Unsupervised Footwear Impression Analysis and Retrieval from Crime Scene Data

Adam Kortylewski(✉), Thomas Albrecht, and Thomas Vetter

Departement of Mathematics and Computer Science,
University of Basel, Basel, Switzerland
{adam.kortylewski,thomas.albrecht,thomas.vetter}@unibas.ch

Abstract. Footwear impressions are one of the most frequently secured types of evidence at crime scenes. For the investigation of crime series they are among the major investigative notes. In this paper, we introduce an unsupervised footwear retrieval algorithm that is able to cope with unconstrained noise conditions and is invariant to rigid transformations. A main challenge for the automated impression analysis is the separation of the actual shoe sole information from the structured background noise. We approach this issue by the analysis of periodic patterns. Given unconstrained noise conditions, the redundancy within periodic patterns makes them the most reliable information source in the image. In this work, we present four main contributions: First, we robustly measure local periodicity by fitting a periodic pattern model to the image. Second, based on the model, we normalize the orientation of the image and compute the window size for a local Fourier transformation. In this way, we avoid distortions of the frequency spectrum through other structures or boundary artefacts. Third, we segment the pattern through robust point-wise classification, making use of the property that the amplitudes of the frequency spectrum are constant for each position in a periodic pattern. Finally, the similarity between footwear impressions is measured by comparing the Fourier representations of the periodic patterns. We demonstrate robustness against severe noise distortions as well as rigid transformations on a database with real crime scene impressions. Moreover, we make our database available to the public, thus enabling standardized benchmarking for the first time.

1 Introduction

Footwear impressions are one of the most frequently secured types of evidence at crime scenes. For the investigation of crime series they are among the major investigative notes, permitting the discovery of continuative case links and the conviction of suspects. In order to simplify the investigation of cases committed by suspects with the same footwear, the crime scene impressions are assigned to a reference impression (see Fig. 1). Through the assignment process, the noisy and incomplete evidence becomes a standardized information with outsole images, brand name, manufacturing time, etc. Currently, no automated systems exist that can assess the similarity between a crime scene impression

C.V. Jawahar and S. Shan (Eds.): ACCV 2014 Workshops, Part I, LNCS 9008, pp. 644–658, 2015.
DOI: 10.1007/978-3-319-16628-5_46

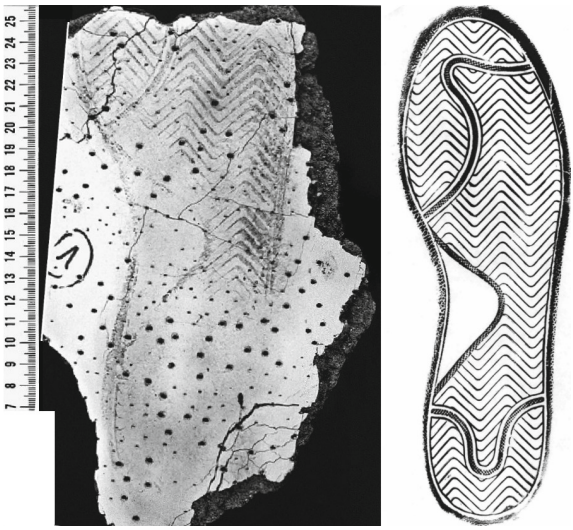

Fig. 1. A crime scene impression and the corresponding reference impression. A main challenge for pattern recognition systems is to isolate the shoe sole pattern from the structured noise. Moreover, the impression is incomplete and the zigzag line is shifted between the impressions. Therefore no point-wise similarity measure can be applied. Furthermore, the relative orientation and translation between the images is unknown.

and reference impressions, due to the severe image degradations induced by the impression formation and lifting process. The main challenges here are the combination of unknown noise conditions with rigid transformations and missing data (see Fig. 1). Moreover, the image modalities could be inverse, meaning that the impression information could be white for one impression and black for the other. Furthermore, training and testing data are scarce, because usually no or few crime scene impressions are available per reference impression. Additionally, in many cases no point-to-point correspondence exists between the impressions because different parts of the shoe sole are produced independently of each other. This results for example in a phase shift of the zigzag line between the impressions in Fig. 1. Therefore, a higher level understanding of the pattern is necessary. In this work, we introduce an unsupervised image retrieval algorithm that overcomes the limitations of existing work by detecting and analyzing periodic patterns in the footwear impressions under unconstraine noise conditions. The only assumption being, that the noise signal is not strictly periodically structured. The basic idea behind our approach is that periodic structures are the most preserved information under unknown noise conditions because of their inherent redundancy. In our reference impression database containing 1175 images, about 60 % of the images show periodic patterns. Given the challenging application scenario, solving the recognition task for this subset of the data is a valuable contribution.

