Measuring Regularity of Network Patterns by Grid Approximations using the LLL Algorithm

A. Hajdu, B. Harangi, R. Besenczi, I. Lázár, G. Emri, L. Hajdu
University of Debrecen
4002 Debrecen, POB 400, Hungary
Email: hajdu.andras@inf.unideb.hu

R. Tijdeman
Mathematical Institute, Leiden University
Postbus 9512, 2300 RA Leiden, The Netherlands
Email: tijdeman@math.leidenuniv.nl

***Abstract*—In a recent work, we have proposed a novel way to approximate point sets with grids using the LLL algorithm, which operates in polynomial time. Now, we show how this approach can be applied to pattern recognition purposes with interpreting the rate of approximation as a new feature for regularity measurement. Our practical problem is the characterization of pigment networks in skin lesions. For this task we also introduce a novel image processing method for the extraction of the pigment network. Then, we show how our grid approximation framework can be applied with specializing it for the recognition of hexagonal patterns. The classification performance of our approach for the pigment network characterization problem is measured on a database annotated by a clinical expert. Throughout the paper we address several practical issues that may help to apply our general framework to other practical tasks, as well.**

I. Introduction

Texture or pattern analysis has been studied by researchers for many years. In several tasks, textures should be classified as regular or irregular ones, where regularity can be determined by appropriate measures. Regularity measurement is usually applied in cases, where we are intended to compare different data sets or derive information from a data set. Its output can be applied to data characterization or data manipulation like in [1] for regularity-based image filtering, or in [2] for the analysis of human motion patterns. Regularity measurement has been considered widely also in medicine. For example, regularity estimation for epileptic seizure data is proposed in [3], while the regularity of EEG data is analyzed on patients suffering from Alzheimers disease in [4]. In this paper, we introduce a new methodology for regularity measurement in terms of finding well approximating grids of the pattern components based on a novel theoretical approach.

As for the core theoretical background, we consider the Lenstra-Lenstra-Lovász (LLL) algorithm [5] which has been making a reasonable impact in mathematics and computer science since its publication in the early 1980s; see e.g. [6] for a comprehensive overview on its applications. In general, LLL is a very efficient polynomial-time algorithm for finding simultaneous rational approximations to real numbers. In our recent work [7], we have made the theoretical foundation of using the LLL algorithm to obtain well approximating grids for an input set of points in arbitrary dimension. Now, we introduce these results into practice with also specializing it to be able to decide whether the approximated point set has an expected regular appearance. Our work has been motivated by the practical problem to classify pigment networks in skin lesions as regular (typical) or irregular (atypical) ones. The proposed framework considers also a novel image processing approach to extract the pigment network components as an input pattern, and apply an LLL approach to check whether it has an expected regular (in this case hexagonal) behavior.

The rest of the paper is organized as follows. Our framework considered for finding well approximating grids for a point set is discussed in section II including its specialization to recognize hexagonal patterns. In section III, we fit our approach to the problem of characterizing pigment networks in skin lesions. We introduce a novel method for the extraction of pigment networks from dermoscopy images in section IV, and present our experimental results regarding the classification of the extracted networks in section V. Finally, some conclusions are drawn in section VI.

II. Finding well approximating grids

A. The general framework

In its original form [5], the Lenstra-Lenstra-Lovász (LLL) algorithm was introduced as a procedure for grid (lattice) basis reduction. The LLL algorithm became very popular both in theoretical and practical problems [6] partly because it operates in polynomial time. With an appropriate interpretation of the original idea behind LLL, in [7] we have introduced a theoretical approach for the approximation of a point set with grids. In this approach, we apply the LLL algorithm that determines well approximating grids if such grids exist, or in other words, the point set is regular. Our method works in arbitrary dimension, however, regarding the current content, we recall the necessary formalization from [7] for 2D only.

