Analysis and Retrieval of Mammographic Images

Stylianos Tzikopoulos*

National and Kapodistrian University of Athens
Department of Informatics and Telecommunications
stzikop@di.uoa.gr

Abstract. In this thesis two computer-aided diagnosis (CAD) systems are presented and implemented and their performance is evaluated. The first system proposed is a fully automated segmentation and classification scheme for mammograms based on breast density estimation and detection of asymmetry. First, image preprocessing and segmentation techniques are applied. Then, features for breast density categorization are extracted and Support Vector Machines (SVMs) are employed for classification, achieving accuracy of up to 85.7%. Most of these properties are used to extract a new set of statistical features for each breast, that are used to detect breast asymmetry between a pair of mammograms. The classifier adopted is an one-class SVM classifier, which resulted in a success rate of 84.47%. This composite methodology has been applied to the miniMIAS database. The results were evaluated by expert radiologists, are very promising and compared to other related works. The second system proposed is an experimental "morphological analysis" retrieval system for mammograms, using Relevance-Feedback techniques. The features adopted are first-order statistics of the Normalized Radial Distance, extracted from the annotated mass boundary. The system is evaluated on an extensive dataset of 2274 masses of the DDSM database, which involves 7 distinct classes. The experiments verify that the involvement of the radiologist as part of the retrieval process improves the results, reaching the precision rate of almost 90%. Therefore, Relevance-Feedback can be employed as a very useful complementary tool to a Computer Aided Diagnosis system.

Subject area: Digital Signal Processing, Medical Image Analysis, Pattern Recognition
Keywords: computer-aided diagnosis (CAD), content-based image retrieval (CBIR), relevance feedback (RF), mammography, classification

1 Introduction

Breast cancer, i.e., a malignant tumor developed from breast cells, is considered to be one of the major causes for the increase in mortality among women,

* Dissertation Advisor: Sergios Theodoridis, Professor

especially in developed countries. More specifically, breast cancer is the second most common type of cancer and the fifth most common cause of cancer death according to [1].

While mammography has been proven to be the most effective and reliable method for the early detection of breast cancer, as indicated by [2], the large number of mammograms, generated by population screening, must be interpreted and diagnosed by a relatively small number of radiologists. In addition, when observing a mammographic image, abnormalities are often embedded in and camouflaged by varying densities of breast tissue structures, resulting in high rates of missed breast cancer cases as mentioned by [3]. In order to reduce the increasing workload and improve the accuracy of interpreting mammograms, a variety of Computer-Aided Diagnosis (CAD) systems, that perform computerized mammographic analysis have been proposed, as stated by [4]. These systems are usually employed as a second reader, with the final decision regarding the presence of a cancer left to the radiologist. Thus, their role in modern medical practice is considered to be significant and important in the early detection of breast cancer.

The new CAD systems try to extract information not only from the external annotation given by the radiologist, but also from the images themselves. This way, they can provide to the user similar cases, by comparing the query image with the images of the available database. This methodology use also the content-based image retrieval (CBIR systems). Over the recent years, these systems are gaining in importance [5, 6]. Such systems extract visual features from the "query" image, e.g. color, texture or shape and perform a comparison of it with the available images in a database, using specific similarity measures. The most similar images are returned to the user.

The scenario described above uses low-level features, which are not capable of capturing the image semantics, e.g. the high-level semantic concept that is meaningful for a user. This is known as the semantic gap. In order to address this gap, Relevance Feedback techniques (RF) have been developed since the early and mid-1990's [7]. In such a system, the user interacts with the search engine and marks the images that he perceives as relevant or non-relevant. Taking into account this feedback information, the engine "learns" and improved results are returned to the user during the next iteration.

Besides image retrieval, RF can also be employed to other systems, such as the retrieval of text documents, music or 3D objects. More recently it was used for medical image retrieval [8, 9]. In such a context, the aim of a retrieval system is to function in conjunction with a Computer Aided Diagnosis (CAD) system. The radiologists can be provided with relevant past cases -according to the query-, along with proven pathology and other information, making the diagnosis more reliable. RF seems an ideal scheme for the improvement of the performance of medical image retrieval systems, as it incorporates the radiologist's judgement, in order to capture the some higher-level semantic concepts of the medical images. The judgement of such an expert is the result of a very complex and vague procedure, combining a multitude of quantitative and qualitative facts, as well as the radiologist's experience, and therefore should be taken into consideration.

2 A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry

2.1 Overview of work

All of the CAD systems require, as a first stage, the segmentation of each mammogram into its representative anatomical regions, i.e., the breast border, the pectoral muscle and the nipple. The breast border extraction is a necessary and cumbersome step for typical CAD systems, as it must identify the breast region independently of the digitization system, the orientation of the breast in the image and the presence of noise. The pectoral muscle is a high-intensity, approximately triangular region across the upper posterior margin of the image, appearing in all the medio-lateral oblique (MLO) view mammograms. Automatic segmentation of the pectoral muscle can be useful in many ways [10]. One example is the reduction of the false positives in a mass detection procedure. In addition, the pectoral muscle must be excluded in an automated breast tissue density quantification method. The location of the nipple is also of great importance, as it is the only anatomical landmark of the breast, as mentioned by [11]. Most CAD systems use the nipple as a registration point for comparison, when trying to detect possible asymmetry between the two breasts of a patient. These automatic methods can also use the nipple as a starting point for cancer detection. Moreover, radiologists pay specific attention to the nipple, when examining a mammogram, according to [12].

