Discrepancy Norm as Fitness Function for Defect Detection on Regularly Textured Surfaces

Gernot Stübl[1], Jean-Luc Bouchot[2], Peter Haslinger[1] and Bernhard Moser[1]

[1]Software Competence Center Hagenberg,
Softwarepark 21, A-4232 Hagenberg, Austria
[2]Department of Knowledge-based Mathematical Systems,
Altenbergerstr. 69, A-4040 Linz, Austria

Abstract. This paper addresses the problem of quality inspection of regular textured surfaces as, e.g., encountered in industrial woven fabrics. The motivation for developing a novel approach is to utilize the template matching principle for defect detection in a way that does not need any particular statistical, structural or spectral features to be calculated during the checking phase. It is shown that in this context template matching becomes both feasible and effective by exploiting the so-called discrepancy measure as fitness function, leading to a defect detection method that shows advantages in terms of easy configuration and low maintenance efforts.

1 INTRODUCTION

This paper is motivated by the demand for performant, highly discriminative and still easy to configure visual defect detection algorithms.

In literature one can find statistical, structural, model-based or filtering approaches for optical quality inspection. See, e.g., Xie[29] or Kumar[17] for a survey. Statistical approaches rely on histograms and first or second order statistics of the intensity image. Features like mean, variance, median, entropy, inertia or contrast can be computed from such statistics. See, e.g., Haralick[13], Ng[25]. Structural approaches, see, e.g., Mirmehdi et al.[20], are based on the principle of defining a structural element and finding its spatial distribution applying morphological operations. Fractal-based methods [10] exploit the concept of fractal dimension as characteristic of textures. Markov Random Fields [9] rely on probabilistic models of the spatial dependencies of the intensity values. Filtering methods aim at characterizing textures by informative spatial or spectral features. Typical examples are approaches based on wavelets [7] or Gabor filters [4]. Approaches based on TEXEM models [30] consider textures as superposition of texture elements in order to represent the texture under consideration.

As a further approach we introduce and study template matching in the context of quality inspection of regular textured surfaces as for instance encountered in industrial woven fabrics (sieves, airbag hose, textiles for automotive interior etc.). The goal is to utilize the template matching principle for defect detection in a way that does not need any particular statistical, structural or spectral features to be calculated. It will be shown that this goal can be achieved by exploiting the so-called discrepancy measure as fitness function.

This paper is structured as follows. Section 2 discusses whether template matching is an appropriate approach for detecting defects in regular textures. In this context the following three aspects are addressed: a) choice of appropriate dissimilarity measure in Section 3, b) design of an algorithm for matching test patches with a reference in Section 4, and c) specification of a discriminative rule for detecting defect candidates in Section 5. This Section proposes a RANSAC inspired algorithm for realizing this template matching concept. The next Section contains experimental evaluations on woven fabrics to show the stability and performance of the algorithm. The conclusion, Section 7, outlines future research potentials.

2 TEMPLATE MATCHING FOR TEXTURE ANALYSIS

In quality inspection a sensed image I is scanned by a sliding window with varying center (i, j). This sliding window crops a test image patch $T_{(i,j)} \subseteq I$ that has to be analysed whether it indicates a defect or not. Template matching requires a notion of (dis-)similarity d that measures to which extent the test patch, $T_{(i,j)}$, matches a given reference image, R. We call the resulting map $\Theta_d(i, j) = d(T_{(i,j)}, R|_{T_{(i,j)}})$ the (dis-)similarity map induced by the (dis-)similarity measure d, $R|_{T_{(i,j)}}$ denotes the restriction of R to the set of pixels of $T_{(i,j)}$.

One way to define a (dis-)similarity map is based on applying a distance measure to extracted features as for example interest points which are widely used in Computer Vision [1, 19]. In this case the dissimilarity measure is defined as a metric in the corresponding feature vector space. The question whether feature-based (dis-)similarity concepts for template matching are reliable or not depends on the number and the distinctiveness of the available feature points. Particularly, for textures, standard methods for template matching usually fail because of the lack of sufficiently distinctive and reproducible feature points. Therefore, generally speaking, the usage of feature points for texture analysis is not recommendable.