Let $A = \{\boldsymbol{a_1}, \ldots, \boldsymbol{a_k}\} \subset \mathbb{R}^2$ is a finite set of points, which do not fit on a line. To approximate A with a grid, our aim is to find $\boldsymbol{o}, \boldsymbol{d_1}, \boldsymbol{d_2} \in \mathbb{R}^2$ such that the distance of $\boldsymbol{a} - \boldsymbol{o}$ to the grid $\Lambda := \boldsymbol{d_1}\mathbb{Z} + \boldsymbol{d_2}\mathbb{Z}$ is relatively small for every $\boldsymbol{a} \in A$. That is, the point set A is approximated by the grid Λ defined by the grid basis vectors $\boldsymbol{d_1}$ and $\boldsymbol{d_2}$ and having $\boldsymbol{o}$ as its origin. As a usual notation for the translation of a set by a vector, we write $U + \boldsymbol{v}$ for $\{\boldsymbol{u} + \boldsymbol{v} : \boldsymbol{u} \in U\}$ for $U \subset \mathbb{R}^2, \boldsymbol{v} \in \mathbb{R}^2$. As for the cardinality of A, we will assume that $k > 3$, otherwise a trivial perfectly approximating grid can be found. Moreover, notice that an arbitrarily well approximating grid can be found if we let $\boldsymbol{d_1}$, $\boldsymbol{d_2}$ extremely small. Therefore, to take the grid size also into consideration, the following measures have been

introduced in [7] to calculate the error of approximating A by a given grid Λ:

$$N_{\Lambda,\boldsymbol{o}}(A) := \max_{\boldsymbol{a}\in A} \frac{|\boldsymbol{a}-\boldsymbol{o}-\Lambda|}{\Delta} \left(\frac{\text{diam } A}{\Delta}\right)^{\frac{2}{k-3}} \quad (1)$$

and

$$N^{(2)}_{\Lambda,\boldsymbol{o}}(A) := \frac{\sqrt{\sum_{\boldsymbol{a}\in A} |\boldsymbol{a}-\boldsymbol{o}-\Lambda|^2}}{\Delta} \left(\frac{\text{diam } A}{\Delta}\right)^{\frac{2}{k-3}}, \quad (2)$$

where $\text{diam } A$ is the diameter of the point set A, and Δ is the square root of the lattice determinant of Λ calculated as the determinant of the 2×2 matrix $(\boldsymbol{d_1 d_2})$. The error measures $N_{\Lambda,\boldsymbol{o}}(A)$ and $N^{(2)}_{\Lambda,\boldsymbol{o}}(A)$ compensate expected distance errors, scaling and the size of the approximating grid; besides other issues, see [7] for more details.

The method introduced in [7] has a parameter $\epsilon \in \mathbb{R}$, which basically controls the sizes (determinants) of the approximating grids found. However, for each fixed ϵ, the LLL algorithm still generates several different grids, whose basis vectors can be rather various regarding their lengths and enclosed angles.

B. Specialization to a pattern with known structure

In this paper, we investigate such patterns, whose grid structures are approximately known. More precisely, we will focus on patterns having honeycomb-like appearance (see Figure 1a), which is a quite frequent structure in nature. Here, we will analyze pigment networks in skin lesions (see Figure 1b), however, according to the result [8] awarded by Nobel prize in 2014, similar investigations could be performed on motion trajectories stimulating the navigating system of the brain (see Figure 1c).

The assumption that we have prior knowledge on the pattern induces some specialization of the general framework described in section II-A. Namely, the following two issues are worth noticing:

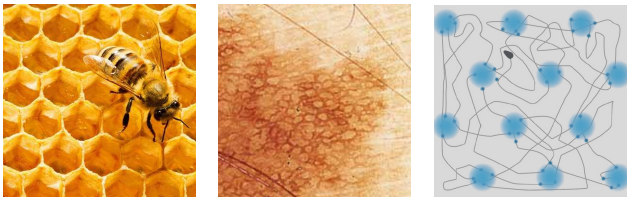
i) If the approximate type of the grid is known (e.g. hexagonal), we can restrict our attention to those grids found by the algorithm in II-A, whose basis vectors fit this condition.

ii) If the approximate size (determinant) of the grid is known, the selection of the parameter ϵ in section II-A might be less critical, since from the prior knowledge we can easily fix an ϵ to support the generation of the approximating grids having the expected determinant.