The main contribution of this work is to extend the periodic pattern model from Lin et al. [8], such that the localization and analysis of periodic patterns under unconstrained, structured noise becomes possible. The reliable extraction of the pattern model enables the compensation of rigid transformations, the estimation of the scale of the repeating texture tile, and the exclusion of background information and noise from the pattern representation. After calculating a local Fourier representation, we robustly segment the periodic pattern from other structures in the image. Finally, during the image retrieval process we use the extracted Fourier features to compute the similarity between images. Our footwear impression database is available to the public at http://gravis.cs.unibas.ch/fid/, enabling research on real case data and standardized comparisons for the first time.

In Sect. 2, we will discuss in detail how the feature extraction is performed. Afterwards, we introduce the detection of periodic patterns and the image retrieval algorithm. In Sect. 3, we introduce our footwear impression database and perform a thorough experimental evaluation. We conclude with Sect. 4.

1.1 Previous Work

Early work on automated shoeprint recognition approached the problem by either analyzing the frequency spectra of the whole images using the Fourier transform [5,6], or by describing the image with respect to its axes through Hu-Moment invariants [1]. Such global shoe print processing methods are particularly sensitive to noise distortions and incomplete data. Pavlou et al. [14,15] and Su et al. [17] proposed to classify shoe print images based on local image features. Their approaches are based on combinations of local interest point detectors and SIFT feature descriptors. However, in general such gradient-based methods are not sufficient for the application of crime scene impression retrieval (see the experiments in Sect. 3). The reason is that the image gradient of the noisy data is strongly distorted compared to clean reference data. Therefore, a reliable gradient-based detection of interest points or the correct rotational normalization based on the image gradient are difficult. Patil et al. [13] divide the image into a block structure of constant size and process each block individually with Gabor features. Their approach is too rigid to capture different shoe print orientations because of the constant sized block grid. Nevertheless, they show that local normalization is a crucial step in processing noisy images. Regarding rotational invariance, DeChazal et al. [5] are tolerant to rotation by brute force rotating the images in one-degree steps. However, this is not usefull in practice because of the computational costs. Nibouche et al. [12] use SIFT combined with RANSAC to compensate for the rotation, but since the feature descriptors are heavily distorted by the structured noise, the method is only applicable to noiseless data. These and most other works in the field work on synthetically generated training and testing data, assuming a very simple noise model such as Gaussian or salt and pepper noise. However, one key challenge of real data is that the noise is unconstrained and therefore cannot be simulated by such

simple noise distributions. Dardi and Cervelli [2,4] published algorithms applied to real data. However, the approach in [4] is based on a rigid partitioning of images and thus is sensitive to image transformations and the approach in [2] is rotationally invariant, but is not robust against noise and incomplete data. Another approach was introduced by Tang et al. [18]. They extract basic shapes (circles, lines and ellipses) out of the shoe print image with a modified Hough transform and store these shapes in a graph representation. Attributes such as distance, relative position, dimension and orientation are encoded into the graph structure, making their recognition algorithm robust to image transformations. But many shoe soles are comprised of more complex patterns which cannot be described by such basic shapes. Additionally, in their experiments, crime scene impressions are mixed with synthetic data, giving no clear performance statements for the real case scenario. A general review on shoeprint recognition methods has been presented by Luostarinen and Lehmussola [11]. However, the experimental setup is very restricted, because they assume known correspondance between crime scene impression and reference impression. Since that information can only be provided after already knowing the correctly matching reference, that work has only limited relevance for the task of footwear impression retrieval.

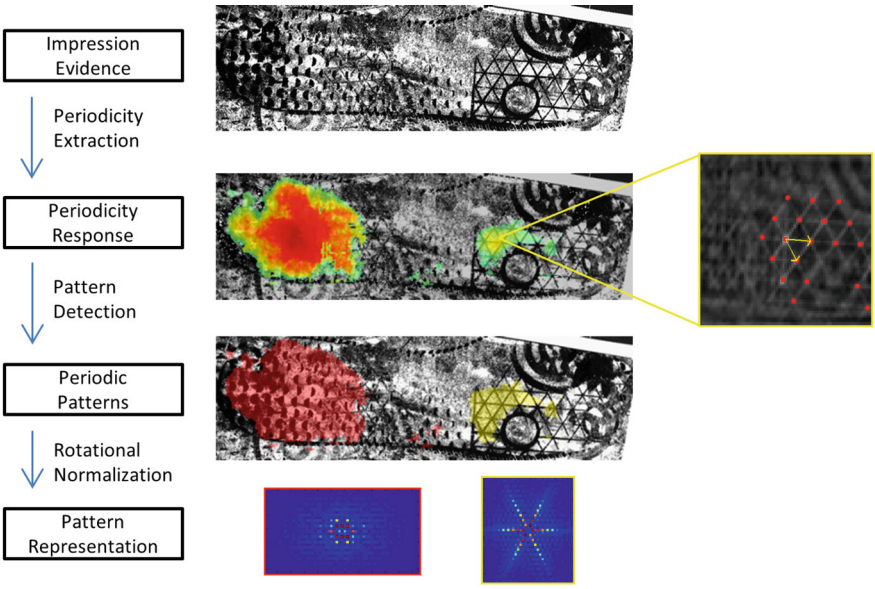

Fig. 2. Graphical abstract of the feature extraction procedure. We first extract the periodicity at each point in the impression evidence. Then, we compute Fourier features and use these to detect the periodic patterns. Finally, the periodic patterns are represented by the rotationally normalized Fourier features.