Measuring the (dis-)similarity directly by applying a (dis-)similarity measure d to the images as sets of (ordered) intensity values is an alternative to the feature-based similarity approach. Such a (dis-)similarity based approach requires a) to choose an appropriate dissimilarity measure (Section 3), b) to choose the size of the test patch, c) to define a matching concept (Section 4), and d) to specify a rule how to distinguish defect from defect-free samples (Section 5).

The choice of the test patch size is crucial for the inspection of periodic and quasi-periodic textures. A too small size might cause undesired registration artefacts whereas a too large size causes unnecessary processing time. The optimal size is closely related to the estimation of the length of the repetitive pattern. However window size estimation is not the topic of this paper and therefore the window size is always calculated with an existing software tool which is based on a modified algorithm of Lizarraga-Morales et al.[18].

The main question is whether and under which circumstances a (dis-)similarity based template matching approach is appropriate to allow a discriminative analysis of defects. The reason why this approach might lead to serious problems is not clearly analysed in the literature so far. This paper contributes to giving an answer to this analysis and, at the same time, to opening up a new approach that mitigates problems inherently induced by commonly used (dis-)similarity measures.

3 CHOICE OF APPROPRIATE DISSIMILARITY MEASURE

This section focusses on the adequateness of similarity measures in the context of defect detection. First of all, let us point out that commonly used (dis-)similarity measures induce the occurrence of local extrema as artefacts. Particularly it can be shown that commonly used (dis-)similarity measures like L_2 norm, mutual information, cross-correlation, Bhattacharyya measure[3] or Kullback-Leibler divergence measure[16] are likely to lead to artefacts in terms of local extrema that corrupt the resulting (dis-)similarity map Θ_d. For details see Moser et al.[23]. As an alternative measure we suggest to exploit Weyl's concept of discrepancy [28], which was introduced in order to measure irregularities of distributions [2, 15]. Due to Moser[22] let us propose

$$\begin{aligned}\|\boldsymbol{I}\|_D := \max \Big\{ &\max_{0\le k\le n, 0\le l\le m} \Big\{ \sum_{i=0}^{k}\sum_{j=0}^{l} I_{(i,j)} \Big\}, \max_{0\le k\le n, 0\le l\le m} \Big\{ \sum_{i=0}^{k}\sum_{j=0}^{l} I_{(n-i,j)} \Big\}, \\ &\max_{0\le k\le n, 0\le l\le m} \Big\{ \sum_{i=0}^{k}\sum_{j=0}^{l} I_{(i,n-j)} \Big\}, \max_{0\le k\le n, 0\le l\le m} \Big\{ \sum_{i=0}^{k}\sum_{j=0}^{l} I_{(n-i,n-j)} \Big\} \Big\},\end{aligned} \quad (1)$$

($I_{(0,0)} := 0$) as an extension of Weyl's discrepancy measure to image data with an image I of the width n and the height m . The indexed variables k and l indicate the current partial sum. Note that (1) can be efficiently computed in $O(n \cdot m)$ by using integral images, see Moser[22].

The interesting point about this is that based on Weyl's discrepancy concept distance measures can be constructed that guarantee desirable registration properties: (R1) the measure vanishes if and only if the lag vanishes, (R2) the measure increases monotonically with an increasing lag, and (R3) the measure obeys a Lipschitz condition that guarantees smooth changes also for patterns with high frequencies.

4 TEMPLATE MATCHING BY REGISTRATION

This section discusses a recently outlined template matching algorithm [5]. It is state-of-the art to process a test image by specifying a so-called sliding window for consecutively cropping image patches and comparing them to reference data. Usually the reference image data are chosen to have the same size as the sliding window.