As a specific example, we will consider the hexagonal grid $\mathcal{H}$ shown in Figure 2, where the grid cells are represented by their centroid and the distance of two grid points is 1. $\mathcal{H}$ can be spanned by the basis $D = (\boldsymbol{d_1 d_2})$ with

$$\boldsymbol{d_1} = (1,0)^T, \text{ and } \boldsymbol{d_2} = \left(1/2, \sqrt{3}/2\right)^T, \quad (3)$$

so it can be naturally considered as a perfect approximation of $\mathcal{H}$ regarding both the error measures defined in (1) and (2). Note that the square root of the grid determinant is $\Delta = \sqrt[4]{3/4} = 0.9306...$, and for the enclosed angle of $\boldsymbol{d_1}$ and $\boldsymbol{d_2}$ we have $\angle(\boldsymbol{d_1 d_2}) = 60°$. The basis vectors given in (3) are



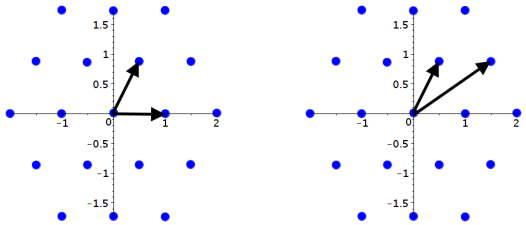
(a) (b) (c)

Fig. 1: Hexagonal grid structures in nature; (a) honeycomb pattern, (b) pigment network in a skin lesion, (c) locations stimulating grid cells of the brain (image source: http://www.nobelprize.org/nobel_prizes/medicine/laureates/2014/press.html).

shown in Figure 2a. The points of the hexagonal grid $\mathcal{H}$, just like other grids, can be spanned by other – actually infinitely many – basis vectors, as well. For example,

$$\boldsymbol{d'_1} = \left(1/2, \sqrt{3}/2\right)^T, \text{ and } \boldsymbol{d'_2} = \left(3/2, \sqrt{3}/2\right)^T, \quad (4)$$

also span $\mathcal{H}$ (see Figure 2a), with $\Delta' = \sqrt[4]{3/4} = 0.9306...$, and $\angle(\boldsymbol{d'_1 d'_2}) = 30°$.



(a) (b)

Fig. 2: The hexagonal grid $\mathcal{H}$ defined by different basis vectors; (a) $\boldsymbol{d_1} = (1,0)^T, \boldsymbol{d_2} = \left(1/2, \sqrt{3}/2\right)^T$, (b) $\boldsymbol{d'_1} = \left(1/2, \sqrt{3}/2\right)^T, \boldsymbol{d'_2} = \left(3/2, \sqrt{3}/2\right)^T$.

That is, for our later decision on the regularity of a grid, we need a proper rule to be able to determine, whether two bases span the same grid or not. For this purpose, we will compare each basis B returned by the LLL algorithm with D using some theoretical considerations. Namely, it is known that if two bases D, B span the same grid, then $\det(D) = \pm\det(B)$ holds. However, comparing only these determinants would not guarantee that the two grids are similar, as e.g. a square grid could easily be defined having the same determinant. Thus, we take advantage of another property, namely that D and B span the same grid, if and only if, there exists an unimodular matrix U (i.e. a square matrix with integer entries and determinant ± 1) such that $B = DU$. Based on this property, for the comparison of D and B we will apply

$$D^{-1}B = D^{-1}DU = U, \quad (5)$$

where D^{-1}denotes the inverse matrix of D. Notice that, the inverse matrices are guaranteed for both D and B, since they are both bases. As a decision rule, we check whether the matrix

U obtained through (5) can be accepted as a unimodular one. That is, if the entries of U are "nearly " integers, and $|\det(U)|$ is close to 1, we consider that B can be accepted to span a hexagonal grid close to that of D.

More precisely, if the respective entries of the matrices of D and U are written as $D = (d_{i,j})$ and $U = (u_{i,j})$, we can introduce three non-negative tolerance parameters with requiring all the following conditions $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3$ to hold to accept U as a unimodular matrix:

- $\mathcal{C}_1$) Determinant: $||\det(U)| - 1| \leq T_{det}$,
- $\mathcal{C}_2$) Entry-wise: $|u_{i,j} - \text{round}(u_{i,j})| \leq T_{ew}, i, j = 1, 2$,
- $\mathcal{C}_3$) Sum of entries: $\sum_{i,j=1}^{2} |u_{i,j} - \text{round}(u_{i,j})| \leq T_{sum}$,

where round(x) stands for the integer closest to the argument x. The appropriate selection of the tolerance parameters should be tuned to the specific application which issue will be addressed later on also in our practical problem.

