Automatic nipple detection in breast thermograms

Mohamed Abdel-Nasser*, Adel Saleh, Antonio Moreno, Domenec Puig

Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Av. Paisos Catalans 26,Tarragona 43007, Spain

ARTICLE INFO

ABSTRACT

Breast cancer is one of the most dangerous diseases for women. Detecting breast cancer in its early stage may lead to a reduction in mortality. Although the study of mammographies is the most common method to detect breast cancer, it is outperformed by the analysis of thermographies in dense tissue (breasts of young women). In the last two decades, several computer-aided diagnosis (CAD) systems for the early detection of breast cancer have been proposed. Breast cancer CAD systems consist of many steps, such as segmentation of the region of interest, feature extraction, classification and nipple detection. Indeed, the nipple is an important anatomical landmark in thermograms. The location of the nipple is invaluable in the analysis of medical images because it can be used in several applications, such as image registration and modality fusion. This paper proposes an unsupervised, automatic, accurate, simple and fast method to detect nipples in thermograms. The main stages of the proposed method are: human body segmentation, determination of nipple candidates using adaptive thresholding and detection of the nipples using a novel selection algorithm. Experiments have been carried out on a thermograms dataset to validate the proposed method, achieving accurate nipple detection results in real-time. We also show an application of the proposed method, breast cancer classification in dynamic images, where the new nipple detection technique is used to segment the region of the two breasts from the infrared image. A dataset of dynamic thermograms has been used to validate this application, achieving good results.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Current breast cancer studies show that detecting breast cancer in its early stage may reduce mortality (Siegel, Miller, & Jemal, 2015). The study of mammographies is the most common method for detecting breast cancer. Besides, the study of breast thermographies has been considered as an adjunct breast cancer screening procedure (Ng & Sudharsan, 2004). Thermograms are infrared images of the human body that can be used to detect breast cancer in its early stage. They can also reveal tumors in dense tissues, outperforming other modalities such as mammographies. However, the accuracy of thermographies is dependent on such factors as the symmetry of breast temperature and temperature stability.

Infrared imaging can be used for breast cancer screening because the metabolic activity and vascular circulation in precancerous tissue and in the area surrounding a developing breast cancer are always higher than in normal breast tissue. Mammographies, sonographies and magnetic resonance images try to find the physical tumor. In turn, thermograms detect the heat produced by increased blood vessel circulation and metabolic changes associated with a tumor's genesis and growth.

A number of computer-aided diagnosis (CAD) systems have been proposed for analyzing thermograms. A basic step in the use of thermal images is the design of CAD systems because they help technicians to execute medical exams using specified routines and protocols (similar to mammography exams). CAD systems give physicians good support in the analysis and interpretation of the outcomes of thermograms.

In the literature, several studies have focused on analyzing thermograms. Borchartt, Conci, Lima, Resmini, and Sanchez (2013) presented a review of the literature on image processing and techniques related to the analysis of thermographies for the detection and diagnosis of breast diseases. In Saniei, Setayeshi, Akbari, and Navid (2015), a computer method for quantifying the bilateral differences between left and right breasts was proposed. Ali et al. (2015) proposed a simple segmentation approach for extracting the region of interest (ROI) from breast thermograms. They extracted statistical and texture features from the segmented ROI and used a support vector machine classifier to discriminate between normal and abnormal breasts. Krawczyk, Schaefer, and Woźniak (2015) analyzed breast thermograms by extracting features describing
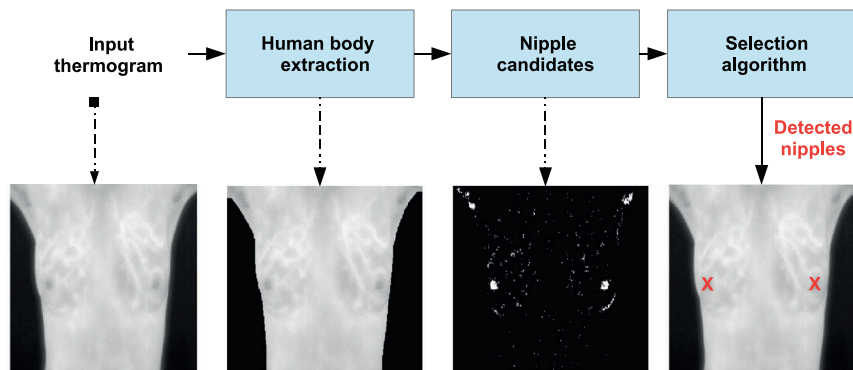

Fig. 1. Proposed system.

bilateral symmetries between the two breasts, and then classifying them into benign or malignant. Schaefer, Závišek, and Nakashima (2009) manually segmented the breast region from each thermogram and then extracted a set of statistical features from them. Finally, they used a fuzzy rule-based algorithm to discriminate between benign and malignant cases.

Motivations and contributions. The nipples are an important anatomical landmark in thermograms. The location of the nipples is invaluable in tasks such as image registration and modality fusion. Indeed, physicians use the location of nipples to find the corresponding masses in breast images that were acquired using different modalities (e.g., a thermogram and a mammogram). For instance, Saniei et al. (2015) manually picked the nipples to register the images of breasts. When analyzing a large number of thermograms, manual nipple detection may be time consuming; in addition, the precision of nipple detection varies with users' accuracy. Therefore, the design of an automatic nipple detection method may be very useful in the analysis of thermograms.

Several methods have been proposed for detecting nipples in mammograms (Chakraborty et al., 2015; Jas, Mukhopadhyay, Chakraborty, Sadhu, & Khandelwal, 2013) or in ultrasound images (Moghaddam, Tan, Karssemeijer, & Platel, 2014; Wang et al., 2014). However, few solutions have been proposed for detecting nipples in thermograms. Koay, Herry, and Frize (2004) proposed a method for classifying thermograms using an artificial neural network. They briefly mentioned an abridged rule-based method for detecting nipples. However, they did not explain any parameter settings or present an evaluation of the method because the main aim of their work was the classification of thermograms into normal, fibrocystic or cancerous. Unlike this study, in this paper we propose an unsupervised, automatic, accurate, simple and fast method for detecting nipples in thermograms using image processing operations and anatomical information. Moreover, we propose a novel selection algorithm that selects the nipples from a set of candidates.