In this paper we concentrate on textures with regular or nearly regular patterns. Examples for this kind of regular textures are woven fabrics as e.g. for automotive interiors, industrial sieves, air-bag hoses, etc. Different localizations of the sliding window yield patches having different offsets with respect to the repetitive pattern. By this the resulting patches are varying in their appearance, although they refer to the same reference pattern. This effect can be interpreted in terms of translational misalignment. As discussed in Section 3 such a misalignment might lead to undesirable artefacts when applying commonly used similarity measures.

The basic idea of Bouchot et al.[5] is to make the usual reference-test image matching more flexible by allowing registration on a larger reference image that covers multiple periods of the repetitive pattern. To avoid the before mentioned artefacts the approach introduces the discrepancy norm as similarity measure and, thereby, fitness function for the registration. At the same time the enlarged reference image size can be chosen also to capture variations in appearance due to other effects like changes in illumination or admissible variations in production.

Mathematically speaking, given a test patch T the registration step aims at identifying optimal transformation parameters $\xi = (\xi_1^*, \ldots, \xi_k^*)$ that minimize the match given by

$$d_T(\xi) = d\left(H_\xi(T), R|_{H_\xi(T)}\right) \tag{2}$$

where d denotes an appropriate dissimilarity measure, R represents the chosen defect-free reference, T the actual patch of a test image, H is a parametrized transformation model (in our case translations) and $R|_{H_\xi(T)}$ denotes the subregion of the reference which is specified by the pixel coordinates of the transformed test patch $H_\xi(T)$. To allow only translational transformations is no restriction for usage in an in-line inspection system of endless material where the camera position is fixed and no rotations and scale variations occur. However with a more complex transformation model also other variations can be covered. In this context it has to be noticed that discrepancy norm is also able to cope with small rotations as demonstrated in Moser[22].

Two drawbacks keep the algorithm of Bouchot et al.[5] from industrial usage. Firstly the window size has to be manually chosen, which is now done with a modified algorithm of Lizarraga-Morales et al.[18]. Secondly although the computational costs are below exhaustive transformation parameters search, they are still too high for usage in an in-line inspection system. To improve speed, this paper adds a statistical level set analysis of the (dis-)similarity map Θ_d, gathered on the defect free reference image as a preprocessing step. Through this a threshold for a RANSAC like global optimization algorithm can be estimated, which accelerates the decision if a patch is defect free or not drastically (from minutes to fractions of a second, on Matlab with a standard PC). This is in strong contrast to the brute-force registration principle used in Bouchot et al.[5].

5 DISCRIMINATIVE RULE FOR DETECTING DEFECT CANDIDATES

A defect can be excluded if registration parameters ξ can be found that reduce the dissimilarity (2) between the test image T with a defect-free reference patch from R below some threshold θ.

The threshold θ can be determined from the cumulative distribution function of dissimilarity values generated by applying defect-free patches to the reference. For example θ can be defined as q-quantile. The lower q the more sensitive the detection, but also the higher the expected rate of pseudo-defects. Its optimal choice depends on the texture characteristics and the type of defects under consideration.

In order to turn this principle into a computational rule for detecting defect candidates let us consider a RANSAC like algorithm. Randomized Sampling Consensus (RANSAC)[12] is a common method to fit models into noisy data by constructing candidate models out of random samples and choosing the model with the best fit. Given a test patch T the problem is to find a position $\xi \in R$ for which $d_T(\xi) \leq \theta$. The key idea of the proposed approach is to randomly compare the test patch to different positions on the reference image and choose the best position to start a local optimization. Suppose that T is defect-free, and let denote $\mathcal{F}$ the set of defect-free patches. The probability $\alpha_\theta = P(d_T(\xi) \leq \theta | \, T \in \mathcal{F})$ to randomly choose a position $\xi \in R$ with $d_T(\xi) \leq \theta$ can be estimated as ratio between the area of the level set $\lambda_\theta = \{\xi \in R | d_T(\xi) \leq \theta\}$ and the area of the reference image R.

