Correction of Geometrical Distortions in Bands of Chromatography Images

Bruno Moreira[1,2]
António Sousa[1,3]
Ana M. Mendonça[1,2]
Aurélio Campilho[1,2]

[1] Instituto de Engenharia Biomédica, Universidade do Porto
[2] Faculdade de Engenharia, Universidade do Porto
[3] Instituto Superior de Engenharia, Instituto Politécnico do Porto

Abstract

This paper presents a methodology for correcting band distortions in Thin-Layer Chromatography (TLC) images. After the segmentation of image lanes, the intensity profile of each lane column is spatially aligned with a reference profile using a modified version of the Correlation Optimized Warping (COW) algorithm. The proposed band correction methodology was assessed using 105 profiles of TLC lanes. A set of features for band characterization was extracted from each lane profile, before and after band distortion correction, and was used as input for three distinct one-class classifiers aiming at band identification. In all cases, the best results of band classification were obtained for the set lanes after band distortion correction.

1 Introduction

Thin-Layer Chromatography (TLC) lanes contain bands that are often affected by geometric distortions resulting from the sample development process. Band deformation can modify significantly the morphology of the lane intensity profile and bias the extraction of features for band characterization, such as band area and band peaks. In extreme circumstances, the distortion can even prevent the occurrence of a band in the intensity profile either by its integration into another band or into the lane background. An example of a typical TLC image can be observed in Fig. 1, where several bands with and without distortion are marked by the continuous and dashed line rectangles, respectively.

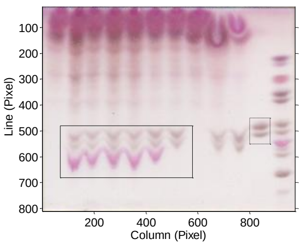

Figure 1: Region of interest of a TLC image. The continuous and dashed line rectangles mark bands with and without distortion, respectively.

Methodologies to deal with band distortion in images similar to TLC images, such as electrophoresis, have been proposed before. The band straightening algorithm developed for the GASepo system [1] significantly reduces band geometric distortions in Epo images. The main idea behind this methodology is an automatic search for an optimum vertical shift in all image columns that leads to a maximum reduction of local geometric distortions. In [2] an algorithm is proposed for removing band distortion in electrophoretic gels using a gradient filter to estimate band orientations. Bands are smoothed and interpolated by robust regression, and integrated to find the unwarping transform. In [3] a two-step method is proposed. Firstly, a simple physicochemical model is formulated and applied for warping each gel image to correct for leakage across the sides during electrophoresis. Secondly, the alignment is obtained by maximizing a penalized likelihood criterion. The likelihood measures the similarity between aligned images and the penalty is a distortion criterion.

This paper describes a methodology for improving the quality of digital images of TLC plates consisting in the correction of band distortions, thus facilitating the subsequent characterization of bands associated with relevant biomarkers. In our application, we are interested in the identification of Fabry disease (FD). For this purpose, we will focus on the detection a specific biomarker, the globotriaosylceramide (Gb3).

2 Methodology

The proposed method for band distortion correction is a modified version of the Correlation Optimized Warping (COW) algorithm [4] that overcomes some of the problems that came out from the application of the original solution to TLC chromatograms of human urine samples.

The COW algorithm was introduced by Nielsen et. al. [4] aiming at the correction of misalignments and shifts between chromatographic 1-D profiles by aligning a sample profile with a reference one by enlarging or compressing sample segments using linear interpolation. An equal number of segments of length (I) is specified in the reference and sample profiles. The maximum shift in the length of a sample segment is controlled by the slack parameter (t).

The modified COW algorithm presented by Skov et al. in [5] includes an automatic method for the selection of the reference chromatogram and an optimization of the I and t parameters through the maximization of the warping effect. This quantitative measure combines two functions with equal weight, the simplicity value and the peak factor [5]. The simplicity value allows to measure how well aligned a set of profiles is. The peak factor tries to guarantee that no strong compression (or enlarging) is required to align the profiles.

The herein proposed methodology for correcting deformations in TLC bands is based on the improved COW algorithm developed by Skov et al. [5]. Our algorithm will be applied in TLC images with a different objective as, instead of aligning different chromatographic profiles, we want to correct band distortions such as smileys, frowns and slopes. These three types of distortion are illustrated in Fig. 2.

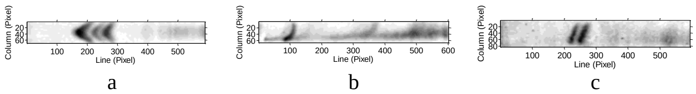

Figure 2: Typical distortions found in TLC bands. (a) Smiley. (b) Frown. (c) Slope.

