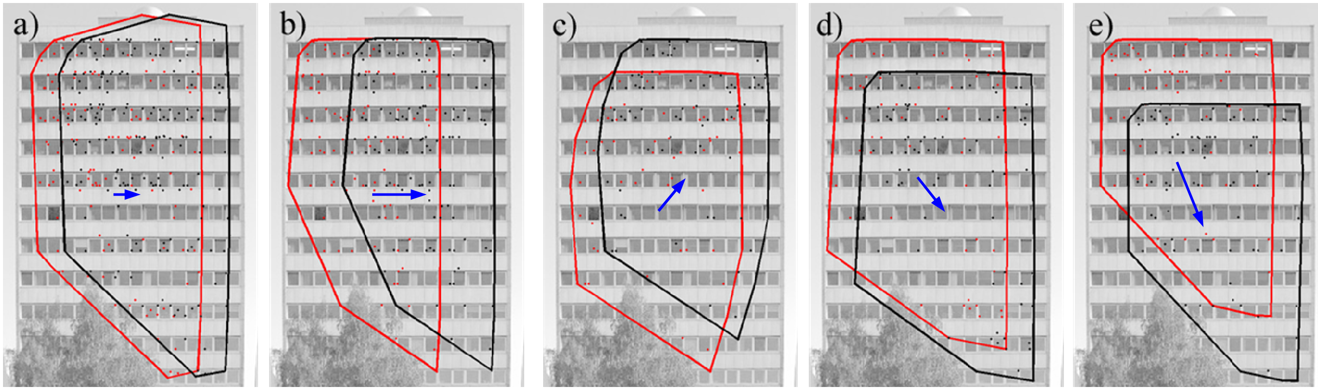

Figure 5: The five first detected repeated elements. The involved features and their convex hulls are represented, together with the translation vector between the red and black groups.

Dominant translations in the image correspond to the maxima of the histogram of the repetition measure. Furthermore, we focus on those translations which are supported by at least t feature pairs. Fig. 5 shows the first five detected repeated structures for this example. In each case, the red features are matched to black features by the same translation. The convex hulls of both feature groups are represented in fig. 5a)-e). For this example altogether 122 repeated groups were detected[1].

For better demonstration of these results fig. 6 shows all detected translations as plot of translation vectors. This representation shows clearly the regularity in the detected translations. We look for a compact description of these repetitions which exactly depicts, respectively, the regularity and underlying pattern.

Inference of the Compact Description

Because we work on rectified images, the main directions of the translations run parallel to the image borders. Therefore, we can reduce the search for a suitable basis to separate searches in the horizontal and in the vertical directions. Then, a typical facade is characterized by perpendicular regularities through rows and columns. There may be different types of repeated elements where bigger elements are compositions of smaller elements. Thus, the repeated elements can be represented in a hierarchical order per direction with depth K, which forms a hierarchical basis. This is illustrated in fig. 7. Note, that we do not restrict the repeated elements to have a certain shape.

Due to the reduction on horizontal and vertical directions, we project all translations on the dx and dy axes and treat these new translations as our observations d_i $(i = 1 : n)$. Thus, they can be described as a linear combination of axis parallel basis translations v_k and the appropriate coefficients α_k, the number of repetition, through

$$d_i = \sum_{k=1}^{K} (\alpha_k \cdot v_k) + \varepsilon_i \ . \qquad (6)$$

The depth K of the hierarchical basis corresponds to the number of elements of the linear combination. A priori the value of K is un-

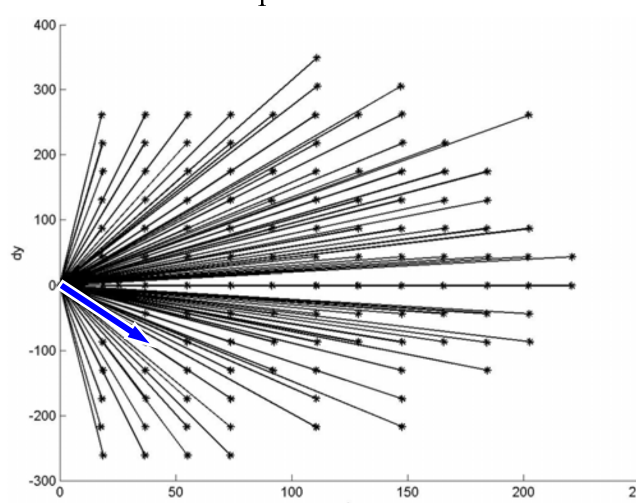

Figure 6: The vectors of translations of the 122 detected repeated groups for the example from fig 5. I.e. the blue arrow represents the translation found in fig 5 e).