Let denote $\{\xi_1, \ldots, \xi_k\}$ a sequence of k independent random trials and consider the conditional error probability $\varepsilon = \varepsilon_{\{\xi_i\},\theta} = P\left(\min_{i=1}^{k} d_{\xi_i} > \theta \,|\, T \in \mathcal{F}\right)$ that all trials yield positions outside λ_θ. Then, we obtain $\varepsilon_{\{\xi_i\},\theta} = (1-\alpha_\theta)^k$. Starting with a probability $p_s = 1 - \varepsilon$, e.g., $p_s = 0.995$, the estimated number k of trials amounts to

$$k = \ln(1-p_s)/\ln(1-\alpha_\theta). \tag{3}$$

Therefore, we obtain the rule:

"If a position $\xi \in R$ with match dissimilarity $d_T(\xi) \leq \theta$ is found with at most k trials given by (3), then the test patch T is considered defect-free, otherwise a defect candidate."

This rule can be modified by taking a maximal number of local optimization iterations into account in order to accelerate the evaluation. Computational experiments showed that the quasi-Newton BFGS method is a reasonable choice for the local optimizer[6].

6 EVALUATION

The applicability of the proposed approach is demonstrated on the basis of defect samples of regular textures taken from the TILDA database and industrial applications of the involved research institutes.

Figure 1 depicts three types of defects: a) high contrast structural defects spreading over many repetitive pattern units, see Figures 1(a) and 1(c), b) more challenging contrast structural defects spreading over many repetitive pattern units in Figures 1(e), 1(g), 1(i) and 1(k), and low contrast defects affecting only a small number of repetitive pattern units in Figure 1(m), 1(o), 1(q) and 1(s). Each image is shown in two versions, the test image and the output of the algorithm, which was not thresholded. The reference images which hold about four to five periods of the original pattern are not depicted. The structure in the output images indicates the number of steps in the search for the best matches due to the coarse-to-fine approach outlined above. For example, Figures 1(a) and 1(c) took less steps than in the other images. What also can be observed is that the defect positions are not always precise. This effect originates from the block processing working principle of the algorithm: if a defect is not fully covered by a single block, the position cannot be located precisely. In future versions this misalignment can be compensated by an additional registration.

These examples of various defect types may demonstrate the potential of the proposed approach also to detect low contrast defects. With a MATLAB implementation on a standard PC the evaluation on one test image of size 756×512 took fractions of a second. This shows that the proposed approach is computationally feasible.

STABILITY Stability analysis is performed on a set of 20 test images. For each test image the algorithm output is split into blocks with window size equal to the length of the period. This is also the finest scale of the coarse-to-fine approach. We choose $p_s = 0.9999$ and $\theta = 1.5\,\sigma + \mu$, where μ and σ denote the mean and standard deviation of the discrepancy values computed from dissimilarity maps of 10 randomly chosen patches from the according reference image. For this configuration we observe an average α_θ of 0.2 which lead to a range of $[26, 88]$ for k pattern comparisons. A block is marked as belonging to a defective region if more than 50 % of its pixels are defective according to the rule of Section 5.

Since the aim of the stability analysis is to show the repeatability of the results, every image is processed 100 times and a defect probability p_i is calculated for each block. In the worst case this probability is 0.5, which means that no clear decision can be made whether the block is defective or not. Therefore as measurement of the stability the entropy $H_i = p_i(1 - p_i)$ is calculated on each block. For the whole test set of 20 images and 7735 blocks this leads to a mean entropy of 0.0036 bit, which demonstrates the high stability of the algorithm despite the random working principle.

PERFORMANCE In the current state of development the proposed algorithm is not supposed to outperform any state-of-the art textile defect detection algorithm. It is rather thought as demonstration on how a novel dissimilarity measure can open up new possibilities in this application context. Nevertheless a comparison with current state-of-the art algorithms in the application field of textile defect detection was done. The evaluation is based on the work of Tolba

