Lane Background Removal in Thin-Layer Chromatography Images Using Continuous Wavelet Transform

Bruno M. Moreira[1,3]
moreira@fe.up.pt

António V. Sousa[1,2]
ats@isep.ipp.pt

Ana M. Mendonça[1,3]
amendon@fe.up.pt

Aurélio Campilho[1,3]
campilho@fe.up.pt

[1] Instituto de Engenharia Biomédica, Universidade do Porto.

[2] Instituto Superior de Engenharia, Instituto Politécnico do Porto

[3] Faculdade de Engenharia, Universidade do Porto

campilho@fe.up.pt

Abstract

This paper describes a new methodology to remove the background of the lanes in Thin-Layer Chromatography (TLC) images aiming at improving subsequent band detection. The storage of the biological samples to be analysed by TLC is usually done via plastic containers. Filter paper is an alternative that allows reduced costs and higher portability, but with consequences in the image analysis stage due to a lane background alteration. In order to overcome this problem, a negative control lane is generated in every chromatographic plate. After pre-processing and lane detection stages a one-dimensional intensity profile is used for integrating lane information and the background influence is removed with the help of the Continuous Wavelet Transform (CWT) decomposition. The proposed method was tested in 78 lane images, A band detection algorithm was applied on lane profiles, and a superior detection rate was achieved for the background removed lanes.

1 Introduction

The objective of this work is the development of a methodology to improve the automatic analysis of digital images of Thin-Layer Chromatography (TLC) [1] plates. The biological samples to be analysed by TLC are usually stored and transported in plastic containers. Instead of this conventional sample storage, filter paper can be used since it presents the main advantages of reduced costs and higher portability. However, the use of filter paper has consequences in the image analysis stage due to the diminished concentration of the compounds, as well as the inclusion of additional substances in the sample, which causes a lane background alteration when compared to what is expected without the use of filter paper. In order to use image analysis techniques for automating the detection of the main compounds that are present in a sample, it is important to minimize the background influence. For this purpose, the material resulting from the application of the extraction process to plain filter paper (without the addition of any kind of biological material) will be included in every TLC plate for generating a negative control lane. As this negative control lane will only contain the result of chromatographic separation of the filter paper components, the associated lane pattern can be used for removing the background of all the other sample lanes.

Fig. 1 presents a digital image of a TLC plate containing several vertical lanes each one corresponding to the development of a sample. The image pattern of each lane is formed by horizontal bands associated with the specific sample compounds. The image clearly illustrates the influence of filter paper usage. The leftmost lane (limited by the blue vertical lines) is the negative control lane; the samples in following 10 lanes were extracted from filter paper, while all the other lanes on the right resulted from the chromatographic development of samples stored using the conventional containers. The rightmost lane, delimited by the black lines in the figure, contains a standard combination of known compounds (standards lane). The two sample lanes delimited by the red lines were obtained using exactly the same quantity of a specific compound (Gb3). As can be observed, besides the distinctive background, the bands corresponding to Gb3 also have a completely different contrast.

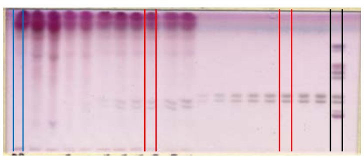


Figure 1: **a**. TLC image. The negative control lane, two sample lane examples and the standards lane are delimited by the blue, red and black lines, respectively.

2 Methodology

Several image and signal processing operations are applied to a raw TLC image to extract a set of lanes. A short description of these techniques can be found in subsection 2.1.

In order to accomplish a satisfactory attenuation of the background present in the lanes, various methodologies were proposed and implemented by the authors. The one herein presented is based on the continuous wavelet transform decomposition of the lane profile.

2.1 Pre-processing and lane detection

The processing flow applied to the original image of a TLC plate is illustrated in the block diagram of Fig. 2. Two distinct regions can usually be found in a typical image: an internal region of interest (ROI) and an external border that has no relevant information for the image analysis process. The ROI is detected using a clustering based algorithm described in [2] and afterwards converted to grey scale. The ROI background is estimated and removed using a closing morphological operator with a large square structuring element. The image is then projected on the direction perpendicular to lane development (vertical projection) to integrate all the image data into an intensity profile. As during background removal the image is inverted, the intensity profile presents maxima regions where the original image is darker (lane zones) and minima regions where this image is lighter (empty zones). The intensity profile is then smoothed using a Savitzky-Golay filter, and lane detection proceeds as described in [3].