Notice that, the above formulation for the decision of the presence of a pattern is not scale-invariant, since it compares the grid basis found by LLL with that of the fixed sized $\mathcal{H}$. Thus, to provide some tolerance, we introduce a corresponding positive scale parameter λ, and test a specific U_λ matrix calculated by the appropriate modification of (5)

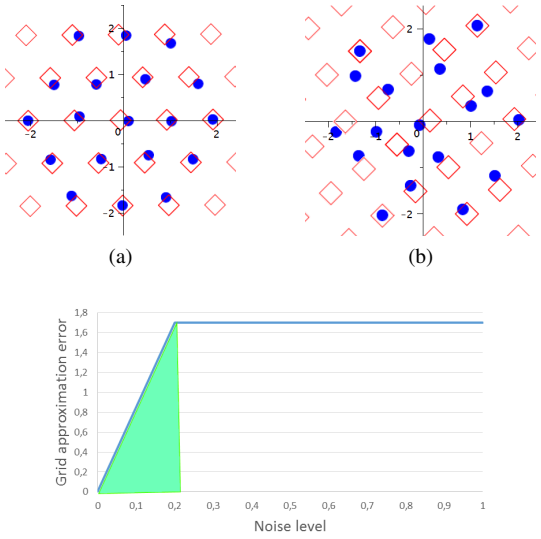
$$\lambda D^{-1} B = U_\lambda, \tag{6}$$

for a specific range of scales $\lambda \in I$. E.g., $I = \{1\}$ provides a strict decision, while $I = \{0.9, 1, 1.1\}$ leads to a bit larger tolerance regarding the expected pattern size.

C. Estimating the rate of regularity

Still if the observed pattern has some grid-like structure, perfect regular appearance can be hardly expected in applications. That is, the pattern is just close to the one we consider as a theoretical model. That is, to be able to decide whether the pattern should be considered as a realization of the theoretic one, we must discover the rate of natural noise to what extent the pattern may differ. After this step, a natural decision on the pattern can be made based on whether its difference from the theoretic grid is below this noise level or not. Naturally, larger reliability can be expected if this noise level can be kept low, that is, the pattern has a rather strict grid-like appearance. On the other hand, the tolerance threshold cannot exceed the level above which the pattern suggests a random behavior.

To check the effect of the natural noise to the appearance of the pattern, we have made some simple simulations with artificially distorting the hexagonal grid $\mathcal{H}$. For a precise formulation, we have selected the point set $A = \alpha \boldsymbol{d_1} + \beta \boldsymbol{d_2}$ for some specific $\alpha, \beta = \pm 2, \pm 1, 0$ to be the 19 points of $\mathcal{H}$ shown in Figure 2 and applied the noise to A in terms of translations of its points with some random figures. The random translation values for both coordinates have been selected from the interval $[-1, 1]$ with noting that a maximum level of noise is already able to exchange two neighboring grid points. After applying different levels of noise, we have executed our LLL algorithm to find well approximating grids. From these grids, we have selected the one having the smallest error with respect to $N^{(2)}_{\Lambda,\boldsymbol{o}}(A)$ defined in (2). Notice that similar results can be obtained by using the error term (1). In Figures 3a and 3b we show two examples for different levels of noises with the best approximating grids found. It can be observed that in the case of small error, the hexagonal structure is still preserved, however, in the case of larger noise it is possible that the best approximating grid found by the LLL method will not be the hexagonal one any more.



(a) (b) (c)

Fig. 3: Best approximating grids found by LLL after adding random translational noise to the hexagonal grid (blue dots: deformed hexagonal grid, red squares: approximating grid); (a) noise level 0.1, (b) noise level 0.3, (c) $N^{(2)}_{\Lambda,\boldsymbol{o}}(A)$ error of grid approximation against noise level (approximating grid is accepted as the hexagonal one in the green region).

We present the minimal $N^{(2)}_{\Lambda,\boldsymbol{o}}(A)$ error occuring among the approximating grids found by LLL against the level of noise in Figure 3c. It is natural that the error grows linearly with the noise level, however, it can be also nicely observed that from the noise level which practically transforms A to be a random point set, the grid approximation error becomes approximately constant. In other words, a threshold for the natural noise tolerance should be selected by no means below this level. In Figure 3c, we have also indicated the noise domain by green, where the best approximating grid could be considered as a hexagonal one. For this selection, we set the tolerance levels for conditions $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3$ as $T_{det} = 0.05, T_{ew} = 0.05, T_{sum} = 0.15$. Also notice that in this simulation we can ignore the scale issue addressed with (6), so can fix the scale parameter as $\lambda = 1$.