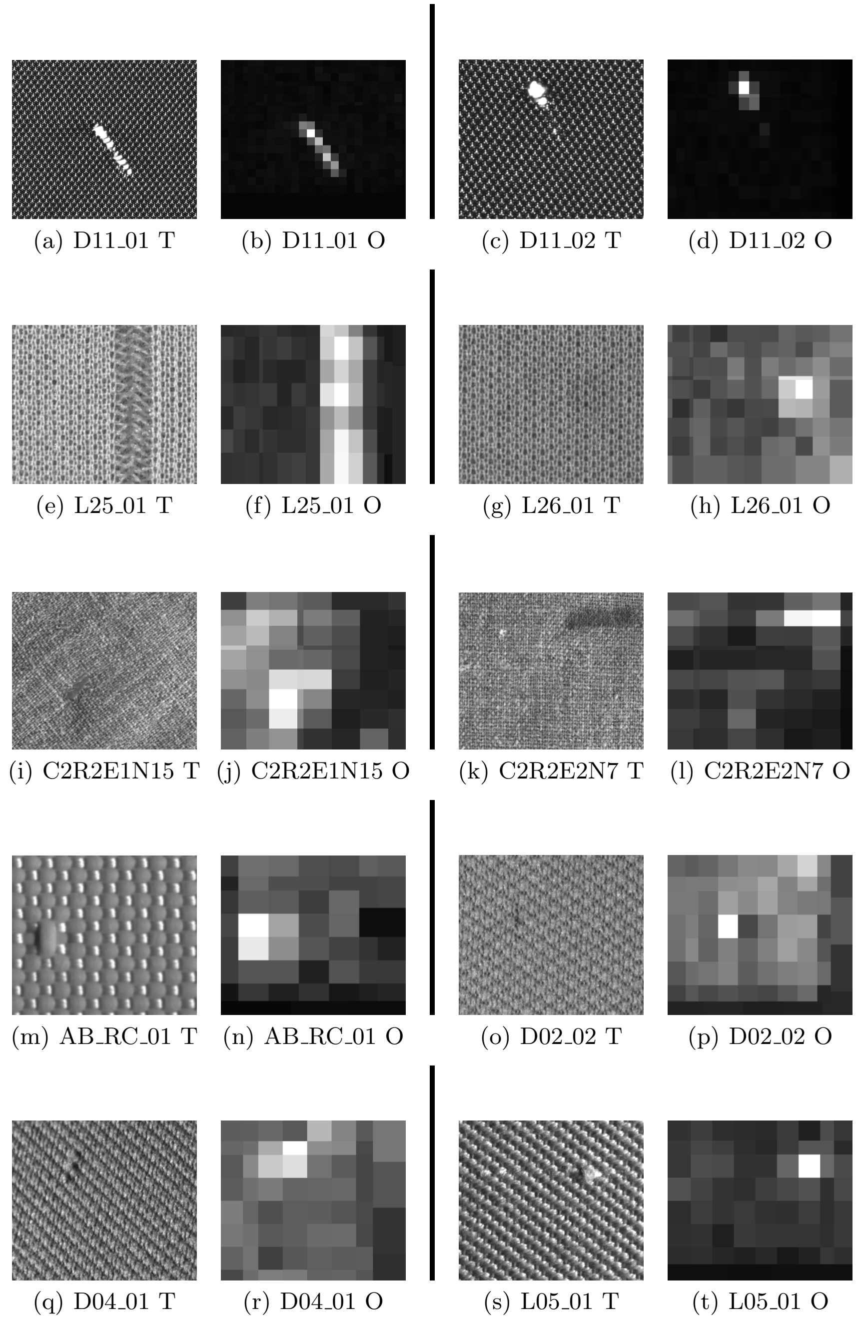

Fig. 1. Examples of the applicability test on textile defect images. For each example the test image (T) as well as the output (O) of the algorithm is shown. Reference images are not shown. The algorithm output is for illustration purposes not thresholded.

et al.[27]. These authors directly compare several state-of-the art algorithms by calculating a performance measure called Percentage of Correct Detection (PCD)

$$PCD = (1 - (FAR + FRR)) \times 100 \tag{4}$$

out of the False Acceptance Rate (FAR) and False Rejection Rate (FRR) reported in papers of the algorithm designers. The performance for our algorithm is evaluated on the above mentioned test set of 20 images using hand-labelled ground truths and the same decision rule as in the stability analysis. This leads to a PCD of 96.1 %. Table 1 summarizes the best performing algorithms listed in

