Auto-Cropping of Phone Camera Color Images to Segment Cardiac Signals in ECG Printouts

Fernando Lozano-Fernández[1], Inmaculada Mora-Jiménez[1], Margarita Sanromán-Junquera[1],
Sergio Muñoz-Romero[1], Arcadio García-Alberola[2], José Luis Rojo-Álvarez[1]

[1] Universidad Rey Juan Carlos, Fuenlabrada, Madrid, Spain
[2] Hospital Virgen de la Arrixaca, Murcia, Spain

Abstract

Digital analysis of the bio-electrical signals allows us to obtain valuable information for clinical and research purposes. Despite current devices record and store the digital electrical signal, in some cases only a printout version is available. Although previous research focused on algorithms to digitize a scanned version of the printed signal, nowadays a color image of the printout can be immediately obtained with a cell phone. Most of these algorithms digitize the signals after manually selecting from the whole image one sub-image per signal to digitize. Here, propose a procedure to automatically crop the image to get as many sub-images as signals are in the ECG printout. First, the perspective distortion is corrected. Then, an image I mainly containing the grid, is obtained from a transformation of the RGB space to another colorspace taking advantage of the grid chromatic characteristics. The grayscale image, G, and I are morphologically processed to emphasize both signal and grid in G, and just grid in I. Boundaries between contiguous signals are set by means of projections and morphological operations.

Thirty one phone camera color images of 12-lead-ECG printouts were used for evaluation. Experiments were conducted with images with different spatial resolutions and quality factors according to the jpeg compressor. Our procedure worked correctly when the tracing of consecutive leads did not overlap in the amplitude axis and the spatial resolution and the quality factor were higher than 0.25 (one-fourth of the original size) and 40%, respectively.

1. Introduction

Digital biomedical signal processing allows to automatically extract useful information to support clinicians' diagnosis or treatments. However, the digital signals are not always available because they are stored on paper or in a digital proprietary format. For those signals printed on paper, it would be very useful to have an automatic procedure converting the printed signal in a digital signal.

Previous scientific research has proposed procedures to digitize the biomedical printed signals, mainly electrocardiograms (ECG) or electrograms, from images obtained with a flat scanner [1–3]. These digitalization procedures required individual images for each lead (signal tracing), and also that the original grid geometry is preserved. Hence, each lead was either manually separated or the original image only covered one lead. The flat scanner is not always accessible and the current cell phone or tablet camera allows taking a RGB image of the signal printouts with resolution as good as the scanner image. In this context, Mitra et al. proposed an algorithm to extract and analyze ECG features from cell phone images, but they did not take into account the perspective distortion introduced by the camera position and the automatic cropping of each lead for the digitalization [4].

Here, we propose a procedure to automatically crop a color photo of an ECG printout in as many sub-images as signals are in the picture. The ECG grid can be solid or dotted, and can also have different colors. Our procedure starts by correcting the perspective distortion. Then, a stage based on projections and mathematical morphology [5] allows to find the boundaries between contiguous signals. Finally, different methods available in the literature can be used to digitize every individual signal [3].

The structure of this paper is as follows. In Section 2, the perspective correction procedure is detailed. Section 3 explains the procedure to cancel the grid. Boundaries for image cropping are found in Section 4. Section 5 presents the results of the auto-cropping procedure on a set of ECG images obtained with a conventional phone camera. Finally, conclusions are summarized.

2. Perspective distortion correction

In an ECG, the electrical activity of the heart is represented in a grid with the horizontal/vertical direction corresponding to time and amplitude, respectively. When the image is taken from a flat scanner, the paper and focal

plane are parallel and theoretically no distortion is introduced in the scanned image. However, when holding manually a cell phone, it is difficult that the paper and the camera focal plane are parallel and then some distortion occurs. This kind of distortion is evidenced because grid lines representing the same feature (time or amplitude) are not parallel, and perpendicularity between original grid lines is also lost. A solution to this issue is to estimate the perspective transformation applied to the ideal image and then reversing this transformation so that perpendicularity between grid lines is preserved as much as possible.

A simple perspective transformation can be modeled as a linear transformation between pairs of points in two images according to matrix $\mathbf{H}$. This way, $\mathbf{P} = \mathbf{HR}$, where $\mathbf{R}$ is a matrix storing the Cartesian coordinates of a set of points in the non-distorted image. Matrix $\mathbf{P}$ has the Cartesian coordinates of the same set of points of $\mathbf{R}$ in the transformed (distorted) image. Algorithms like Direct Linear Transform [6] just require four correspondences to find matrix $\mathbf{H}$, with the restriction that no 3 points can be collinear.