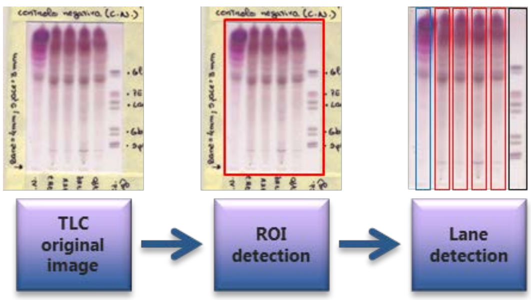


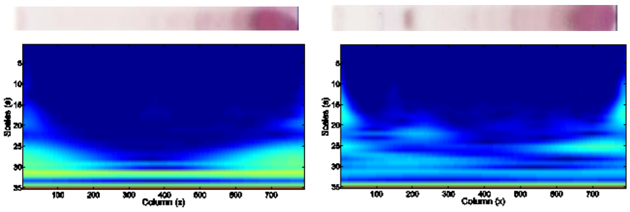
Figure 2: Main steps of the methodology used for obtaining a set of lanes ready for classification.

2.2 Lane background removal

The information of each image lane is integrated into an intensity profile by projection along the sample development direction. The profiles are afterwards smoothed using a procedure based on the Continuous Wavelet Transform (CWT) decomposition. The CWT performs signal decomposition through a set of orthonormal functions generated by the scaling and translation of two functions: the mother wavelet and the daughter wavelet. As a result of the wavelet transform, a set of coefficients hierarchically organized is produced, each one associated with a spatial position and a frequency component (scale) of the original profile [4].

The scales in the CWT are not constrained and the wavelets are non-orthogonal. This property, while making CWT redundant, provides a finely detailed description of a signal in terms of both time and frequency [5]. These characteristics allow an accurate selection of scales that will be important in the smoothing of the intensity profiles, since we pretend to adapt the level of smoothing to the type of lane (sample or negative control). The smoothing level can be controlled by selecting a cutoff in the scale domain and by forcing all the coefficients at lower scales to zero. A cutoff value at a high scale (low frequency) will ensure greater smoothing than a cutoff at a low value (high frequency).

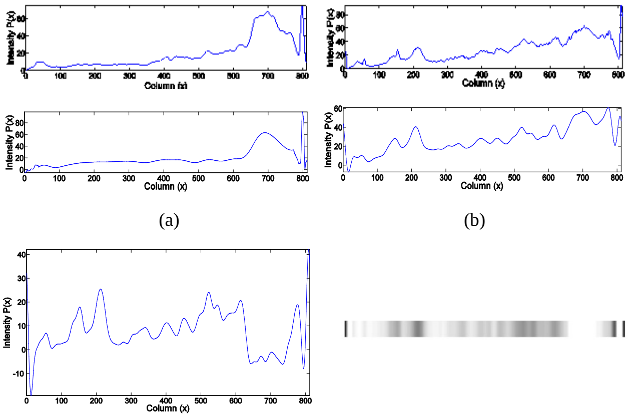
Fig. 3 depicts the images of two lanes (top) and their wavelet decomposition (bottom), with a negative control lane shown on the left and a sample lane on the right. This representation helps to perceive which range of scales contain the most important coefficients for each type of lane. From the analysis of several examples, it was possible to conclude that the majority of the high amplitude coefficients in the negative control lane (Fig. 3a) are located at high scales (low frequencies). In this case, it is advantageous to use a high cutoff value to remove most of the noise and retain just the low frequency components which are associated with the lane background that is expected to occur also in all the other lanes. In what concerns the sample lanes (Fig. 3b), there are relevant low scale (high frequency) coefficients related with the presence of bands. For this type of lanes, the cutoff value should be lower than the one used for the negative control lanes. The cutoff values were estimated based on a randomly selected set of negative control and sample lanes.



(a) (b)

Figure 3: **a**. Negative control lane **b**. Sample lane.