Method	*PCD (%)*	*Reference*
Decision Fusion	98.64	[27]
GLCM	97.09	[21]
GLCM + Gabor + wavelet packets (selected from 219 features)	96.90	[14]
Selected from Gabor and GLCM	96.90	[11]
Discrepancy Norm Based Template Matching	**96.10**	**-**
Clustering	91.60	[8]
Wavelet97	88.15	[7]
Local Binary Patterns	85.83	[26]

Table 1. Performance comparison of texture characterization approaches using Percentage of Correct Detection (PCD), for details see Tolba et al.[27]. The original Table contains multiple entries per reference, here only the best performing ones are listed. Furthermore the top performing method of Murino et al.[24] is skipped because it is a pure classification algorithm without detection. The discrepancy norm based algorithm can be compared with the class of Grey-Level Co-occurrence Matrices (GLCM) and filter (Gabor, wavelet) based feature extraction algorithms.

Tolba et al.[27] together with our method. It shows up that the performance is in the same class as a combination of Grey-Level Co-occurrence Matrices (GLCM) and filter (Gabor, wavelet) based feature extraction algorithms. However in contrast to the work of Drimbarean and Whelan[11], who use a neural network classifier, the proposed approach does not need any training, but only relies on one sample image. Another issue with state-of-the art methods is the sometimes expensive computation of features, as can be found in Karoui et al.[14], Monadjemi[21] and Tolba et al.[27]. The latter even suggest to outsource the feature generation on FPGAs. Our proposed approach does not have the computational complexity problem. Nevertheless it does not reach the detection performance of other methods since the failure location is due to the block processing nature not always precise. Therefore it cannot compete with highly tuned state-of-the-art methods. To solve the location problem is a future research topic. Furthermore it has to be mentioned that our approach is limited to regular or near regular textures. Despite that the algorithm performs surprisingly good, having in mind that it is a not yet optimized straight forward approach with only one sensitivity configuration parameter for the coarse-to-fine working principle.

7 CONCLUSION

The template matching principle in the context of quality inspection of regular textures has been addressed. The motivation was to come up with a method that effectively can be implemented, shows distinctiveness also for low contrast defects and still is easy to configure. The approach outlined only requires the configuration of a sensitivity parameter which reduces the efforts of configuration. Experimental results indicate its usefulness and motivate further research to improve the defect localization. The approach outlined in this paper can also be combined with other methods e.g. Decision Fusion [27] which remains future research.

ACKNOWLEDGEMENTS

This work was supported in part by the Austrian Science Fund (FWF) under grant no. P21496 N23, and by the European Fund for Regional Development under Regionale Wettbewerbsfähigkeit OÖ 2007-2013. We also want to thank the reviewers for their comments and suggestions.

References

1. Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: SURF: Speeded up robust features. Computer Vision and Image Understanding 110, 346 – 359 (2008)
2. Beck, J., Chen, W.W.L.: Irregularities of distribution. Cambridge University Press, New York, NY, USA (2009)
3. Bhattacharyya, A.: On a measure of divergence between two statistical populations defined by probability distributions. Bull. Calcutta Math. 35, 99–109 (1943)
4. Bodnarova, A., Bennamoun, M., Latham, S.: Optimal Gabor filters for textile flaw detection. Pattern Recognition 35, 2973–2991 (2002)
5. Bouchot, J.L., Stübl, G., Moser, B.: A template matching approach based on the discrepancy norm for defect detection on regularly textured surfaces. In: Proceedings of the SPIE 10th International Conference on Quality Control by Artificial Vision. SPIE, Saint Etienne, France (June 2011)
6. Broyden, C.G.: The convergence of a class of double-rank minimization algorithms 1. General considerations. IMA Journal of Applied Mathematics 6(1), 76–90 (March 1970)
7. Chen, C.M., Chen, C.C., Chen, C.C.: A comparison of texture features based on SVM and SOM. In: Proceedings of the 18th International Conference on Pattern Recognition - Volume 02. pp. 630–633. ICPR '06, IEEE Computer Society, Washington, DC, USA (2006)
8. Cheng, H.D., Sun, Y.: A hierarchical approach to color image segmentation using homogeneity. IEEE Transactions on Image Processing 9(12), 2071–2082 (2000)
9. Cohen, F., Fan, Z., Attali, S.: Automated inspection of textile fabrics using textural models. IEEE Transactions on Pattern Analysis and Machine Intelligence 13, 803–808 (1991)
10. Conci, A., Proença, C.B.: A fractal image analysis system for fabric inspection based on a box-counting method. Computer Networks and ISDN Systems 30, 1887–1895 (1998)