III. Estimating the tolerable noise level for pigment networks

The pigment network is an important feature in dermoscopy for the detection of skin cancer (malignant melanoma). It is

a grid-like or honeycomb-like structure consisting of round, pigmented lines and lighter hypo-pigmented holes. As for its classification, the pigment network can be either typical or atypical. Slowing growing benign pigmented lesions produce uniform patterns lending a more regular honeycomb-like (hexagonal) appearance to the network. In the typical case, the lines are uniformly populated with benign pigment cells that grow at a slow rate. On the other hand, atypical pigment networks relate more to melanomas consist of malignant melanocytes that vary in size and degree of pigmentation. Malignant cells move through the epidermis in all directions at varying rates, which movement results in structures with bizarre patterns. In such atypical networks, cells are numerous in some locations and sparse in others. More details on this topic can be found in [9].

The above dermoscopic differences of pigment networks suggest the idea to classify them as typical or atypical based on measuring their regularity. Namely, we apply our LLL approach to check whether a point set extracted from the pigment network can be approximated by a regular hexagonal grid or not. We start with showing how the parameters of our framework should be adjusted according to this task.

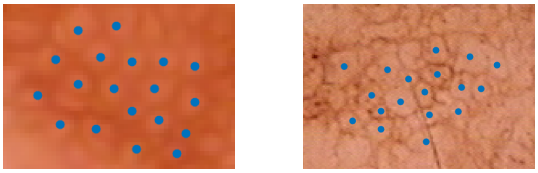
Though the literature describes the typical pigment network as having a honeycomb-like pattern, it has a natural deviation from the regular hexagonal grid. To be able to distinguish this natural noise from atypical appearance, a local dermatologist selected manually the centers of pigment cells in some typical and atypical pigment networks. For these manually annotated point sets, we applied the LLL algorithm to find well approximating grids. From these tests we could determine the appropriate tolerance levels regarding conditions $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$ for the allowed distortion of the regular hexagonal grid. Namely, we have set $T_{det} = 0.3, T_{ex} = 0.15$ and $T_{sum} = 0.5$ to distinguish the typical and atypical appearances.

Moreover, we must address the scale issue described in (6). That is, we have to find that basis B among the ones returned by the LLL approach, which spans the hexagonal grid having the basis λD with minimal error considering (1) or (2). Here, the scale parameter λ falls within an allowable range $[\tilde{\lambda} - \delta, \tilde{\lambda} + \delta]$ with some $\delta > 0$ to have the closest input points to be adjacent in the approximating grid. More precisely, for each input point set $A = \{\boldsymbol{a_1}, \ldots, \boldsymbol{a_k}\}$ we set

$$\tilde{\lambda}_A = \frac{\sum_{i=1}^{k} \min_{j=1, j\neq i}^{k} (\|\boldsymbol{a_i}, \boldsymbol{a_j}\|_2)}{k}, \tag{7}$$

where $\|\boldsymbol{a_i}, \boldsymbol{a_j}\|_2$ denotes the Euclidean distance of the points $\boldsymbol{a_i}$ and $\boldsymbol{a_j}$. As for the proper figures, we had $10 \leq \tilde{\lambda}_A \leq 20$ according to the difference of the sizes of the pigment networks, and δ was adjusted to be $\tilde{\lambda}_A/2$.

When the LLL approach did not find a grid which can be accepted as a hexagonal one with the above setup, the pigment network is considered as an atypical one. On the other hand, if LLL found a basis B meeting the tolerance levels with being sufficiently hexagonal, its approximation error is checked. Based on the manually annotated pigment network images, we consider the network to be typical, if for its approximation error $N^{(2)}_{\Lambda,\boldsymbol{o}}(A) < 3.1$ holds. Notice that, with corresponding scaling this value fits the noise level 0.15 in Figure 3c. In Figure 4a, a typical pigment network is shown which can be approximated by a regular hexagonal grid, while an atypical lesion can be seen in Figure 4b.



(a) (b)

Fig. 4: Manual annotated pigment networks to adjust tolerance levels for hexagonal grid approximation; (a) a typical network, (b) an atypical network.

IV. Pigment network extraction

In order to use our regularity measure for pigment network classification, we should extract the centroids of the pigment cells by an automated method. In this section, we present a new algorithm for the extraction of the pigment network, whose regularity will be further analyzed by our grid matching approach described earlier.