Nipple detection is an important step in CAD systems because the location of the nipple can be used for several relevant purposes. For example:

- The basic steps of classifying a thermogram into normal or abnormal are the extraction of the ROI, feature extraction from the ROI, and using a machine learning method for classification. The ROI can be automatically extracted around the location of the nipple, determined using the proposed method.
- The distance from the location of the nipple can be used to find the corresponding masses in images of the same patient (mono- or multi- modality images).
- The location of the nipple can be used to register (align) the breast region of two thermograms of the same patient. The work in Saniei et al. (2015) manually detected the nipples to

register the breast regions in a sequence of dynamic thermograms. This process is time consuming and sensitive to the accuracy of the users. However, the proposed method can be used to automatically detect the nipples in real-time with a high accuracy.

Finally, we present an application of the proposed nipple detection method: breast cancer classification in dynamic thermographies, where the nipple detection method is used to segment the regions of interest (the two breasts) from the whole infrared image.

The rest of this paper is organized as follows. Section 2 explains the proposed method. Section 3 shows and discusses the results. Section 4 presents an application of the proposed method. Section 5 includes the conclusion and presents some lines of future work.

2. Proposed method

The stages of the proposed method are summarized in Fig. 1. As shown, there are three sequential stages: human body segmentation, determination of nipple candidates using adaptive thresholding and detection of the nipples using a selection algorithm.

2.1. Human body segmentation

Thresholding is used to segment the human body from a given image by setting all pixels whose intensity values are above a threshold to a foreground value (one) and all the remaining pixels to a background value (zero). A threshold of 50 is used in this study to generate the human body mask (binary image), which is then smoothed by morphological operations. We applied morphological closing to the binary mask using a disk-shaped structuring element with a radius of 3 to fill the gaps. We then used a morphological dilation with a disk-shaped structuring element with a radius of 10. Fig. 2(a) shows an example of an input thermogram and how the generated human body binary mask (Fig. 2(b)) is used to segment the human body image (Fig. 2(c)).

2.2. Determination of nipple candidates using an adaptive threshold

To determine nipple candidates, we process the image that contains the human body by means of *an adaptive thresholding* algorithm. Adaptive thresholding typically takes an image as input and produces a segmented image and the labels of the segmented regions. It is a method that can separate the foreground from the background with non-uniform illumination (He & Yung, 2004).

For each pixel in the image, a threshold has to be calculated. If the pixel value is below the threshold it is set to the background value; otherwise, it is set to the foreground value. To find
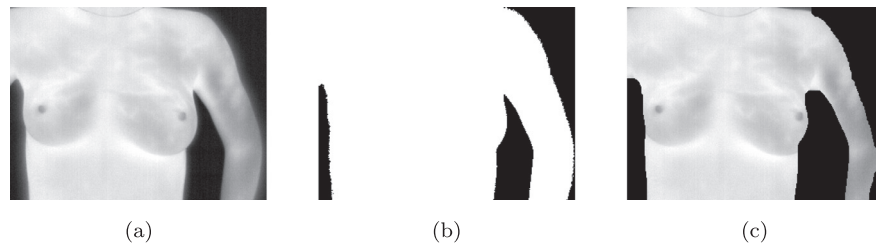
(a)   (b)   (c)

Fig. 2. Human body segmentation. (a) the input thermogram image, (b) the human body mask, and (c) the segmented image.

Algorithm 1 Adaptive thresholding algorithm.

1: Input: a gray scale image
2: Output: a binary image
3: **procedure** ADAPTIVE THRESHOLDING
4:     Convolution of the image with a suitable statistical operator (the median or mean).
5:     Subtraction of the original image from the convolved one.
6:     Thresholding the difference image with a constant *C*.
7:     Invert the thresholded image.
8: **end procedure**

---

the local threshold, the algorithm statistically examines the intensity values of the local neighborhood of each pixel. Functions such as mean or median of the local intensity distribution can be used. In Algorithm 1, we present the steps of the adaptive thresholding.

In this study we used the median operator with a local neighborhood of 15 pixels and *C* was 0.03. Fig. 3 shows nipple candidates extracted from the human body image using the adaptive thresholding algorithm. As shown, most candidates are a long way from the real position of the nipples. Moreover, some regions correspond to external objects, like necklaces, bracelets or wristwatches (e.g., in Fig. 3 we can see a necklace on the upper part of the image). Thus, we need a selection algorithm to pick the correct location of nipples. In the subsection below we explain the proposed nipple selection algorithm.

2.3. Nipple detection

To select the correct position of the nipples from a set of candidates we applied several selection rules. We used our knowledge of the features of thermograms to design these rules. Breast nipples have some unique anatomical and visual features that can be used to design the selection rules. We used the following facts:

- Nipples lie inside the outer boundary of the human body.
- The human body has only two nipples, which occupy small regions.
- Nipples do not lie in the lowest or uppermost parts of thermograms.
- A thermogram has only two nipples, one on the left and another on the right.
- A breast nipple is approximately circular.

Exploiting the first fact. The human body boundary is extracted by defining the outermost edges in the human body mask. To do so, the rows and columns are scanned twice. Each row of the mask is scanned from the left to determine the right most edge points, and then they are scanned from the right to determine the left most edge points. Each column of the mask is scanned from the top to determine the bottom most edge points and then scanned from the bottom to determine the top most edge points.