Since the original algorithm in [5] could not achieve satisfactory results, it was adapted to cope with the limitations that were found. The main problems related to the original algorithm are depicted in Fig. 3. Fig. 3a presents the original images and Fig. 3b shows the images after band correction with the approach described in [5] (using the original code provided by the authors): in the top row the contour of an original flawless band becomes irregular; in the middle row, band positions are modified by the automatically selected reference and the column alignment is spread to other lane regions; in the bottom row, the segment boundary separates the band and the deformation correction is unsuccessful. Fig. 3c presents the results obtained with the application of the herein proposed methodology, which is not affected by the described problems.

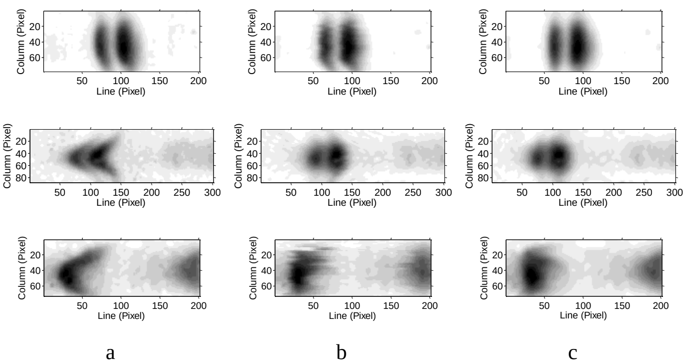

Figure 3: Correction of band distortions. (a) Original images. (b) Results obtained with the original algorithm described in [5]. (c) Results obtained with our approach.

The lanes in the TLC plate image are segmented using the methodology described in [6]. Each lane column is represented as an intensity profile to be aligned with the lane reference. The product of the correlation coefficients between all individual profiles was suggested in [5] as an appropriate solution for selecting the reference column from the lane. However, this approach is not included in our algorithm as in a deformed band the central columns are normally those that better describe the correct position. So, the reference profile, $R(y)$, for a lane image $L(x,y)$ is obtained by

$$R(y) = \frac{1}{2S+1}\sum_{i=-S}^{S} L\left(\frac{N}{2}+i, y\right), \qquad y = 1 \cdots M \tag{1}$$

where M is the number of rows and N is the number of columns in the lane image. S is the number of columns on each side of the lane centre that is included in the averaging of the reference profile. In this work, a value of S=3 was selected.

The profile partition into equal length segments as performed in COW can result in the separation of a single band in two segments to be disjointedly aligned, which makes the result highly dependent on the localization of segment limits. To overcome this problem, the implemented algorithm only applies the warping function in selected lane regions corresponding to groups of bands, identified as local maxima regions in the lane intensity profile. As a result of this region selection process, nearby bands are grouped together and aligned as a whole, but at the same time each group of bands is warped independently from all the others.

In our implementation, for calculating the warping effect value different weights are assigned to the simplicity and peak factor, being the simplicity weight three times higher than the peak factor weight. In band correction, a strong compression (or stretching) is often necessary due to the high displacement between the profiles, which inevitably creates a low peak factor. This overly influences the warping effect and prevents the alignment of the columns further away from the centre. Thus, it was decided that the simplicity value should have a greater impact in the choice of the correct values for I and t.

3 Results

The application of the proposed methodology is exemplified in Fig. 4. A sample lane from Fig. 1 (lane 6) is shown in Fig. 4a. The result of the aligning process on the lane bands can be observed in Fig. 4b, where the "smiley" effects have disappeared. Figs. 4c and 4d show the reference and a column intensity profile, before and after band correction, respectively. The methodology allows profile displacement and peak alignment, as can be observed in Fig. 4d. The mean lane profile that is obtained after the column projection over the lane development direction is presented in Figs. 4e and 4f, before and after the band correction process, respectively. The middle band is visible in the corrected lane profile and the peaks present a higher intensity value than in the original profile.