[1] We selected the matching criterion of the Lowe-matcher *distRatio = 0.9* very sensitively concerning variances (shade, curtains etc.) of the facade elements. Thus, relatively large distances between the descriptors of the features lead to a positive match. For more details about the parameters for the matching of two SIFT feature descriptors, especially about *distRatio* see [6].

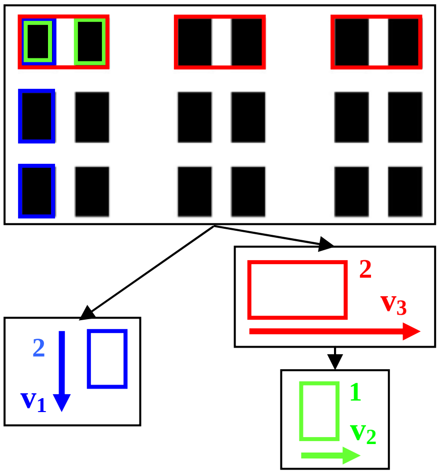

Figure. 7: A typical facade is characterized by perpendicular regularities through rows and columns. In this example, there is a hierarchy of repetition for the horizontal direction and a single two-fold repetition for the vertical direction. Thus in horizontal direction the compact image description consists of a hierarchy *(K=2)* of basis elements with the amount of the translation and the number of repetitions.

known, but we assumed typical urban facades in its complexity do not exceed the value $K=4$. Neither the integer-valued coefficients α_k nor the real-valued basis translations v_k are known. Furthermore, each observation d_i is afflicted with a residual ε_i.
We look for a hierarchical basis, consisting of K basis elements, which explain the observed translations in the best possible way, including the minimisation of the residues and the complexity K of the solution.

Since we could not find a direct solution for this problem, we decided for a heuristic procedure. Therefore, we determine the differences between all observed translations. We calculate a histogram via these second differences of the positions. The peaks of this histogram are potential candidates for the basis translations that we look for. For these c candidates, we form all

$$C = \binom{c}{K} \tag{7}$$

combinations of possible bases $\mathbf{v}$. Then, we determine the appropriate coefficients α_k for each of these potential solutions ${}^j\mathbf{v}$ $(j=1{:}C)$ and for each observation d_i. The residual vector ${}^j\boldsymbol{\varepsilon}$ is obtained for the results of every solution ${}^j\mathbf{v}$. The best solution minimises the residuals with the smallest model complexity.

If a certain data set can be described by a compact model, then only the model parameters and possible deviations of the data from this model need to be encoded. This consideration leads to the MDL criterion, proposed in [8]:

$$MDL = -\log \prod_{i=1}^{n} P(x_i|\pi) + \frac{K}{2}\log(n) \tag{8}$$

We look for that model (π, K), that describes the observed data x_i with the smallest complexity K and the largest data probability

$$\prod_{i=1}^{n} P(x_i|\pi) \tag{9}$$

where π are the parameters of the modell. On the assumption of normally distributed residuals the criterion can be represented as

$$MDL = \frac{1}{2}\Omega + \frac{K}{2}\log(n) \tag{10}$$

The consideration of outliers is based on Huber, cf. [3], with the optimisation function

$$\kappa(\varepsilon) = \begin{cases} T^2 & \text{if } (\varepsilon/\sigma)^2 \geq T^2 \\ (\varepsilon/\sigma)^2 & \text{if } (\varepsilon/\sigma)^2 < T^2 \end{cases} \tag{11}$$

and

$$\Omega = \sum_{i=1}^{n} \kappa(\varepsilon_i) \tag{12}$$

According to the critical value T traditionally chosen on the basis of the significance level of hypothesis test, we select the threshold value for outliers as $T = 3\sigma$ with $\sigma = 1.5$.
The convex hull of the feature pairs, which support the selected model $\mathbf{v}$, defines the region that can be described by these basis elements. Thus, we get a compact description of the repetitive structure in the form of basis elements and the associated regions in the image.