Exploiting the second fact. We count the number of pixels in the region of each candidate. If the number of pixels in a given
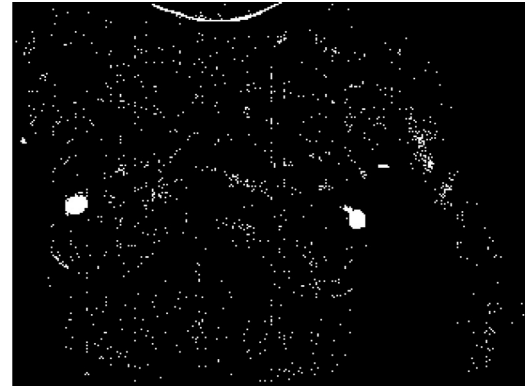

Fig. 3. Determining nipple candidates using adaptive thresholding.

region is less than a predefined threshold N_p, we delete it from the nipple candidates list. In this study N_p was set to 20.

Exploiting the third fact. We generate a binary mask ($mask_{UL}$) to exclude the regions that lie in the uppermost and lowermost parts of the image. Fig. 4 shows the upper region ($region_{up}$) and lower region ($region_{lw}$). Candidate regions that lie in or touch the $region_{up}$ or $region_{lw}$ are excluded from the nipple candidates list. The regions that correspond to $region_{up}$ and $region_{lw}$ in $mask_{UL}$ are set to zero; the rest of the image is set to one. Assume that H_{im}, H_{up} and H_{lw} are the height of the input thermogram, $region_{up}$ and $region_{lw}$, respectively. H_{up} and H_{lw} have been defined as follows.

- H_{up} was set to $0.35 * H_{im}$.
- H_{lw} was set to $0.3 * H_{im}$.

A similar setting was used in Ali et al. (2015) to segment the breast region.

Exploiting the fourth fact. To determine the left and right regions of the input thermogram, we determine the center-line (L_{cnt}) as shown in Fig. 4. L_{cnt} divides the input image into two equal regions: one region includes the left breast and the other includes the right breast.

Exploiting the fifth fact. As nipples are approximately circular, we also use this feature to select the correct nipples from the candidates list (N). In this study we use the *roundness* measure to identify the candidates that are round in shape. To measure the roundness (R) of a given region we used the following metric (Crocker, Jones, & Curran, 1983):

$$R = 4\pi . A/P^2 \tag{1}$$

In this equation A and P are the area and the perimeter of the region of a given candidate, respectively. As the value of R approaches 0, it indicates an oblong (non-circular) object. A value of 1.0 indicates a perfect circle. The area of a region is measured simply by counting the number of pixels. Given the locations of the points of the boundary of a certain region $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, we estimate the region's perimeter
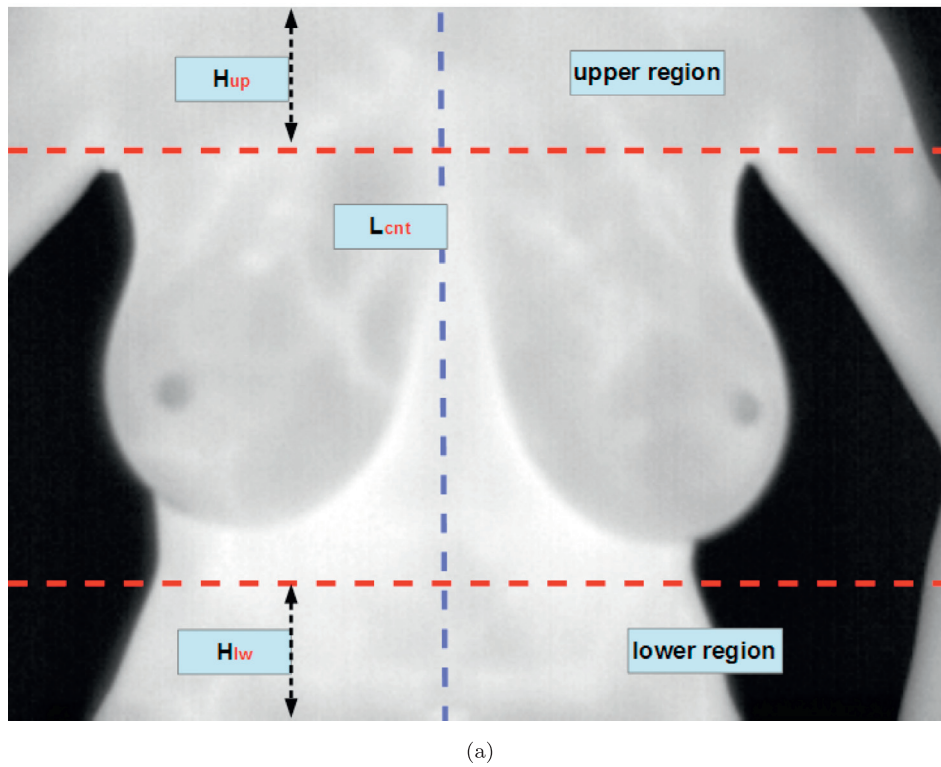
Fig. 4. Determining the upper region ($region_{up}$), the lower region ($region_{lw}$) and the center-line (L_{cnt}).
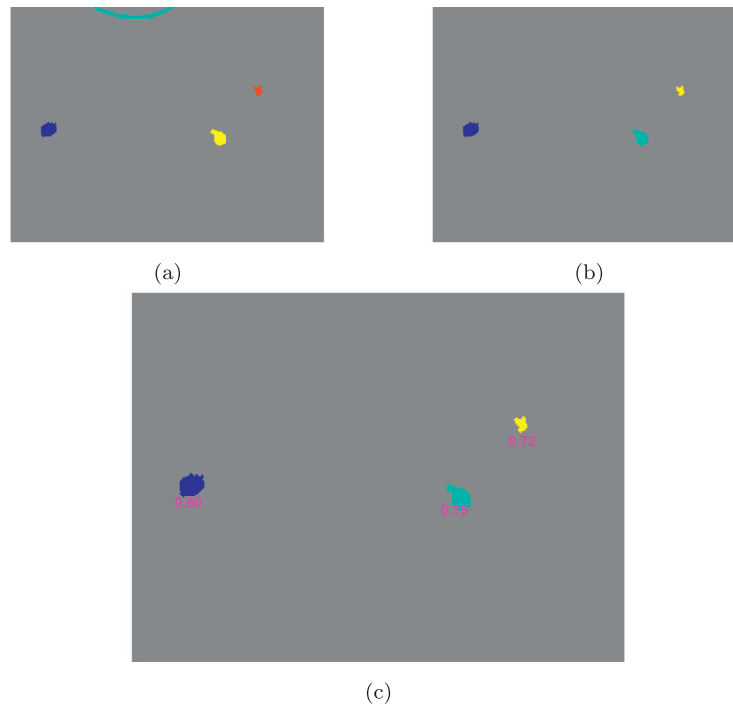