2 Impression Analysis Based on Periodic Patterns

The goal of this work is to compare images in the presence of unconstrained noise, rigid transformations, without point-to-point correspondence and across different modalities, based on their periodic patterns. A periodic pattern is fully defined by a basic texture tile that is spatially distributed in two dimensions according to two fixed distribution vectors. We begin with periodicity detection by fitting a model of translational symmetries to local autocorrelation responses. Then, we calculate translation-invariant Fourier descriptors and separate the periodic patterns from other structures in the image by point-wise classification. Afterwards, the individual periodic patterns can be rotationally normalized with respect to their inherent translational symmetry distribution. Finally, the comparison between impressions is achieved by first recomputing the Fourier representation of the crime scene pattern based on the extracted periodicity models of the reference impressions, and the subsequent comparison of the feature representations. An overview about our approach is depicted in Fig. 2.

2.1 Periodicity Extraction

The basic building block of the proposed image retrieval pipeline is the ability to measure the periodicity at a point X in the image. This presumes a robust extraction of translational symmetries in the local neighbourhood of X. We start by extracting a quadratic image patch around X and correlating it within a local region of interest around X by normalized cross correlation. Since the scale of the pattern is initially unknown, we repeat this procedure with multiple patch sizes and average the resulting correlation maps together. The advantage of the local, patch-based correlation over a global autocorrelation is that translational symmetries stay sharp in the correlation map, even if other structures are present in the region of interest. Afterwards, the positions of peaks in the correlation map are extracted through non-maximum-suppression. Since many spurious peaks are extracted by the non-maximum-suppression, we propose to filter the peaks with a threshold τ on the correlation value. Compared to the filtering approach in [9], this is more robust against distortions in the region of interest by strong noise or other image structures (Fig. 3). The resulting list C of candidate peak positions c_k for translational symmetries, still does not solely contain correct translational symmetries of the pattern, as can be seen in the first and third row of Fig. 3. We propose a periodicity extraction method that extracts the correct periodicity despite the remaining spurious peaks.

Rigid periodic patterns follow a grid-like spatial distribution. Therefore, we search for the two shortest linearly independent vectors that can describe the spatial distribution of the points in C, with an integer linear combination of themselves. To this end, we build on the Hough transform approach of [8] and constrain the candidate vectors to originate at the center of the region of interest and point to one of the candidate peaks c_k. Thus introducing a vector space with the origin set at the center of the region of interest. Additionally, the number of possible candidate vectors is reduced to the number N of points in C. We modify

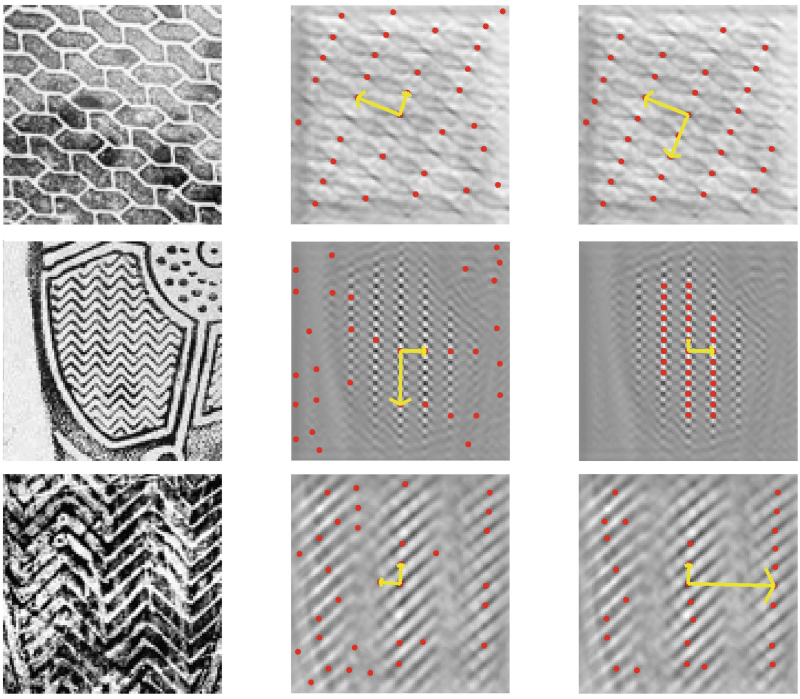

Fig. 3. Detection of translational symmetries. Left column: region of interest; middle column: result of the approach of [9]; right column: result of our approach. The results show the autocorrelation maps in gray, the candidate peaks for translational symmetries in red and the detected periodicity indicated by two vectors in yellow. The proposed approach extracts peaks more reliably and gives better periodicity estimates compared to the approach of [9].