Our extraction procedure works on an intensity image, thus, it starts with a grayscale conversion of the input color image. In the first phase of its operation, the proposed method applies a pixel-wise directional scheme. For every pixel, the intensity values along discrete line segments of different orientations centered at the pixel under examination are considered. The orientations of these segments uniformly cover the range $[0, 180°]$ with a fixed rotation angle $\delta\phi$ between them that was set to $6°$ in our implementation. The vector consisting of the intensity values along such a line segment is referred to as a cross-sectional intensity profile. Its length, i.e., the number of its composing pixels is $2r + 1$. Let P denote such a profile and $P[i]$ the intensity of its its i-th pixel, where $i = -r, \ldots, r$. Thus, the central value, i.e., the intensity of the pixel whose surroundings is being examined is $P[0]$. In our implementation, r was set to 10, however, it will be later shown that this is not a critical parameter. The operation described in the following is applied on all cross-sections of a pixel, resulting in a response value. This way, a directional response vector (DRV) is assigned to every pixel.

The second order derivative of the Gaussian (SDG) filter is widely used in signal and image processing tasks. Its negative normalized variant is often referred to as the Mexican-hat due to its shape. The two dimensional counterpart is known as the Laplacian of Gaussian (LoG) filter and is a fundamental tool in corner, blob and edge detection tasks. The SDG filter including a scale parameter σ is given as

$$G''_\sigma(x) = \frac{x^2 - \sigma^2}{\sigma^4} \exp\left(\frac{-x^2}{2\sigma^2}\right). \tag{8}$$

Favorable properties of the SDG filter are being zero-mean and flattening out to zero. Hence, the response is less dependent on the length of the cross-section than it would be in the case of e.g., zero-mean Gaussian masks. The only constraint is that the cross-section length should be sufficiently large to allow the SDG filter with the largest scale to flatten out. The general way to calculate the matched filter response is the discrete cross-correlation of the signal and the mask. If both data are real and the mask is central symmetric, this is the same as discrete convolution. Thus, both expressions occur in the literature. In the case of the proposed technique, the size of the cross-section profile and the mask are equal. Since only the response at the central position is relevant, the operation simplifies to the sum-product of the two vectors. That is, the matched filter response of an SDG on a profile P is given as

$$M_{P,\sigma} = \sum_{i=-r}^{r} P[i]G''_{\sigma}(i). \tag{9}$$

The most important drawback of this correlation based matching of directional SDGs is the high response for asymmetric bright intensity transitions, such as those occurring at the boundary of air bubbles in dermatoscopic images. The solution proposed here is to split the sum-product (9) to a left and right part, excluding the central element, and subtracting the absolute difference of these two from the entire response. Thus, a symmetry constraint is incorporated. In the case of an SDG filter, the sum-products may be negative, yielding erroneous effects. In such a case, the response is considered to be zero. The formal description can be given as:

$$\begin{gathered} M_{P,\sigma} = M^L_{P,\sigma} + P[0]G''_{\sigma}(0) + M^R_{P,\sigma}, \text{ where} \\ M^L_{P,\sigma} = \sum_{i=-r}^{-1} P[i]G''_{\sigma}(i), \; M^R_{P,\sigma} = \sum_{i=1}^{r} P[i]G''_{\sigma}(i). \end{gathered} \tag{10}$$

This way, the symmetric matched filter response (SM) is expressed as

$$SM_{P,\sigma} = \max(M_{P,\sigma} - |M^L_{P,\sigma} - M^R_{P,\sigma}|, 0). \tag{11}$$

To handle the varying thickness of pigment network lines, SDG masks corresponding to different scale parameters σ are matched against the profile, and the maximal response is selected. Formally, the final multiscale symmetric matched filter (MSM) response for cross-section profile P is given as

$$MSM_P = \max_{\sigma}(SM_{P,\sigma}). \tag{12}$$

In our implementation, the scale values σ range from 1.0 to 3.0 with a step of 0.1.