Fig. 5. Applying the selection rules to detect the nipples. (a) regions having a number of pixels greater than N_p, (b) the regions that lay outside $region_{up}$ and $region_{lw}$, and (c) the roundness of the selected regions.

with the expression

$$P = \sum_{i=1}^{n} \sqrt{((x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2)} \qquad (2)$$

in which n is the number of points of the boundary.

After applying the above rules, we can detect the regions of the nipples. Fig. 5(a) shows the regions that have a number of pixels greater than N_p and that lie inside the boundary of the body (in this study N_p was set to 20). Fig. 5(b) shows the regions that do not lie in $region_{up}$ or $region_{lw}$ of the thermograms. This step excludes the regions that lie away from the breasts. Fig. 5(c) shows the roundness measure of the regions selected in Fig. 5(b). As shown, the three regions have a roundness measure of 0.72, 0.75 and 0.80. The regions that have the two highest values correspond to the nipples.

Table 1
Notation used in Algorithm 2.

Notation	Meaning
N	List of nipple candidates.
HBB	Human body boundary.
A	The area of a given region.
R	The roundness of a given region.
L_{cnt}	The center-line of the image.
$region_{up}$	Indices of the upper region.
$region_{lw}$	Indices of the lower region.
L_{left}	List of left nipple's candidates.
L_{right}	List of right nipple's candidates.
max	A function that finds the index of the region that has the maximum roundness or biggest area.
m_l	Index of the region that has the maximum roundness in L_{left}.
m_r	Index of the region that has the maximum roundness in L_{right}.
m'_l	Index of the region that has the biggest area in N_{left}.
m'_r	Index of the region that has the biggest area in N_{right}.
$nipple_{left}$	The left nipple.
$nipple_{right}$	The right nipple.

In Algorithm 2, we present the steps of nipple selection from the given candidates. The algorithm receives the list of nipple candidates and outputs the location of the left and right nipples. To understand the proposed nipple selection algorithm, we summarize and explain the notations used in Algorithm 2 in Table 1. The main steps of the algorithm are the following:

- **Steps 4–8.** The algorithm suppresses the regions that lie outside the human body boundary (HBB) and that touch or lie $region_{up}$ or $region_{lw}$ from the list of candidates (N).

Algorithm 2 Nipple selection algorithm.

1: Inputs: The list of nipple candidates N, HBB and the center-line (L_{cnt})
2: Outputs: Locations of the two nipples ($nipple_{left}$ and $nipple_{right}$)
3: **procedure** NIPPLE SELECTION
4:     **for** each candidate $i \in N$ **do**
5:         Delete i, if any element of $i \notin HBB$
6:         Delete i, if any element of $i \in region_{up}$
7:         Delete i, if any element of $i \in region_{lw}$
8:     **end for**
9:     Find left nipple candidates N_{left} that lay at the left of L_{cnt}
10:     **if** $length(L_{left}) > 1$ **then**
11:         $m_l = max\{R(L_{left})\}$
12:         **if** $length(m_l) > 1$ **then**
13:             $m'_l = max\{A(L_{left}(m_l))\}$
14:             $nipple_{left} = L_{left}(m'_l)$
15:         **else** $nipple_{left} = L_{left}(m_l)$
16:         **end if**
17:     **else** $nipple_{left} = L_{left}$
18:     **end if**
19:     Find right nipple candidates (L_{right}) that lay at the right of (L_{cnt})
20:     **if** $length(L_{right}) > 1$ **then**
21:         $m_r = max\{R(L_{right})\}$
22:         **if** $length(m_r) > 1$ **then**
23:             $m'_r = max\{A(L_{right}(m_r))\}$
24:             $nipple_{right} = L_{right}(m'_r)$
25:         **else** $nipple_{right} = L_{right}(m_r)$
26:         **end if**
27:     **else** $nipple_{right} = L_{right}$
28:     **end if**
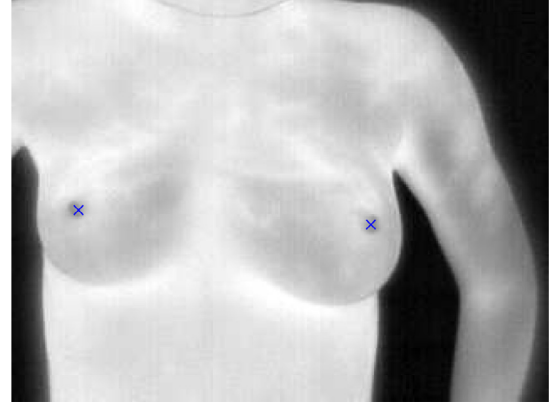29: **end procedure**

Fig. 6. Detected nipples.

- **Steps 9–18.** The algorithm identifies a list containing the left nipple candidates (L_{left}) and then finds the region that has the maximum roundness. However, if two regions have the same maximum roundness, it selects the region that has the biggest area. Note that, if L_{left} includes one region, the algorithm directly identifies it as a left nipple.
- **Steps 19–28.** The algorithm repeats steps 9–18 to detect the right nipple.

Finally, we calculate the centroid of each region and plot it on the input thermogram to highlight the location of the detected nipple. Fig. 6 shows the final position of the nipples in the input thermogram.