the scoring function of [8] by incorporating the correlation values at the peak positions $NCC(c_k)$ and by penalizing the length of both distribution vectors, instead of just the largest one (Eq. 1). The function $nint(x)$ in Eq. 1 rounds a scalar x to the next integer.

$$\begin{aligned} h(c_p, c_q, C) = {} & \alpha \cdot (NCC(c_p) + NCC(c_q)) \\ & + \sum_{c_k \in C,\, k \neq p,q} \frac{(1 - 2\max(|a_k - nint(a_k)|, |b_k - nint(b_k)|))\, NCC(c_k)}{\|c_p\| + \|c_q\|} \\ & \begin{bmatrix} a_k \\ b_k \end{bmatrix} = \begin{bmatrix} c_p, c_q \end{bmatrix}^{-1} \begin{bmatrix} c_k \end{bmatrix}; \quad c_k, c_p, c_q \in C \end{aligned} \tag{1}$$

The variables a_k and b_k are coefficients related with the linear combination of c_p and c_q for the peak c_k. The factor α performs the tradeoff between the importance of the correlation values and the shortness of the distribution

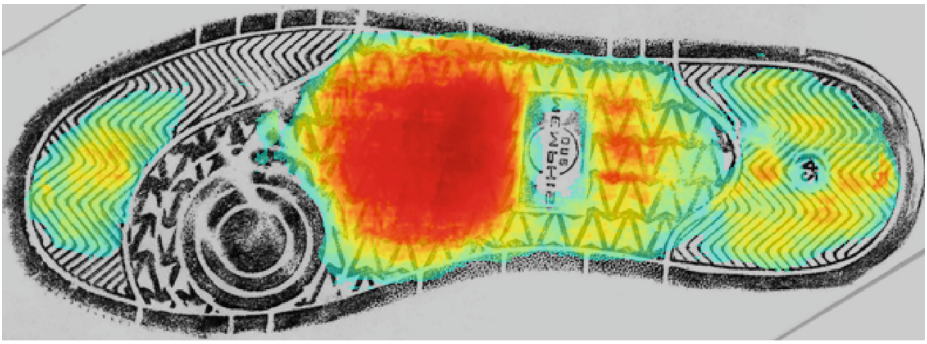

Fig. 4. Illustration of periodicity scores for different points in a reference impression; the maximal value is colored in red. The periodicity score reflects the amount and the similarities of translational symmetries in a local region of interest.

vectors. The periodicity at X is then determined by searching for the pair of vectors $(c_p, c_q) = \text{argmax}_{p,q}\{h(c_p, c_q, C)|p = 1..N, q = 1..N\}$. In Fig. 4 the periodicity score h is illustrated for each point in a reference impression. It can be seen, that maxima of the score are reached predominantly at the centers of the periodic patterns. We have illustrated the results of our approach and the one of [9] in Fig. 3. The combination of the multi-scale autocorrelation maps with the modified scoring function is robust against local correlation maxima as well as distortions through noise and additional structures in the region of interest.

2.2 Periodic Pattern Detection and Representation

On the basis of the periodicity measurement introduced in Sect. 2.1, we compute a low-dimensional representation of the periodic patterns by pooling redundancies into one feature descriptor and subsequently determine the number of periodic patterns in the image. We propose to extract Fourier based features to encode the appearance and the periodicity of a periodic pattern. However, the computation of discriminative local Fourier features presumes the selection of the right window size. Especially for small window sizes this is critical, since image structures that do not belong to the periodic pattern have a great distortive impact on the frequency spectrum, as do the discontinuities at the window boundary. In the following, we describe a method to determine a suitable window for the Fourier transform based on the periodicity information. The Fourier features are then applied to detect the number of periodic patterns contained in the image in an iterative procedure.

We start by computing the periodicity at each point $P = (i, j)$ in the image $I \in \mathbb{R}^{n \times m}$. The maximal periodicity scores h and the corresponding distribution vectors v_1 and v_2 are stored in a Matrix $H = \{H(P) = [h(P), v_1(P), v_2(P)] \mid i = 1..n, j = 1..m\}$. Based on that information, we apply an iterative grouping procedure to determine if and how many periodic patterns are available in the image.

Algorithm 1. Pseudocode of the Periodic Pattern Detection

```
1: Input: H ← Periodicity, I ← Image
2: PPatterns ← Empty
3: while max(H,h) > 0 do
4:     X ← maxpos(H,h)
5:     f_X ← CalculateFourierDescriptor(X,H,I)
6:     H(X)=0;
7:     for (Y)∈I do
8:         if s(f_X,CalculateFourierDescriptor(Y,H,I))>ρ then
9:             H(Y)=0;
10:    PPatterns ← f_X
11: Output: PPatterns

12: function CalculateFourierDescriptor
13:    Input: X, H, I
14:    [v1,v2] ← H[X]
15:    R_{I,X} ← Rect(X,v1,v2,I)
16:    f_X ← |F(G(X, Σ_{R_{I,X}}) · R_{I,X})|
17:    Output ← f_X
```