In order to obtain a single value that describes the likelihood of a pixel being part of the pigment network, a score for every pixel using the statistical measures of its response vector is calculated. A formula that considers the mean μ_{DRV}, standard deviation σ_{DRV} and maximal value $\max_{DRV}$ of the response vector is considered here. The idea behind this approach is that points of elongated structures, such as those of the pigment network, have low or zero response at directions matching the orientation of the corresponding structure. Thus, the variation of responses over the directions is higher. Therefore, the score of a pixel is calculated by multiplying the response vector mean with its standard deviation, and then divide the product by the maximal response value to achieve a certain level of normalization, i.e.

$$score(DRV) = \frac{\mu_{DRV}\sigma_{DRV}}{\max_{DRV}}. \tag{13}$$

To obtain a binary network mask, a hysteresis thresholding technique is applied to the score map. This means that two threshold values are considered, a high ($thigh$) and a low ($tlow$) one, and those pixels are marked as foreground ones, whose score is greater than $tlow$, and are connected to at least one point with a score greater than $thigh$ through an 8-connected path of points with scores greater than $tlow$. The high threshold was set as the mean score, and $tlow = thigh/2$.

The recognition of the pigment network is based on analyzing the holes of the binary network mask, i.e., those 4-connected components of background pixels that do not touch the image borders. Since the network segmentation is not perfect, several such holes will be detected even in images that do not contain pigment networks. Consequently, the holes need to be filtered. First, holes consisting of less than 5 or more than 400 pixels are automatically rejected (at image resolution 768×560). For the remaining holes, a simple convexity measure, the ratio of the area of the hole, and the area of its convex hull is calculated. This measure is known as the solidity, and those hole components are kept, whose solidity is at least 0.8. The remaining holes are decomposed into disjoint groups, based on proximity analysis using a simple maximal distance clustering. In each such group, every hole must be within this maximal distance to at least another hole component. The group is considered to correspond to a pigment network if it is composed of at least four hole components. The number of holes in groups that fulfill this criteria are summed up, and if this number exceeds a certain threshold, then the lesion is considered to contain a pigment network. This threshold was set to 50 in our test implementation. The steps of our proposed pigment network extraction approach are also shown in Figure 5.

V. Experimental results

Currently, few dermoscopic image sets are publicly available like PH^2 [10] or the ones involved in the International Skin Imaging Collaboration (ISIC): Melanoma Project [11]. For our testing purposes, PH^2 was more suitable, since it contains manual annotations regarding the typical/atypical appearance at image-level. These labels are derived from the presence/absence of atypical pigment networks. Namely, if the specialist can detect atypical pigment network in any subregion of the lesion, the whole lesion is considered atypical. That is, if the lesion contains only typical networks, then it is typical but an atypical lesion may contain typical networks, as well. This behavior requires a local investigation on the pigment networks detected in the image, which is supported by the method described in section IV.

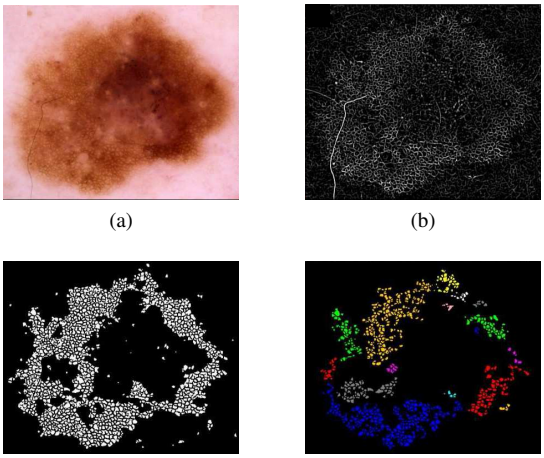


Fig. 5: Steps of extracting the pigment network; (a) a sample dermatoscopic image; (b) the calculated network score map (contrast enhanced for better visualization), (c) the holes of the segmented network, (d) the result of the maximal distance clustering of the filtered hole components.

PH^2 contains 200 8-bit RGB color dermoscopic images (80 common nevi, 80 atypical nevi, and 40 melanomas) taken at $20\times$ optical zoom and resolution 768x560. Only 94 images (39 typical, 55 atypical at lesion-level) contain pigment networks which images are divided into 136 sub-regions containing pigment network segments. However, no annotation is available in PH^2 at network-level regarding typical/atypical behavior. Thus, these network segments were labeled manually by a local dermatologist (62 typical and 74 atypical).