3. Experimental results and discussion

The proposed nipple detection method has been tested using a breast thermography dataset that consists of 148 thermograms (for 148 healthy and sick women). We collected the dataset from the Proeng database that is available at http://visual.ic.uff.br/en/proeng/. In a thermogram each pixel corresponds to the temperature of the acquired scene. The range of temperatures is associated with the range of gray values of the images. The size of the images is 640×480 pixels. The thermograms were captured by a FLIR thermal camera, model $SC620$, which has a sensitivity of less than $0.04°$ range and capture temperatures from $-40°$ C to $500°$ C (Silva et al., 2014).

3.1. Evaluation

We used a training set of 8 images to optimize the threshold value, the upper and lower limits (H_{up} and H_{lw}), and N_p. We then used the optimized values to test the proposed method using the rest of the images. We used the Recall, Precision and F-score to evaluate the proposed method. If the centroid of the selected region is inside the region of the true nipple, then it is a true positive (TP). On the other hand, if the centroid of the selected region is outside the region of the true nipple, it is a false positive (FP). To calculate the recall, precision and F-score, we determine the following identities:

- TP = correctly identified nipples.
- FP = incorrectly identified nipples.
- P_{pos} = the number of positive instances (number of nipples).

Then, the Recall, Precision and F-score can be calculated as follows:
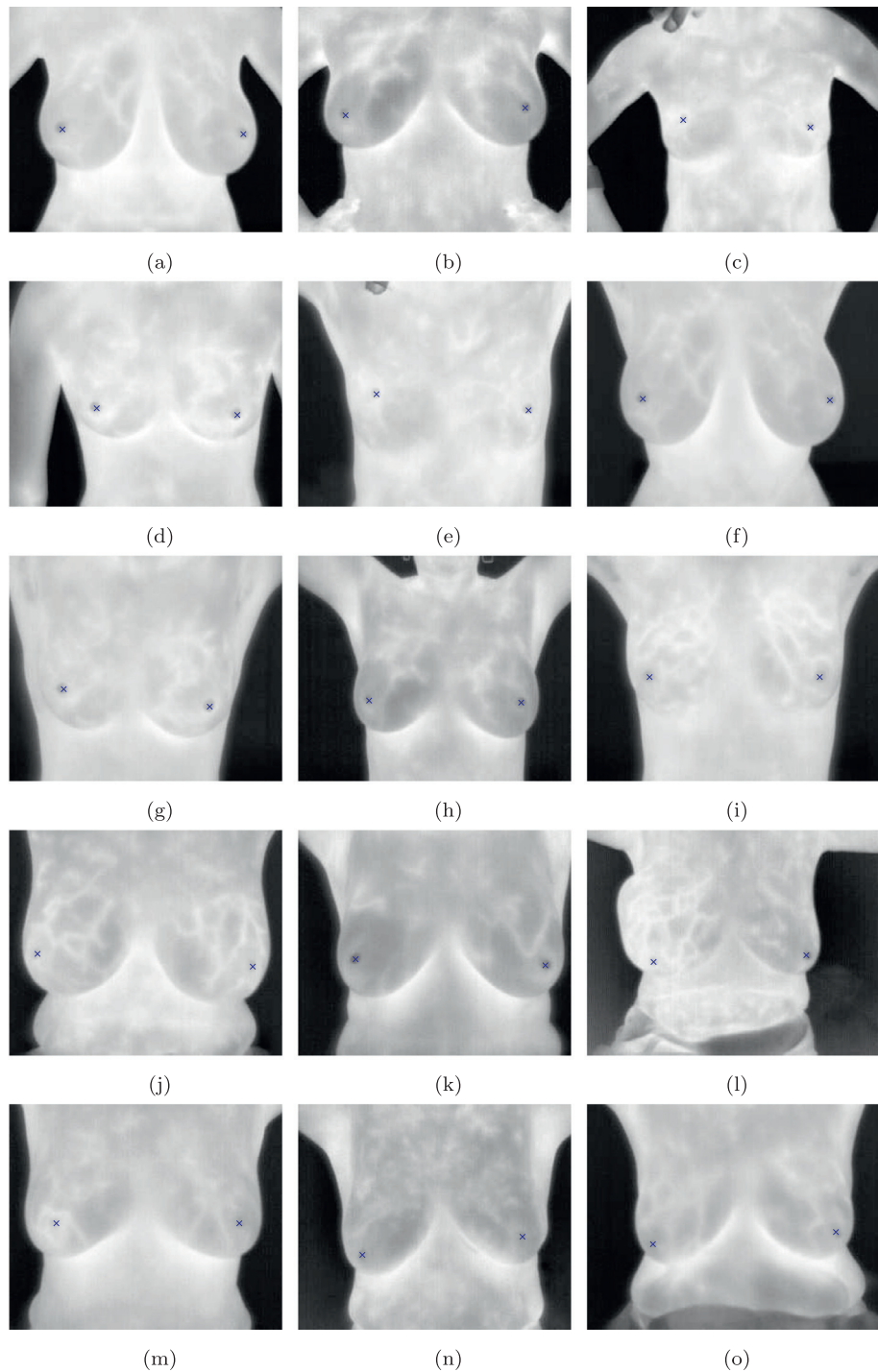
$$Recall = TP/P_{pos} \tag{3}$$

Fig. 7. Examples of correctly detected nipples using the proposed method.

$$Precision = TP/(TP + FP) \qquad (4)$$

$$\text{F-score} = \frac{2.precision.recall}{(precision + recall)} \qquad (5)$$

3.2. Results

Fig. 7 shows examples of correctly detected nipples on thermograms for different women. We show the results of the proposed method with small breasts (Fig. 7(c,d,g)), medium breasts (Fig. 7(e,h,i)) and big breasts (Fig. 7(a,b,f,j–o)). As shown, in all cases the proposed method correctly determines the location of the nipples.