For background removal, the smoothed profile of the image negative control lane is subtracted to the smoothed sample lane profile. Fig. 4 shows the process of background removal for the lane on the right of Fig. 3. Fig. 4a and Fig. 4b present the original (top) and smoothed (bottom) profiles for the negative control lane and the sample lane, respectively. The profile resulting from background removal is presented in Fig. 4c, while Fig. 4d shows the reconstructed lane image obtained by profile replication.



(a) (b)

(c) (d)

Figure 4: **a** Negative control lane original (top) and smoothed (bottom) **b**. Sample lane original (top) and smoothed (bottom) profiles. **c**. Result of background removal. **d**. Reconstructed lane image.

3 Results

The proposed approach was tested in a total of 78 lanes, each one with 4 regions where a band could exist (312 potential bands). The lanes were previously examined by a specialist for marking the true bands (175).

For band detection, the lane profile in each of the 4 potential band positions is analyzed for calculating band densities. The two local profile minima that define the band width are found, and the area limited by the profile and the line connecting these two minima is used for band area calculation. Band density is obtained as the ratio between band area and band width. The results of this process are depicted in Fig. 5. A band is detected when the calculated density is above a threshold value derived from the corresponding band density in the standards lane.

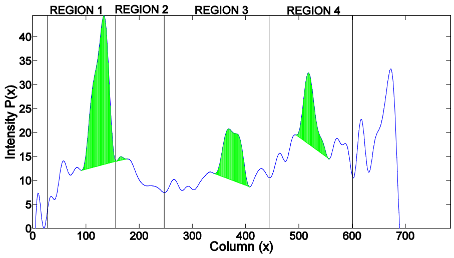


Figure 5: Example of band detection for a generic sample lane profile.

In order to evaluate the lane background removal process, the results of band detection obtained with and without background were used to obtain the two ROC curves shown in Fig. 6, which were obtained by a variation of the threshold value used for band detection.

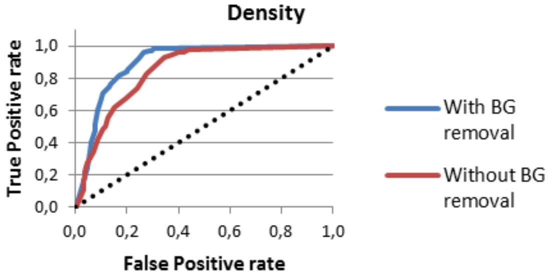


Figure 6: ROC curves for band detection obtained with and without lane background removal.

4 Conclusions

The additional problems found in the analysis of chromatography images when filter paper is used to store and transport the biological samples were addressed by including a negative control lane in every TLC plate. The integration of lane information into a one dimensional profile along with an efficient smoothing proved to be effective for the estimation and removal of the lane background influence.

As future work we intend to improve the methodology for obtaining the intensity profile by a previous compensation of some distortions that occurs in the bands, such as bending and smiling effects. We also intend to measure other features besides band density to improve the robustness of the band detection process.

Acknowledgements

The authors would like to thank Prof. Maria Clara Sá-Miranda from IBMC for providing the datasets used in this work. This work was financed by FEDER funds through the Programa Operacional Factores de Competitividade – COMPETE and by Portuguese funds through FCT – Fundação para a Ciência e a Tecnologia in the framework of the project PTDC/SALBEB/ 100875/2008.

References

[1] B. Fried and J. Sherma, Thin-Layer Chromatography, Marcel Dekker, New York, 1999

[2] A.V. Sousa, M. C. Sá-Miranda, A. M. Mendonça and A. Campilho, "Classification-Based Segmentation of the Region of Interest in Chromatographic Images," in A.C. Campilho and M.S. Kamel (eds.) ICIAR 2011. LNCS, Springer, Vol. 6754, pp. 68-78, 2011.

[3] B.M. Moreira, A. V. Sousa, A. M. Mendonça and A. Campilho. Automatic Lane Detection in Chromatography Images. LNCS 7325, Part II, pp. 180-187i, Springer, 2012.

[4] F. Chau, Y. Liang, J. Gao and X. Shao (2004) Chemometrics – From Basics to Wavelet Transform, John Wiley & Sons, New Jersey.

[5] J. S. Walkers, A primer on wavelets and their scientific applications, Chapman & Hall/CRC, 2008.