The procedure is summarized in pseudocode in Algorithm 1. An iteration starts at the point X in the image with maximal periodicity score (see Algorithm 1, l.4). High periodicity indicates that translational symmetries exist in the local neighborhood. This also implies that other image structures are less likely to occur in the surrounding area. Based on the corresponding distribution vectors $v_1(X)$ and $v_2(X)$, a window R is determined as the smallest rectangle containing the points $\{X-v_1(X)-v_2(X); X-v_1(X)+v_2(X); X+v_1(X)-v_2(X); X+v_1(X)+v_2\}$. Then, the image patch $R_{I,X}$ of size R centered at position X is extracted. An important advantage of determining the window R based on the periodicity of the pattern is that we are able to capture approximately an integer number of periods of the pattern. Additionally, no other structures are contained in $R_{I,X}$, thus making it a particularly good basis for further feature extraction processes. In practice it is still beneficial to reduce distortions in the frequency spectrum through small discontinuities at the boundary, by multiplying $R_{I,X}$ with a Gaussian window of size R centered at X before the Fourier transformation. We denote the Gaussian window by $G(X, \Sigma_{R_{I,X}})$. Two periods of the pattern are extracted in each direction of repetition because that increases the sampling rate during the Fourier transformation and thus leads to more discriminative Fourier features. A central property of periodic patterns is that only the phase of the frequency spectrum varies throughout the pattern, but the magnitude stays approximately constant. We exploit this fact by using just the magnitude of the frequency spectrum of $R_{I,X}$ as a feature descriptor for the periodic pattern:

$$f(R_{I,X}) = \left|\mathcal{F}(G(X, \Sigma_{R_{I,X}}) \cdot R_{I,X})\right|. \tag{2}$$

With this translationally invariant descriptor we are able to classify if another point Y belongs to the same pattern as X by evaluating the normalized

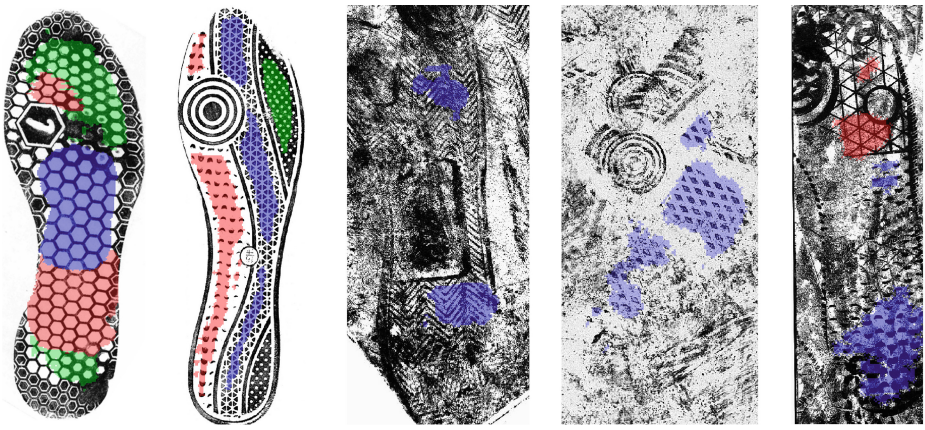

Fig. 5. Results of the grouping procedure on reference images as well as crime scene impressions. Different patterns in the same image are coded in different colors. Periodic patterns on very different scales are well separated from other structures in the images even under strong noise conditions.

cross-correlation of the Fourier descriptors at these points:

$$s(f(R_{I,X}), f(R_{I,Y})) = \frac{(f(R_{I,X}) - \mu_{f(R_{I,X})})(f(R_{I,Y}) - \mu_{f(R_{I,Y})})}{\sigma_{f(R_{I,X})}\sigma_{f(R_{I,Y})}}. \quad (3)$$

An important detail to notice is that both Fourier descriptors are calculated with the same window size R, ensuring that the feature descriptors have the same dimensionality. By ignoring the phase of the frequency spectrum, the feature descriptor gets invariant to inverse image modalities.

Based on the extracted Fourier feature $f(R_{I,X})$ we classify each point in the image by computing the similarity measure of Eq. 2, followed by a thresholding with a constant value ρ (see Algorithm 1, l.8). For the points with a similarity greater than ρ, we set the corresponding periodicity score in H to zero. This procedure is repeated until the maximum remaining periodicity score is zero. After this procedure, the image is decomposed into a set of periodic patterns, each represented by its corresponding Fourier feature. In Fig. 5, example results illustrate the extracted patterns for different impressions containing periodic patterns from a wide variety of scales and appearances. It can be seen that different periodic patterns are well separated from the background even under structured noise, indicating that the extracted Fourier features are a good representation of the corresponding patterns.