We evaluated the proposed method with the whole dataset. Note that the thermograms used contain at least one nipple in their profile. Indeed, we have excluded the thermograms that do not include any nipples in their profiles. However, the dataset still contains very difficult cases, where nipple detection is difficult even for humans. Table 2 shows the evaluation of the proposed method with the whole dataset. The nipple detection results were outstanding (precision close to 0.99). Note that the dataset used
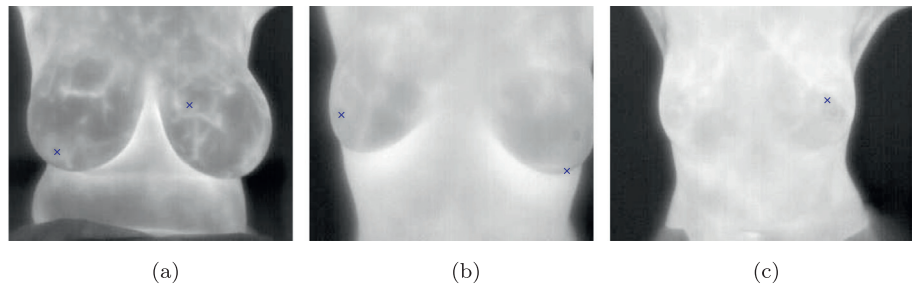
(a)　　　　　　　　　(b)　　　　　　　　　(c)

Fig. 8. Examples of false positives.

Table 2
Evaluation of the performance of the proposed method.

Method	Recall	Precision	F-score	Time (s)
Proposed	**0.930**	**0.989**	**0.958**	**0.30**
(Koay et al., 2004)	0.500	0.652	0.566	0.91

includes several cases in which it is difficult to identify the nipples even with the human eye.

Fig. 8 shows examples of false positive nipples. As shown, the algorithm fails to detect some nipples in these cases. Indeed, the regions of undetected nipples are not clear in the images, so detecting the nipples with the human eye is also difficult. Consequently, the nipples can not be detected using the proposed method in these cases.

3.3. The effect of hyper-parameters on the accuracy of the proposed method

The proposed method has a few hyper-parameters to be tuned: the threshold value, the upper and lower limits (H_{up} and H_{lw}), and N_p. However, these parameters do not affect the accuracy of the proposed method. The threshold value is used to segment the human body from the background, which is a very simple operation because we segment the human body region from a black background. Even if there is a segmentation error due to the threshold value or H_{up} and H_{lw}, it will not affect the accuracy of the proposed method because this error is happening at the boundary of the human body, which is far from the location of the nipples (at the middle of the human body). As mentioned above, a training set of 8 images was used to find the optimal values for these hyper-parameters. We then used these values to test the proposed method using the rest of the images.

3.4. Comparison with related work

Indeed, very few methods have been proposed to detect nipples in thermograms. In this subsection we compare our method with the approach suggested in Koay et al. (2004), in which the authors used an adaptive thresholding and morphological operations to detect the nipples. They also removed big contour regions and regions in the upper breasts, and assumed that the nipples have an eccentricity close to one. Indeed, they did not present any parameter settings and made no evaluation of nipple detection because the main goal of their work was to classify thermograms into normal, fibrocystic or cancerous. We implemented the aforementioned rules and tuned the values of the parameters to obtain the best performance.

As can be seen in Table 2, the proposed method easily outperforms the method of Koay et al. (2004) in terms of recall, precision and F-score. Fig. 9(a-b) shows the performance of the two methods in a thermogram containing two nipples. The proposed method

easily detects the two nipples in the cases presented (Fig. 9(a)) while (Koay et al., 2004) does not (Fig. 9(b)). Fig. 9(c-d) shows the performance of the two methods in a thermogram containing one clear nipple (right nipple) in its profile (the left nipple is fuzzy). The proposed method successfully detects the right nipple (Fig. 9(c)) while (Koay et al., 2004) does not (Fig. 9(d)). These results demonstrate that the rules used in Koay et al. (2004) produce poor detection results.

The two methods were implemented using MATLAB on an Intel processor core 2 Quad at 2.5GHz and 8 GB of RAM. As presented in Table 2, the proposed method takes 0.3s to determine the nipples. This time is approximately one-third of the execution time of Koay et al. (2004), indicating that the proposed method is more suitable for working with real-time applications.

3.5. Limitations

Like any computer approach, the proposed method has some limitations:

- It does not work in thermograms in which nipples do not appear or are not clearly visible.
- It fails to detect nipples in thermograms that contain fuzzy regions, and in which even the human eye would have trouble detecting them.

3.6. Possible applications of the proposed method

As we have shown above, our method accurately detects the nipples in real-time (0.3s). Thus, it could be used in such applications as the following:

- It can be used to find two seed points to segment the breast region in thermograms. Thus, we can integrate our method with the cancer detection approach suggested by Schaefer et al. (2009) where the breasts were manually segmented. We suggest replacing this step by an automatic segmentation method based on the method suggested in this paper. This makes the approach proposed in Schaefer et al. (2009) fully automatic.
- Given two thermograms, the proposed method can be used to register one breast to the coordinates of another. Thus, it can be used to improve the registration step used in Saniei et al. (2015).
- It can be used to find the correspondence location of masses in images that were acquired using different modalities (e.g., thermograms and mammograms) for the same breast.
- It will be used to design a robust thermogram registration framework. In Abdel-Nasser, Moreno, and Puig (2016) we used an optimization approach to find the optimal location of the origin of the curvilinear coordinates to register mammograms; in turn, the nipples detected by our method will be used as origins to establish the curvilinear coordinates in order to align thermograms.
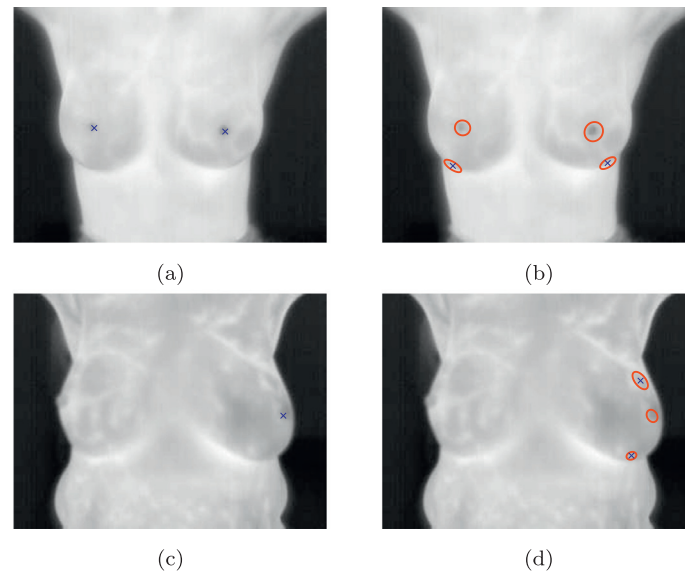
Fig. 9. Comparing the performance of proposed method (a,c) and the approach proposed in Koay et al. (2004) (b,d).
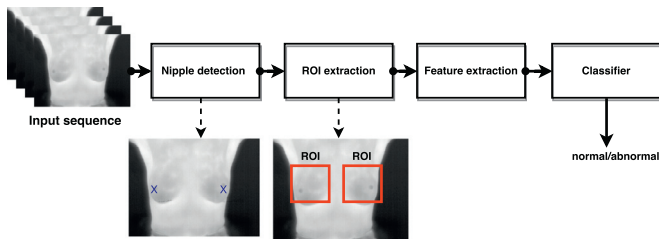