Our procedure for perspective correction makes use of the properties of the printed ECG grid and requires the intervention of the user to mark four feature points in the distorted image. These points correspond to the intersection points of two pair of perpendicular grid lines, which delimit a rectangle in the printed grid. Thus, the user provides the four vertices of a quadrilateral in the image taken by the camera, and they are stored in $\mathbf{P}$. The ideal rectangle (vertices in $\mathbf{R}$) is built such that its sides pass through the midpoint of every side in the quadrilateral.

Then, the homography matrix $\mathbf{H}$ is estimated from $\mathbf{P}$ and $\mathbf{R}$, and the inverse transformation $\mathbf{H}^{-1}$ is applied to the distorted image. Fig. 1 (a) shows the distorted image and the quadrilateral built with the vertices pinpointed by the user. The image after perspective correction is in Fig. 1 (b). In this work, the function *FindHomography* of the OpenCV library [7] has been used to estimate $\mathbf{H}$, and *perspectiveTransform* to get the image I without perspective distortion. We will call I_R to the grayscale conversion of I.

3. Grid cancelation

Many algorithms for ECG signal extraction assume that the image only contains the tracing of one signal. Since the perspective correction algorithm needs at least four points, the user should mark four points per signal. To avoid so much user intervention, we propose an automatic procedure to crop the image obtained from Section 2 in the same number of sub-images as tracings from different signals.

Our procedure to find the boundaries between signals starts by obtaining a binary image just containing the signal tracings. From a conceptual point of view, this image can be obtained by canceling the pixels associated to grid lines, which mostly are reddish or green (though other col-
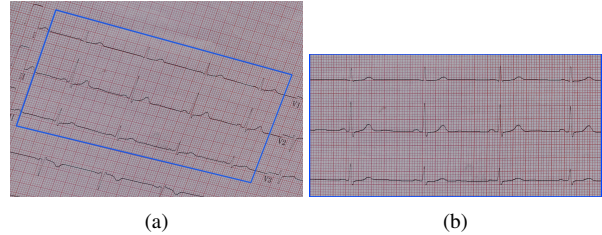

Figure 1. (a) Distorted original photograph; (b) Corrected image. Artificial quadrilateral and rectangle are in blue.

ors are also possible). Since pixels belonging to both signal tracing (blank ink) and paper (white) have achromatic colors, an strategy to find the grid lines is to determine chromatic pixels. Though this is a theoretically valid reasoning, it does not hold when the image is saved by the smartphone. This, along with the fact that the camera lens can also add some chromatic distortion, makes it necessary a transformation from the *RGB* space to another space where the grid color is emphasized and therefore pixels belonging to the grid are more easily determined.

Since both pixels of grid lines and signal tracing have the lowest intensity, we propose to emphasize them by applying a morphological bottom-hat to I_R. For this purpose, the size of the structuring element (SE) must be higher than the size of the elements to highlight. Though this size can be arbitrarily large, we have chosen it to be the size of the larger square in the ECG graph paper. It is computed from the number of pixels between consecutive grid lines, as explained in the next paragraph. After morphological processing, contrast is enhanced so that 2% of the transformed intensity values are saturated. The resulting image, with highlighted grid and tracings, is named I_{gt}.

To find the average number of pixels between consecutive grid lines, the horizontal projection of I_R is computed and normalized to a maximum of 1, obtaining p_h. Maxima of p_h are computed, and those not exceeding a threshold are filtered. The threshold is not constant and it is obtained using a median sliding window (length 0.1 times the projection length) on p_h and adding an offset (experimentally set to 0.1). The distance between consecutive maxima will determine the number of pixels between consecutive horizontal grid lines, so the average distance δ_{gl} is computed.

Since the image can have shadows and so the intensity is not uniform, a normalization of the *RGB* colorspace to the *rgb* colorspace [5] is explored. The *rgb* colorspace has the advantage of representing chromaticity invariantly with respect to intensity. Components r and g of the normalized space should be adequate to determine the grid pixels in reddish and green grids, respectively. Other space colors separating intensity and chroma were also considered. We experimentally checked that the C_r component of the YC_bC_r colorspace allowed us to determine the red-
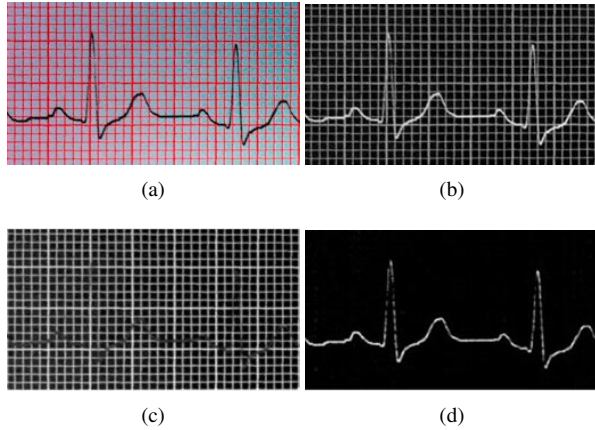
Figure 2. Example of the grid cancelation process on a fragment of the ECG: (a) I; (b) I_{gt}; (c) I_g; (d) I_t.

dish grid better than r, since less pixels belonging to signal tracing were identified with C_r. Then, C_r is chosen for reddish grids and g for green ones. For other grid colors, the relative color difference (rcd) is used [8]. Then, C_r, g or rcd is chosen (depending on the grid color) and the top-hat morphological operator is applied with a flat, disk-shaped SE of diameter δ_{gl}. Image I_g, which mainly preserves the grid, is obtained after contrast enhancement.