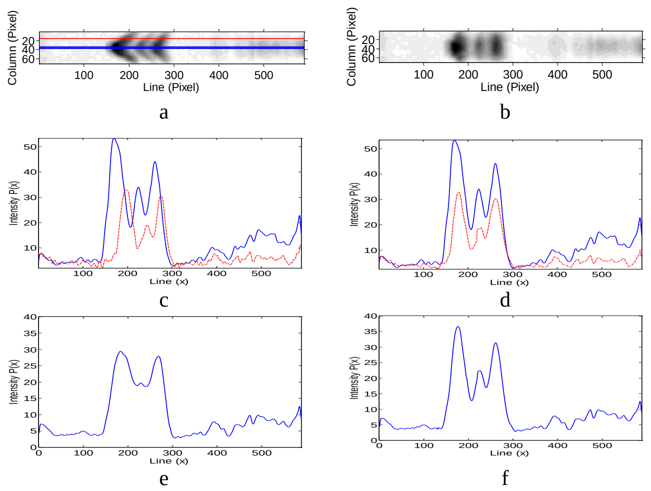

Figure 4: Illustration of the results obtained with the band correction methodology. (a) Original lane of a TLC image. (b) Lane reconstruction after band correction. (c-d) Reference profile (continuous line) and profile obtained from a column (dashed line) before alignment (left) and after alignment (right). (e-f) Integration of lane information before band correction (left) and after band correction (right).

The quantitatively evaluation of band correction was achieved by identifying the Gb3 band in specific regions of the profile, before and after the application of the methodology. A set of 105 regions (53 bands and 52 empty regions) was used. Band identification was based on a set of four features extracted from each lane profile, before and after band distortion correction, and was used as input for three distinct one-class classifiers [8]. The mean and standard deviation values, obtained after 100 runs of 10-fold cross validation method, are shown in Table 1. Both the mean error and variance decreased after band correction using the proposed algorithm. The random behaviour of the original COW prevented the achievement of satisfactory results.

Table 1. Mean error and standard deviation

	Gaussian	Parzen	SVM
Before band correction	8.2% ± 0.9%	6.9% ± 1%	8.7% ± 0.5%
Original COW [5]	7.9% ± 0.6%	8.3%± 1%	8.9% ± 0.9%
Proposed algorithm	5.2% ± 0.2%	4.9% ± 0.1%	5.9% ± 0.2%

4 Conclusion and Future Work

A new methodology for the automatic correction of band deformations in chromatography images based on the Correlation Optimized Warping has been proposed. The method overcomes some limitations of the original algorithm when applied to TLC images. The advantages of band correction were validated both visually by inspecting the reconstructed lanes and quantitatively by using the same set of features, measured before and after the application of the methodology, to identify Gb3 bands. The improvement of characteristics related to the presence of bands, such as the peak intensity values, lead to an increase in the number of correct identifications.

Acknowledgements

The authors would like to thank the Portuguese Foundation for Science and Technology (FCT) (Ref. FCT PTDC /SAU-BEB/100875/2008) and FEDER (Ref. FCOMP-01-0124-FEDER-010913) for co-funding this research project.

References

[1] I. Bajla, I. Holländera, M. Minichmayra, G. Gmeinerb and Ch. Reichelb "GASepo—a software solution for quantitative analysis of digital images in Epo doping control." Computer Methods and Programs in Biomedicine, Vol 80(3), pp. 246-270, 2005.

[2] C. Glasbey, L. Vali, and J. Gustafsson, "A statistical model for unwarping of 1-D electrophoresis gels". Electrophoresis, Vol 26(22), pp. 4237-4242, 2005.

[3] J.S. Gustafsson, A. Blomberg, and M. Rudemo, "Warping two-dimensional electrophoresis gel images to correct for geometric distortions of the spot pattern". Electrophoresis, Vol 23(11), pp. 1731-1744, 2002.

[4] N.P.V. Nielsen, J.M. Carstensen, and J. Smedsgaard, "Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimised warping." Journal of Chromatography A, Vol 805(1–2), pp. 17-35, 1998.

[5] T. Skov, F. van den Berg, G. Tomasi and R. Bro "Automated alignment of chromatographic data". Journal of Chemometrics, Vol 20(11-12), pp. 484-497, 2006.

[6] B. Moreira, A.V. Sousa, A. M. Mendonça and A. Campilho, "Automatic Lane Detection in Chromatography Images in Image Analysis and Recognition", in A. Campilho and M. S. Kamel, (eds.) ICIAR 2012, LNCS 7325, Springer Berlin / Heidelberg. pp. 180-187, 2012.

[7] P. Soille, A.B. Smith and E.F. Roberts, Morphological Image Analysis – Principles and Applications, Springer, 2004.

[8] D. Tax and R. Duin, "Combining One-Class Classifiers", in J. Kittler and F. Roli, (eds), Multiple Classifier Systems, LNCS 2096, Springer Berlin / Heidelberg. pp. 299-308, 2001.