Fig. 10. An application: breast cancer detection in dynamic thermograms.

In the section below, we show the use of the nipple detection algorithm in a breast cancer detection method (classification of the images into normal or abnormal).

4. Application: breast cancer detection in dynamic thermographies

Dynamic infrared imaging may improve the performance of the conventional static infrared thermographies when an external stimulus is applied to enhance the thermal contrast. Dynamic infrared imaging produces a sequence of temperature matrices, which can be converted into images. Each matrix is acquired at time t and contains the temperatures of the middle area of a woman's body (from the neck to the waist).

The dataset of dynamic thermographies of Silva et al. (2014) has been used in this application. The main step of the acquisition protocol is the cooling of the regions of the breasts and armpits of the patient using an electric fan for several minutes in an environment with controlled temperature (20° C to 22° C). After the cooling process, 20 images were captured during 5 minutes using a FLIR thermal camera (model SC620). The camera had a sensitivity smaller than 0.04° C, the detectable temperature range was between −40° C and 500° C and the images have a dimension of 640 × 480 pixels. The dataset includes 37 sequences of cancer patients (histopathologically proven cancer) and 19 sequences of healthy cases (each sequence contains 20 images, in total, 1120 images).

Fig. 10 shows the steps of the proposed application: nipple detection, ROI extraction, feature extraction and classification. In this application, given a sequence of dynamic images for a given pa-
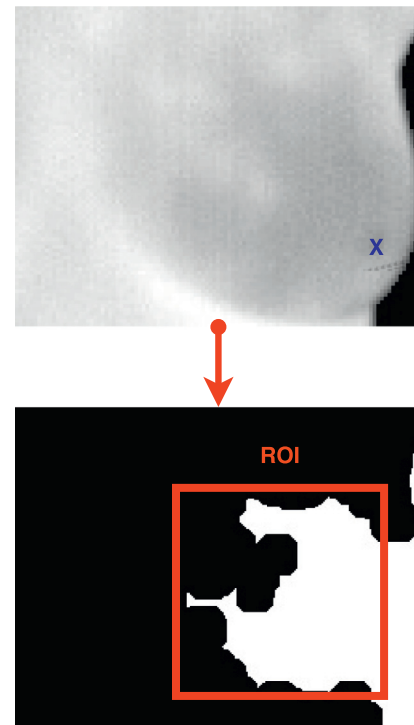

Fig. 11. The extracted region of interest.

tient, we extract a region of interest around the location of the detected nipples from each image. To extract the regions of interest, a region-based segmentation method can be used to roughly segment the breast area, and then we determine the rectangular ROI that contains the segmented area. We used Chan-Vese active contours for segmentation (Chan, Sandberg, & Vese, 2000; Chan & Vese, 2001). In general, active contour methods require a mask in order to initialize the segmentation process. In several studies the location of this mask is set manually. In contrast, we initialize a mask ($n1 \times n1$ neighborhood) around the detected nipple, and then the active contour method uses this mask to segment the breast region. Fig. 11 shows an example of ROI segmentation. The white region to the right of the ROI lies outside the human
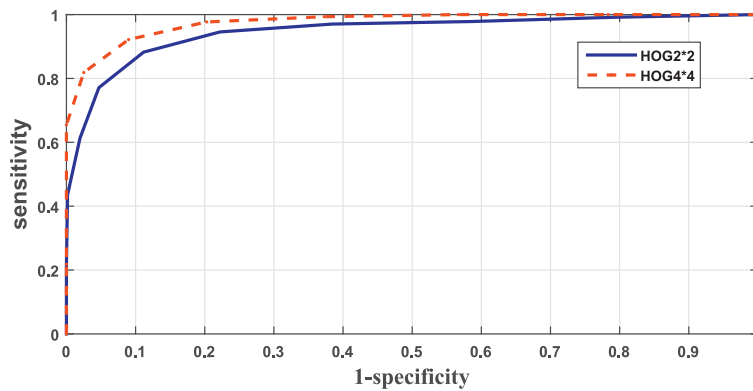
Fig. 12. The ROC curve of the proposed breast cancer detection method with HOG2*2 and HOG4*4.

Table 3
Breast cancer detection method based on the nipple detection algorithm.

Method	AUC
HOG2*2	0.974
HOG4*4	**0.980**

5. Conclusion

An unsupervised, automatic, accurate, simple and real-time method for detecting nipples in thermograms is proposed in this paper. The main stages of the proposed method are: human body segmentation, determination of nipple candidates using adaptive thresholding and finally detection of the nipples using rules derived from the anatomical structure of the human body. The proposed method takes 0.3s to detect the nipples. It gives very accurate results in real-time. We also presented an application of the nipple detection algorithm (breast cancer classification in dynamic images), in which good results were obtained.