2.3 Similarity and Footwear Retrieval

We propose to compute the similarity between a crime scene impression and a reference impression by means of the similarity between their periodic patterns. Before two patterns are compared, we first actively normalize their rotation

based on their translational symmetry structures computed above. We do so instead of e.g. making the feature descriptor rotationally invariant, in order to retain the discriminative power of the descriptors. The normalization is achieved by rotating the image so that the smallest distribution vector points upright. In cases where both vectors have the same length we align to the vector along which the image gradient is greatest. After this, we recalculate the Fourier features on the rotationally normalized images. In practice, the estimated window sizes R for the Fourier transformation do not always have exactly the same size for the same patterns in crime scene impressions and reference impressions. This is due to differences in the noise conditions and the pixel discretization of the image. By using the window size of the reference pattern for both impressions, we ensure that the features have the same dimensionality. Although this can lead to distortions of the frequency spectrum when the patterns differ, the effect is negligible for similar patterns. Afterwards, we compare the features as described in Eq. 2. The feature extraction for the crime scene impressions can also be interpreted as an interest point detection with rotational normalization, as the Fourier features are recomputed during the matching with the window sizes of the respective reference impressions. The similarity between a reference impression $A = \{f(R^1_{I_A,X_1}), ..., f(R^n_{I_A,X_n})\}$ and a crime scene impression $B = \{I_B; Y_1, ..., Y_m\}$ can now be computed by the following similarity measure:

$$S(A,B) = \frac{1}{m}\sum_{j=1}^{m} \max_i \; s(f(R^i_{I_A,X_i}),\; f(R^i_{I_B,Y_j})). \tag{4}$$

As the crime scene impression may not contain all periodic patterns from the shoe outsole, only the subset of periodic patterns from the reference with maximal similarity is evaluated.

3 Experiments

We test our approach on a database of real crime scene impressions. The database consists of 170 crime scene impressions, among which 102 show periodic patterns. Out of the 170 impressions the pattern extraction algorithm detects for 133 impressions periodic patterns, including all true periodic patterns and 31 false positives. During the evaluation, we show the performance on all 133 impressions with detected periodic pattern in order to evaluate a fully automated retrieval setting. Those 37 impressions that are not included in the evaluation are also to be published with the database and are marked to enable a correct performance comparison of future works. The reference database consists of 1175 reference impressions. In a random subset of 230 impressions, 142 showed a periodic pattern, thus we assume that roughly 60 % of the reference impressions show periodic patterns.

3.1 Setup

In our approach, we do not account for scale changes, because this information is provided by a ruler on each impression image and can thus be given manually.

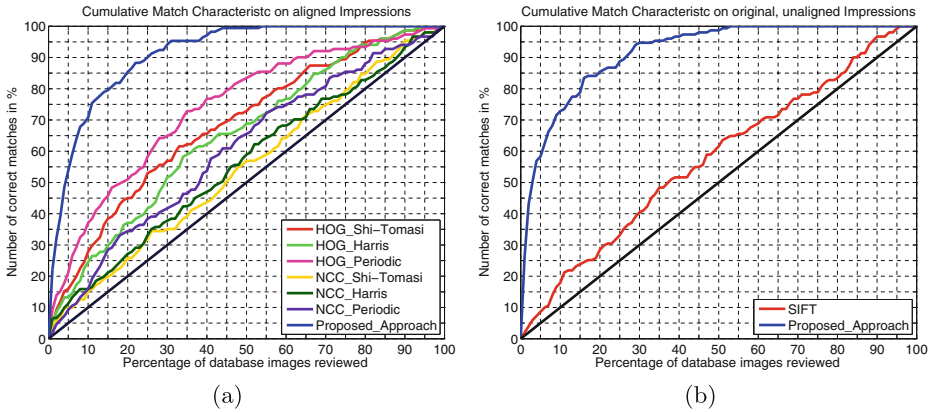

Fig. 6. Cumulative match characteristic for (a) crime scene impressions aligned to the corresponding reference impression and (b) the original unaligned data.

For the experiments, the images are scaled to 10 pixels per centimeter. Also, the crime scene images are cropped so that the impression is roughly centered in the image frame (see Fig. 7). During the experiments we choose the parameter $\alpha = 2$. The patch sizes for the localized multi-scale autocorrelation are chosen to vary from 11 to 37 pixels in a four pixel step size. With this configuration, we capture all scales of periodic patterns that show at least two repetitions in both directions. We constrain the angle between the distribution vectors to be between 60 and 90 degrees as in [9]. This reduces the number of possible distribution vector combinations and thus acts as a regularization and saves computational time. Despite this constraint, we are still able to describe the translational symmetries of all periodic patterns in the database. During the experiments, we first compute the similarity between the query impression and all reference impressions in the database. We sort the reference images by similarity, thus producing a ranking list of the most similar reference impressions. As a performance measure, we apply the cumulative match characteristic. This score reflects the question: "What is the probability of a match if I look at the first n percent of the ranking list?" (adapted from [5]). We split the experiments into two parts. For the first part, we manually align all crime scene impressions to the corresponding reference impressions. Thus, the remaining difference between the impressions is of a structural nature. The normalization makes it possible to compare our approach with robust methods that are not invariant to rigid transformations, such as the normalized cross correlation (NCC) or the histogramm of oriented gradients (HOG) [3]. In the second part, we compare the SIFT algorithm [10] with the proposed approach on the original unaligned data.