11. Drimbarean, A., Whelan, P.F.: Experiments in colour texture analysis. Pattern Recognition Letters 22(10), 1161–1167 (2001)
12. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM 24(6), 381–395 (Jun 1981)
13. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics 3, 610–621 (1973)
14. Karoui, I., Fablet, R., Boucher, J.M., Pieczynski, W.: Fusion of textural statistics using a similarity measure: application to texture recognition and segmentation. Pattern Analysis and Applications 11(3-4), 425–434 (2008)
15. Kuipers, L., Niederreiter, H.: Uniform distribution of sequences. Dover Publications, New York, NY, USA (2005)
16. Kullback, S., Leibler, R.A.: On information and sufficiency. The Annals of Mathematical Statisitcs 22(1), 79–86 (Mar 1951)
17. Kumar, A.: Computer-vision-based fabric defect detection: A survey. IEEE Transactions on Industrial Electronics 55, 348–363 (2008)
18. Lizarraga-Morales, R., Sanchez-Yanez, R., Ayala-Ramirez, V.: Homogeneity cues for texel size estimation of periodic and near-periodic textures. In: Martnez-Trinidad, J., Carrasco-Ochoa, J., Ben-Youssef Brants, C., Hancock, E. (eds.) Pattern Recognition, Lecture Notes in Computer Science, vol. 6718, pp. 220–229. Springer Berlin / Heidelberg (2011)
19. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal on Computer Vision 60, 91 – 110
20. Mirmehdi, M., Marik, R., Petrou, M., Kittler, J.: Iterative morphology for fault detection in stochastic textures. Electronic Letters 32, 443–444 (1996)
21. Monadjemi, A.: Towards efficient texture classification and abnormality detection. Ph.D. thesis, University of Bristol, UK (2004)
22. Moser, B.: Similarity measure for image and volumetric data based on Hermann Weyl's discrepancy measure. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(11), 2321–2329 (2011)
23. Moser, B., Stübl, G., Bouchot, J.L.: On a non-monotonicity effect of similarity measures. In: Similarity-Based Pattern Recognition. pp. 46–60. Springer (2011)
24. Murino, V., Bicego, M., Rossi, I.A.: Statistical classification of raw textile defects. In: Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04). ICPR '04, vol. 4, pp. 311–314. IEEE Computer Society, Washington, DC, USA (2004)
25. Ng, H.F.: Automatic thresholding for defect detection. Pattern Recognition Letters 27, 1644–1649 (2007)
26. Pietikinen, M., Ojala, T.: Nonparametric texture analysis with simple spatial operator. Spectrum (1999)
27. Tolba, A.S., Khan, H.A., Mutawa, A.M., Alsaleem, S.M.: Decision fusion for visual inspection of textiles. Textile Research Journal 80 (2010)
28. Weyl, H.: Über die Gleichverteilung von Zahlen mod. Eins. Mathematische Annalen 77, 313–352 (1916)
29. Xie, X.: A review of recent advances in surface defect detection using texture analysis techniques. Electr. Letters on Computer Vision and Image Analysis 3, 1–22 (2008)
30. Xie, X., Mirmehdi, M.: TEXEMS: Texture exemplars for defect detection on random textured surfaces. IEEE Transactions on Pattern Analysis and Machine Intelligence 29, 1454–1464 (2007)