Our future work will focus on integrating the proposed method with several applications, such as breast region segmentation and finding the correspondences between the masses in mammograms and thermograms of the same patients. In addition, we will use the proposed nipple detection method to find the optimal location of the origin of the curvilinear coordinates, which will be used to register thermograms.

body boundary, therefore, we excluded it. Indeed, the location of the nipple is a reasonable and accurate position for setting the mask.

From each ROI, we extract a histogram of oriented gradients (HOG), and then concatenate the HOG descriptors extracted from all ROIs in the sequence, thus generating one feature vector describing the whole sequence. Finally, we input the feature vectors into a multi-layer perceptron (MLP) to build a model that discriminates a normal ROI from an abnormal one.

HOG is a robust texture analysis method because it produces distinctive features in the case of illumination changes (Dalal & Triggs, 2005). In the HOG method, the occurrences of edge orientations in a local image window are counted. The image is divided into blocks (small groups of cells) and then a weighted histogram is computed for each of them. The frequencies in the histograms are normalized in the interval [0,1] to compensate changes of illumination. The final HOG descriptor is formed by concatenating the histograms of all blocks.

The performance of the proposed method is measured in terms of the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. Table 3 presents the results of the proposed cancer detection method with the MLP classifier (50 neurons) in terms of AUC with HOG2*2 (2 blocks, 2 cells) and HOG4*4 (4 blocks, 4 cells). Fig. 12 shows the ROC curves of the proposed breast cancer detection method with HOG2*2 and HOG4*4. As shown, HOG4*4 gives an AUC value and accuracy better than the ones of HOG2*2. The main reasons of the good performance of HOG are:

- HOG descriptor is less sensitive to the noise.
- HOG accumulates the occurrence of the gradients of the input data into different orientations, and therefore it provides a good description of the gradients of temperatures in the input images.

Despite the fact that HOG gave good results, other feature extraction methods could have been used. However, determining the best feature extraction method for this problem is out of the scope of this paper.

References

Abdel-Nasser, M., Moreno, A., & Puig, D. (2016). Temporal mammogram image registration using optimized curvilinear coordinates. *Computer Methods and Programs in Biomedicine, 127*, 1–14.
Ali, M. A., Sayed, G. I., Gaber, T., Hassanien, A. E., Snasel, V., & Silva, L. F. (2015). Detection of breast abnormalities of thermograms based on a new segmentation method. In *Federated conference on computer science and information systems (fedcsis)* (pp. 255–261). IEEE.
Borchartt, T. B., Conci, A., Lima, R. C., Resmini, R., & Sanchez, A. (2013). Breast thermography from an image processing viewpoint: A survey. *Signal Processing, 93*(10), 2785–2803.
Chakraborty, J., Midya, A., Mukhopadhyay, S., Rangayyan, R. M., Sadhu, A., Singla, V., … Azevedo-Marques, P. M., et al. (2015). Detection of the nipple in mammograms with Gabor filters and the Radon transform. *Biomedical Signal Processing and Control, 15*, 80–89.
Chan, T. F., Sandberg, B. Y., & Vese, L. A. (2000). Active contours without edges for vector-valued images. *Journal of Visual Communication and Image Representation, 11*(2), 130–141.
Chan, T. F., & Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing, 10*(2), 266–277.
Crocker, J., Jones, E., & Curran, R. (1983). A comparative study of nuclear form factor, area and diameter in non-Hodgkin's lymphomas and reactive lymph nodes. *Journal of Clinical Pathology, 36*(3), 298–302.
Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In *IEEE computer society conference on computer vision and pattern recognition: 1* (pp. 886–893). IEEE.
He, X.-C., & Yung, N. H. (2004). Curvature scale space corner detector with adaptive threshold and dynamic region of support. In *Proceedings of the 17th international conference on pattern recognition: 2* (pp. 791–794). IEEE.

Jas, M., Mukhopadhyay, S., Chakraborty, J., Sadhu, A., & Khandelwal, N. (2013). A heuristic approach to automated nipple detection in digital mammograms. *Journal of Digital Imaging, 26*(5), 932–940.

Koay, J., Herry, C., & Frize, M. (2004). Analysis of breast thermography with an artificial neural network. In *26th annual international conference of the ieee engineering in medicine and biology society: 1* (pp. 1159–1162). IEEE.

Krawczyk, B., Schaefer, G., & Woźniak, M. (2015). A hybrid cost-sensitive ensemble for imbalanced breast thermogram classification. *Artificial Intelligence in Medicine, 65*(3), 219–227.

Moghaddam, M. J., Tan, T., Karssemeijer, N., & Platel, B. (2014). Automatic nipple detection on 3d images of an automated breast ultrasound system (ABUS). *Spie medical imaging*. International Society for Optics and Photonics. 903405–903405

Ng, E. Y., & Sudharsan, N. (2004). Computer simulation in conjunction with medical thermography as an adjunct tool for early detection of breast cancer. *BMC cancer, 4*(1), 17.

Saniei, E., Setayeshi, S., Akbari, M. E., & Navid, M. (2015). A vascular network matching in dynamic thermography for breast cancer detection. *Quantitative InfraRed Thermography Journal*, (ahead-of-print), 1–13.

Schaefer, G., Závišek, M., & Nakashima, T. (2009). Thermography based breast cancer analysis using statistical features and fuzzy classification. *Pattern Recognition, 42*(6), 1133–1137.

Siegel, R. L., Miller, K. D., & Jemal, A. (2015). Cancer statistics, 2015. *CA: A Cancer Journal for Clinicians, 65*(1), 5–29.

Silva, L., Saade, D., Sequeiros, G., Silva, A., Paiva, A., Bravo, R., & Conci, A. (2014). A new database for breast research with infrared image. *Journal of Medical Imaging and Health Informatics, 4*(1), 92–100.

Wang, L., Böhler, T., Zöhrer, F., Georgii, J., Rauh, C., Fasching, P. A., … Uder, M., et al. (2014). Fully automated nipple detection in 3d breast ultrasound images. In *Breast imaging* (pp. 64–71). Springer.