The method described in section IV extracts the pigment cells of the network. The input point set $A = \{\boldsymbol{a_1}, \ldots, \boldsymbol{a_k}\}$ is composed from the centroids of these components and well approximating grids are found by the LLL algorithm. If a grid spanned by a basis B meets the tolerance levels to be sufficiently hexagonal and provide low approximation error, we consider the pigment network typical, otherwise atypical. In this way, the accuracy of automated classification of the 136 sub-regions was found to be 0.71 (with sensitivity 0.82 and specificity 0.58). When a lesion is classified at image-level based on whether it contains atypical sub region or not, we reached 0.76 accuracy. These figures currently cannot be compared with other works, since existing methods like [12], [13] and [14] focus only on the decision on the presence of pigment networks and not on their classification. The advantage of the proposed pigment network extraction method comes from the more extensive use of the information about the local surroundings of the individual pixels through the usage of directional response vectors.

VI. Conclusion

The proposed way for measuring pattern regularity based on grid matching could be useful in other fields, as well. Our effort to develop the theoretical foundation of grid matching was motivated by the clinical problem to classify pigment networks in skin lesions as typical or atypical ones. We have showed how the LLL algorithm can be used for finding well approximating grids if we have some preliminary conditions regarding the expected structure. Moreover, we have introduced a new technique for pigment network extraction. The output of this method is further analyzed using the LLL framework to decide whether a well approximating hexagonal grid can be found ro recognize typical pigment networks. Our experimental results show that our approach is reasonable in this field, and the LLL approach with its polynomial-time complexity is very efficient also for this purpose. However, a proper extraction of the network components is important for accurate grid approximation, and thus, the corresponding image processing algorithms need further improvements.

Acknowledgment

This work was supported in part by the OTKA grant K115479, the projects GINOP-2.1.1-15-2015-00376 and VKSZ_14-1-2015-0072, SCOPIA: Development of diagnostic tools based on endoscope technology supported by the European Union, co-financed by the European Social Fund.

References

[1] D. Chetverikov, "Pattern regularity as a visual key," *Image and Vision Computing*, vol. 18, no. 12, pp. 975–985, 2000.

[2] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabási, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.

[3] N. Radhakrishnan and B. Gangadhar, "Estimating regularity in epileptic seizure time-series data," *Engineering in Medicine and Biology Magazine, IEEE*, vol. 17, no. 3, pp. 89–94, 1998.

[4] D. Abásolo, R. Hornero, P. Espino, J. Poza, C. I. Sánchez, and R. de la Rosa, "Analysis of regularity in the EEG background activity of Alzheimer's disease patients with Approximate Entropy," *Clinical Neurophysiology*, vol. 116, no. 8, pp. 1826–1834, 2005.

[5] A. Lenstra, H. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," *Math. Ann.*, vol. 261, pp. 515–534, 1982.

[6] P. Q. Nguyen and B. Valle, *The LLL Algorithm: Survey and Applications*, 1st ed. Springer Publishing Company, Incorporated, 2009.

[7] A. Hajdu, L. Hajdu, and R. Tijdeman, "Finding well approximating lattices for a finite set of points," *submitted to Mathematics of Computation*, 2016, arXiv:submit/1539667 [math.NT] 20 Apr 2016.

[8] T. Hafting, M. Fyhn, S. Molden, M. B. Moser, and E. I. Moser, "Microstructure of spatial map in the entorhinal cortex," *Nature*, vol. 436, pp. 801–806, Aug. 2005.

[9] "Dermnet skin disease image atlas," http://www.dermnet.com, accessed: 2016-04-11.

[10] T. Mendona, P. M. Ferreira, J. S. Marques, A. R. S. Marcal, and J. Rozeira, "PH^2 - a dermoscopic image database for research and benchmarking," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2013, pp. 5437–5440.

[11] "International skin imaging collaboration (ISIC): Melanoma project," https://isic-archive.com/#.

[12] C. Barata, J. S. Marques, and J. Rozeira, "A system for the detection of pigment network in dermoscopy images using directional filters," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 10, pp. 2744–2754, Oct 2012.

[13] M. Sadeghi, M. Razmara, T. K. Lee, and M. Atkins, "A novel method for detection of pigment network in dermoscopic images using graphs," *Computerized Medical Imaging and Graphics*, vol. 35, no. 2, pp. 137 – 143, 2011, advances in Skin Cancer Image Analysis. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0895611110000674

[14] C. Barata, J. S. Marques, and J. Rozeira, "A system for the automatic detection of pigment network," in *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, May 2012, pp. 1651–1654.