Finally, image subtraction of I_{gt} and I_g is computed to get I_t. Fig. 2 shows the images obtained along the cancelation grid process. Note that this procedure has the side effect of producing gaps in the tracing, as shown in Fig. 2 (d).

4.   Segmentation boundaries

The goal is to find a set of horizontal lines that best segment I in n_l subimages, each containing only one signal. To find the position of these lines, that we call segmentation boundaries, our approach considers projections [5] and maxima. To get maxima with high contrast, images I_g and I_t are thresholded by the scheme in [9], resulting in binary images BW_g and BW_t, respectively.

First, the horizontal projection of BW_t, denoted as $p^h_{BW_t}$, is obtained. Note that the highest maxima of this projection will be at the central position of the signal tracings. Therefore, boundaries must be located between two adjacent maxima and out of the range of the tracings.

As mentioned in Section 3, the grid cancelation may cause that signal pixels are in the background of BW_t, producing undesired gaps (minima) in $p^h_{BW_t}$. To prevent these minima from setting a false boundary, we propose to fill them by applying a closing to $p^h_{BW_t}$. Since the size of these minima is related to the width l_s of the grid lines, we consider that l_s is an adequate length for the flat SE. The l_s value is estimated from the vertical projection of
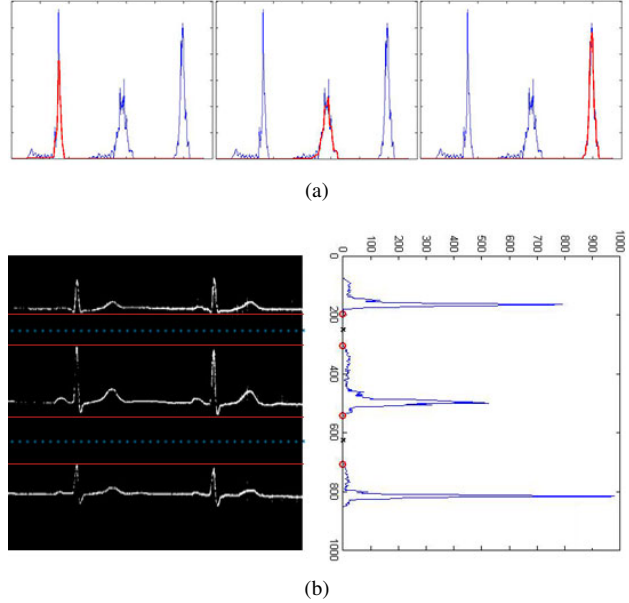

Figure 3.   (a) Projection and individual reconstructions (red), one per maximum considered as a valid candidate; (b) Projection, signal limits (red circles) and boundaries (black crosses) associated to the image in the left panel.

BW_g, named $p^v_{BW_g}$. The idea is to compute the area a_0 under $p^v_{BW_g}$ and monitor its changes when applying openings with increasing size of the SE. Thus, successive openings on $p^v_{BW_g}$ using SEs of increasing size are performed, and the corresponding area is calculated. When the area is for the first time lower than $0.1a_0$, then the corresponding SE length is considered to be an estimation for l_s.

Next, the closing of $p^h_{BW_t}$ with SE of length l_s is obtained in $p^{hc}_{BW_t}$, and maxima are searched to find the central position of the signal tracings. A morphological measure named contrast extinction value (CEV) was computed [10] for each maximum. We consider a maximum is a valid candidate to be the central point of a signal tracing when its CEV is greater than 10% of the highest CEV. The CEV of a maximum is the maximal size of a height filter such that the maximum still exists after the filtering.

Since boundaries should be placed between tracings of two consecutive signals, the range of rows $[r_{ini}, r_{end}]$ covered by every signal is next obtained. For this aim, we propose to perform n_l morphological reconstructions [11] of $p^h_{BW_t}$ using one marker per reconstruction. The marker is just one point at the position of each maximum considered as a valid candidate. This process is shown in Fig. 3 (a), where $n_l = 3$ reconstructions (red color) were obtained.