Fig. 8 shows the results of the inference of the compact description for two facade images. On the left, where we continue the example from fig. 5, a basis which consists only of one element has been determined for both axis directions. The convex hulls of the features that take part in this basis cover the entire facade region (with exception of the region covered by the tree). On the right of fig. 8, we present an ex-

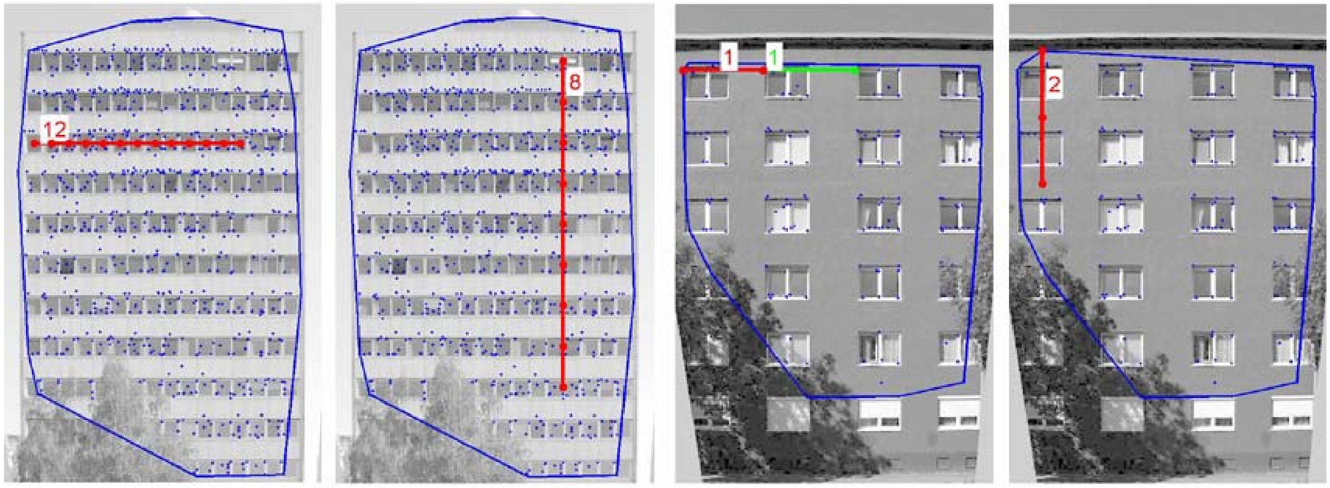

Figure 8. Two results of inference of the compact description of the structure. For the horizontal and vertical direction the regions are shown that are described by the basis elements together with the found basis elements. The number of repetitions of basis vectors is given by the maximum number of coefficients for the linear combinations of basis vectors of all observations.

ample of a hierarchical basis in the horizontal direction. The four columns of windows do not have the same distance from each other, but the two window columns on the left have the same distance as the two window columns on the right. Thus, we obtain two different translation vectors according to the real structure of the facade.

Conclusion

We showed how the approach from [7] can be extended to the detection of multiple repeated groups. From the detected translations in the image we derived a model for a compact description of the repetitive structure in facade images using a heuristic search method and the criterion of the minimum description length. So far, our algorithm only works on images, which show only one regular part of facades. The matching procedure is very sensitive to the repeated objects in the regular part of a facade due to the very generous choice of the matching criterion. Especially, similar structures in the neighbourhood of the facades trouble our approach. We need to refine our method, in particular regarding the robustness against disturbances in the picture.

References

1. C. Brenner, N. Ripperda. Reconstruction of Façade Structures Using a Formal Grammar and RjMCMC. DAGM 2006. P. 750-759.
2. J. Čech, R. Šára. Language of the structural models for constrained image segmentation. Technical Report TN-eTRIMS-CMP-03-2007, 2007.
3. W. Förstner. Image Analysis Techniques for Digital Photogrammetry. Schriftenreihe des Inst. f. Photogrammetrie der Univ. Stuttgart. 1989. Vol. 13. P. 205-221.
4. J. Hays, M. Leordeanu, A. Efros, Y. Liu. Discovering Texture Regularity as a Higher-Order Correspondence Problem. ECCV 2006, P. 522-535.
5. T. Leung, J. Malik. Detecting, localizing and grouping repeated scene elements from an image. ECCV 1996. P. 546-555.
6. D. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision. 2004. Vol. 60, No. 2. P. 91-106.
7. G. Loy, J.-O. Eklundh. Detecting Symmetry and Symmetric Constellations of Features. ECCV 2006. P. 508-521.
8. J. Rissanen. Stochastic Complexity in Statistical Inquiry. World Scientific: Series in Computer Science. 1989. Vol. 15.
9. F. Schaffalitzky, A. Zisserman. Planar grouping for automatic detection of vanishing lines and points. Image and Vision Computing. 2000. Vol. 18(9). P. 647-658.
10. T. Tuytelaars, A. Turina, L. Van Gool. Noncombinatorial detection of Regular Repetitions under Perspective Skew. PAMI. 2003. Vol. 25(4). P. 418-432.
11. S. Wenzel. Resorting the elements of the SIFT-feature descriptor for symmetry detection. Technical Report TB-igg-p-07/1. 2007.