3.2 Footwear Impression Retrieval Performance

The comparison methods in the first experiment are histogramms of oriented gradients and normalized cross correlation. We combine HOG and NCC with

Table 1. Summary of the experiments in terms of footwear impression retrieval performance measured by the cumulative match characteristic.

		Recall@				
	Method	1 %	5 %	10 %	15 %	20 %
Aligned	HOG + Shi-Tomasi	3.8	15.8	25.6	37.6	44.4
	HOG + Harris	4.5	13.5	24.8	30.1	37.6
	HOG + Periodic	9.8	22.6	36.8	45.9	50.4
	NCC + Shi-Tomasi	3.8	9.8	15.0	19.6	26.3
	NCC + Harris	7.5	12.8	17.3	22.6	28.6
	NCC + Periodic	2.3	10.5	17.3	30.1	36.8
	This work	**27.1**	**56.4**	**70.0**	**76.7**	**85.0**
Unaligned	SIFT	1.5	8.3	18.1	24.1	28.6
	This Work	**27.1**	**59.4**	**74.4**	**79.7**	**85.7**

interest point detections by the Harris [7] and Shi-Tomasi [16] corner detectors and with the feature locations detected by our approach. For the corner detection we use the algorithms implemented in Matlab. The patch size is fixed to 21×21 pixels since that leads to the best performance in the experiments. Our approach estimates the window size by itself during the feature extraction. Although the alignment to the correct reference impression is not possible in forensic practice, this experiment is of interest because it measures the performance of HOG and NCC and interest point detectors under structured noise. During the experiments, we detect interest points on the crime scene impressions and subsequently compute the descriptors at the interest point locations. Since the crime scene impression and the reference impression are in correspondence, we repeat this procedure at the same locations in the reference impressions. The number of extracted features on a query impression is fixed to the number of features detected by our approach, in order to allow for a comparison of the results. The results are illustrated in Fig. 6a. NCC performs nearly equally for all interest points, as it does not account for any structural information in the image patches. HOG features perform better than NCC on average as they utilize local structural information to a certain degree and gain robustness to small transformational deviations compared to NCC by pooling information locally into histrogramms. Combined with the periodic interest points detected by our approach, it also performs better than on the Harris and Shi-Tomasi interest points. That is mainly because the periodic interest points always lie inside the shoe impression, whereas the others also detect less discriminative points on the impression boundary. Through the criterion of high periodicity during the interest point extraction, these points are also less likely to be distorted by noise. Despite that the alignment clearly introduces a positive bias on the matching performance, the results in general show that neither HOG features nor NCC perform well under the noisy environment. Our approach performs significantly better compared to

the other feature detector and descriptor combinations. Especially under the first 10 % in the ranking list, the performance gain is about 33 %. Important is that our procedure does not profit from the alignment of the data since the rotaion is normalized during feature extraction. In Fig. 6b, we compare our approach on the original unaligned data with the SIFT feature extraction and description approach. For the experiments we use the SIFT implementation from Andrea Vedaldi and Brian Fulkerson [19]. Although, SIFT performs above chance rate, especially for the more important first fifth of the ranking list, the performance is still not satisfactory. Our proposed approach performs significantly better, so that e.g. the Recall@10 % is improved from 18.1 % to 74.4 % compared to SIFT. The performance results are summarized in Table 1. Another observation is that our approach reaches faster to 100 % matching score than the other approaches. This is because the algorithm does not detect a periodic pattern for 558 reference impressions and can thus exclude these from the candidate list early. Note that the performance results of the proposed approach are nearly the same for aligned and unaligned data, underlining the rotational invariance of the feature extraction. The low result for the Recall@1 % originates from the fact the proposed approach focuses on the periodic patterns in the impressions and no other structures. Since many reference impression have e.g. grid like patterns, these are all grouped together at the front of the ranking list (see Fig. 7). But since no other structures are included in the similarity measure, their order is not clearly defined and depends on nuances in the scale or noise of the impressions. We have illustrated two image retrieval results in Fig. 7. Despite strong structured noise and even different modalities in the first example, the grid-like and circular structures of the periodic patterns are clearly reflected in the retrieval results. The missing 25.6 % at Recall@10 % can be ordered in two categories. One part are impressions that are smeared through liquid on the ground, such that the rigid transformation assumption in the feature extraction does not hold anymore. The other part are double-prints, meaning that two impressions are overlayed on top of each other such that the algorithm is not able to extract the correct periodicity.