The $n_l - 1$ boundary points are obtained by computing, for every reconstruction i, the edges r^i_{ini} and r^i_{end} corresponding to the first and last value greater than zero. Next, all reconstructions from two consecutive maxima (labeled

| Scale | JPEG Quality | | | |
Factor	25%	40%	65%	100%
0.25	93.55	98.39	100.00	100.00
0.50	100.00	100.00	100.00	100.00
0.75	100.00	100.00	100.00	100.00
1.00	100.00	100.00	100.00	100.00

Table 1. Accuracy rate (in %) in the boundary position according to the scale factor and jpeg quality.

as i and $i+1$, $i = 1, \cdots, n_l - 1$) are considered, and boundary points are obtained as $(r_{end}^i + r_{ini}^{i+1})/2$. The projection in the right panel of Fig. 3 (b) shows both the edges of every reconstruction (circles) and the boundaries (crosses) associated to the image in the left panel of Fig. 3 (b).

5. Results

A set of thirty one 12-leads ECG printouts were used to evaluate the performance, both with solid and dotted grid lines. Grid lines are reddish, but the color is not the same for all printouts because the paper comes from different manufacturers. Images were taken using the auto mode with a mid-range smartphone, in indoor scenarios, with natural lighting and shadows. The phone was manually held and the paper was placed on a flat surface, providing images of 4208×3120 pixels in jpeg format. Average spatial resolution was of 1000 dpi. Signal tracings in the same printout do not overlap in the corresponding amplitude ranges, though they can be quite close. The definition of the signal tracing is not uniform, sometimes it is well defined and others it is just slightly perceptible, specially at high frequency deflections.

Table 1 presents the accuracy rate in the boundary position. We consider the boundary position is adequate when it is between two signal tracings, i.e. without fragmenting a signal. The influence of the spatial resolution (decimation of the original image up to one-fourth the original one) is evaluated through the scale factor (rows in Table 1). JPEG Quality is associated to the compression rate, with the right-hand column corresponding to the maximum quality (original image). For the same percentage of 'jpeg quatily', higher compression rates were achieved for images with dotted grid lines. This affected to the estimation of the boundary position, since errors occurred just in images with dotted grid lines, values of 'jpeg quatily' lower than 65% and the lowest spatial resolution (0.25 scale factor, equivalent to 250 dpi).

6. Conclusions

This work has proposed an automatic procedure to crop color ECG printout images in individual sub-images containing a unique lead in each sub-image. Unlike previous digitalization ECG procedures, these color EGC images were taken by using the camera of a mobile phone. Thirty one 12-leads ECG images were used in order to evaluate this procedure. These images were decimated to different spatial resolutions and compressed by modifying the 'jpeg quality' parameter. Though both parameters were important for a successful segmentation, results evidenced that images with solid grid lines were more robust when these parameters decreased their values. In fact, errors just occurred for images with dotted grid lines, the lowest spatial resolution (250 dpi) and values of the 'jpeg quality' parameter lower than 65%.

Acknowledgements

References

[1] Badilini F, Erdem T, Zareba W, Moss A. Ecgscan: a method for conversion of paper electrocardiographic printouts to digital electrocardiographic files. Journal of Electrocardiology 2005;38(4):310–318.

[2] Ravichandran L, Harless C, Shah A, Wick C, Mcclellan J, Tridandapani S. Novel tool for complete digitization of paper electrocardiography data. IEEE J of translational engineering in health and medicine 2013;1:1800107.

[3] Sanromán-Junquera M, Mora-Jiménez I, Caamaño A, Almendral J, Atienza F, Castilla L, García-Alberola A, Rojo-Álvarez J. Digital recovery of biomedical signals from binary images. Signal Processing 2012;92(1):43–53.

[4] Mitra R, Pramanik S, Mitra S, Chaudhuri B. A robust technique for delineation and features extraction of ecg signal from mobile-phone photography. In Intl Conf Communications, Devices and Intelligent Systems. 2012; 121–124.

[5] Gonzalez RC, Woods RE. Digital image processing. Prentice Hall, 2008.

[6] Hartley R, Zisserman A. Multiple view geometry in computer vision. Cambridge University Press, 2003.

[7] The opencv open source computer vision library. http://www.intel.com/technology/computing/opencv/index.htm. Accessed: 2016-08-22.

[8] Lillo-Castellano J, Mora-Jiménez I, Figuera-Pozuelo C, Rojo-Alvarez J. Traffic sign segmentation and classification using statistical learning methods. Neurocomputing 2015; 153:286–299.

[9] Sauvola J, Pietikäinen M. Adaptive document image binarization. Pattern Recognition 2000;33(2):225–236.

[10] Vachier C, Meyer F. Extinction value: a new measurement of persistence. In IEEE Worksh on Nonlinear Signal and Image Processing, volume 1. 1995; 254–257.

[11] Vincent L. Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. IEEE Trans on Image Processing 1993;2(2):176–201.