4 Conclusion

In this work, we have proposed an image retrieval algorithm based on periodic patterns. The algorithm is robust under unconstrained noise conditions by separating the meaningful pattern information from the structured background. Additionally, it is robust against incomplete data and it overcomes the problem of absence of point-to-point correspondence between impressions by extracting a translation-invariant pattern representation. Furthermore, it is able to match rotated data by actively normalizing the pattern representations with respect to the intrinsic tranlational symmetry structure of the periodic patterns. Our experiments demonstrate a significant performance gain over standard image retrieval techniques for the task of footwear impression retrieval. By making the database with real crime scene impressions and reference impressions publicly available,

we open a new application to the field of computer vision, concerning the issue of how to separate patterns from structured noise despite incompleteness and spatial transformations. Thus, our publication enables standardized benchmarking in the field for the first time. In the future, we plan to make the approach scale invariant and, since regular patterns are only available on about 60 % of the data, we will also focus on how to robustly incorporate other structures from footwear impressions into the retrieval process.

Fig. 7. Results of the image retrieval algorithm for two crime scene images. The column shows the query image. The second to fifth columns show the top results in the ranking list. And the last column shows the correct references. The correct references are found at position five and 13 in the ranking lists.

Acknowledgement. This Project was supported by the Swiss Comission for Technology and Innovation (CTI) project 13932.1 PFES-ES. The authors thank the German State Criminal Police Offices of Niedersachsen and Bayern and the company forensity ag for their valuable support.

References

1. AlGarni, G., Hamiane, M.: A novel technique for automatic shoeprint image retrieval. Forensic Sci. Int. **181**(1), 10–14 (2008)
2. Cervelli, F., Dardi, F., Carrato, S.: A translational and rotational invariant descriptor for automatic footwear retrieval of real cases shoe marks. Eusipco (2010)
3. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, vol. 1, pp. 886–893. IEEE (2005)
4. Dardi, F., Cervelli, F., Carrato, S.: A texture based shoe retrieval system for shoe marks of real crime scenes. In: Foggia, P., Sansone, C., Vento, M. (eds.) ICIAP 2009. LNCS, vol. 5716, pp. 384–393. Springer, Heidelberg (2009)
5. De Chazal, P., Flynn, J., Reilly, R.B.: Automated processing of shoeprint images based on the fourier transform for use in forensic science. IEEE Trans. Pattern Anal. Mach. Intell. **27**(3), 341–350 (2005)
6. Gueham, M., Bouridane, A., Crookes, D., Nibouche, O.: Automatic recognition of shoeprints using fourier-mellin transform. In: NASA/ESA Conference on Adaptive Hardware and Systems, AHS 2008, pp. 487–491. IEEE (2008)
7. Harris, C., Stephens, M.: A combined corner and edge detector. In: Alvey Vision Conference, Manchester, UK, vol. 15, p. 50 (1988)
8. Lin, H.-C., Wang, L.-L., Yang, S.-N.: Extracting periodicity of a regular texture based on autocorrelation functions. Pattern Recogn. Lett. **18**(5), 433–443 (1997)
9. Liu, Y., Collins, R.T., Tsin, Y.: A computational model for periodic pattern perception based on frieze and wallpaper groups. IEEE Trans. Pattern Anal. Mach. Intell. **26**(3), 354–371 (2004)
10. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
11. Luostarinen, T., Lehmussola, A.: Measuring the accuracy of automatic shoeprint recognition methods. J. Forensic Sci. (2014)
12. Nibouche, O., Bouridane, A., Crookes, D., Gueham, M., et al.: Rotation invariant matching of partial shoeprints. In: 13th International Machine Vision and Image Processing Conference, IMVIP 2009, pp. 94–98. IEEE (2009)
13. Patil, P.M., Kulkarni, J.V.: Rotation and intensity invariant shoeprint matching using gabor transform with application to forensic science. Pattern Recogn. **42**(7), 1308–1317 (2009)
14. Pavlou, M., Allinson, N.M.: Automatic extraction and classification of footwear patterns. In: Corchado, E., Yin, H., Botti, V., Fyfe, C. (eds.) IDEAL 2006. LNCS, vol. 4224, pp. 721–728. Springer, Heidelberg (2006)
15. Pavlou, M., Allinson, N.M.: Automated encoding of footwear patterns for fast indexing. Image Vis. Comput. **27**(4), 402–409 (2009)
16. Shi, J., Tomasi, C.: Good features to track. In: 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Proceedings CVPR 1994, pp. 593–600. IEEE (1994)
17. Su, H., Crookes, D., Bouridane, A., Gueham, M.: Local image features for shoeprint image retrieval. In: British Machine Vision Conference, vol. 2007 (2007)
18. Tang, Y., Srihari, S.N., Kasiviswanathan, H., Corso, J.J.: Footwear print retrieval system for real crime scene marks. In: Sako, H., Franke, K.Y., Saitoh, S. (eds.) IWCF 2010. LNCS, vol. 6540, pp. 88–100. Springer, Heidelberg (2011)
19. Vedaldi, A., Fulkerson, B.: Vlfeat: An open and portable library of computer vision algorithms. In: Proceedings of the International Conference on Multimedia, pp. 1469–1472. ACM (2010)