Another important characteristic of a mammogram is the breast parenchymal density with regard to the prevalence of fibroglandular tissue in the breast as it appears on a mammogram. The relation between mammographic parenchymal density levels and high risk of breast cancer was first shown by [13], using four distinct classes for breast parenchymal density categorization. Thus, mammographic images with high breast density value should be examined more carefully by radiologists, for both physiological and imaging risk factors, creating a need for automatic breast parenchymal density estimation algorithms. In [14], such algorithms are presented and a new technique, introducing a histogram distance metric, achieves good results. Some existing algorithms, e.g., [15, 16], use the texture information of mammograms, in order to extract more features for breast density estimation.

Radiologists try also to detect possible asymmetry between the left and the right breast in a pair of mammograms, as it can provide clues about the presence of early signs of tumors such as parenchymal distortion. Many CAD systems analyze automatically the images of a mammogram pair and provide results for the detection of asymmetric abnormalities by applying some type of alignment and direct comparison, as implemented by [17]. In the works of [18, 19], directional analysis methods are proposed, using Gabor wavelets, in order to detect possible asymmetry.

In this work, we propose a fully automated and complete segmentation methodology as the first stage of a multi-stage processing procedure for mammographic images [20–22]. Specifically, we have chosen to implement and apply the algorithm presented by [14] for breast boundary extraction, as the first step of the composite processing procedure; for the second step of pectoral muscle estimation, we enhanced the algorithm presented by [10] in order to achieve improved results; as a third step, we propose a new nipple detection technique, using the output of the breast boundary extraction procedure, when the nipple is in profile; that is, when it is projected on the background area of the mammogram, which is the recommended and usual case. The last algorithm, that is proposed in this work, besides locating the nipple point, can also serve as an improvement for the existing breast boundary algorithm, which misses the nipple if it is in profile. The improvement is obtained when updating the breast boundary, in order to include the detected nipple. Furthermore, as a fourth step, a new breast parenchymal density estimation algorithm is proposed, using segmentation of the inner-breast tissue, first-order statistics and fractal-based analysis of the mammographic image for the extraction of new statistical features, while the classification task is performed using Support Vector Machines (SVMs). Finally, a new algorithm is proposed for breast asymmetry detection, using the feature values already extracted from the breast parenchymal density estimation step, using an one-class SVM classifier. Both techniques achieve high success rates, often higher than the corresponding values of other algorithms in the relevant literature, while simpler and faster feature extraction methods have been employed. Our methodology has been tested on all the 322 mediolateral oblique view mammograms of the complete miniMIAS database, which is provided by [23], giving prominent results according to specific statistical measures and evaluation by expert radiologists, even in the case of such a difficult (very noisy) mammographic dataset.

2.2 Results and discussion

The complete system described was used for processing all the images of the miniMIAS database. All the intermediate results, i.e., breast boundary detection, pectoral muscle detection, nipple detection, asymmetry detection and breast density estimation, were examined in detail and evaluated by expert radiologists. It should be noted that the high level of noise, added to the images during the digitization process and the creation of the initial database images, makes the fully automated segmentation process a very challenging task.

The pre-processing techniques, which were selected to be applied in this work, were in general proved to be effective and successful, as the noise is correctly detected in most cases and sufficiently removed from the remaining stages of processing the images. The implemented breast boundary detection technique, which is based on a simple inference, gives satisfactory results. This is obvious by a careful observation of the detected boundary of the images and also verified accordingly, as it is compared to the ground truth boundary using specific statistic measures, such as the Tannimoto Coefficient and the Dice SimilarityCoefficient. The pectoral muscle estimate is accurate and further improved through the modification we propose, according to specific statistical measures extracted by the evaluation of the images from an expert radiologist. The new nipple detection technique tries to overcome the drawback of the breast boundary estimation method, i.e., not detecting the nipple, when this is in profile. In this way, it can serve as an improvement for the already established breast boundary, and in addition as a key point for further processing of the image, due to the importance of the nipple area in a mammographic image. Note that this technique can not be objectively compared to the algorithms proposed in previously published relevant literature, since the most similar one is the work by [12], which uses only a small subset of the miniMIAS database and has a different target than ours. The results were evaluated by expert radiologists and are promising enough to expect even better results, when applied to high quality digital mammograms.

The proposed algorithm for mammographic breast density estimation was tested on all the images of the miniMIAS database, fully annotated according to the 3 breast density classes. The results showed an accuracy of up to 85.71%, using the leave-one-out evaluation methodology. The achieved results are better compared to the relevant work of the bibliography [14, 16], although the latter one uses only a selected small portion of the miniMIAS database. The work of [15] achieves higher success rates, albeit it uses a different approach with higher-order textural features, which are computationally very expensive. The work we propose in this paper uses simple first-order statistical features and a new technique for the power spectrum estimation, making the whole process suitable for on-line training updates and real-time applications.

The asymmetry detection scheme was applied to all the images of the miniMIAS database, which is fully annotated, by characterizing each pair of mammograms as symmetric or asymmetric. The proposed methodology achieved a success rate of up to 84.47%. This success rate is similar to or even higher than the levels reported in the relevant literature, although it uses the complete set of images of the miniMIAS database, instead of a small subset, as the work of [18, 19]. Therefore, our experimental results can be considered more reliable and consistent. Furthermore, the use of the one-class classification algorithm turned out to be a simple yet effective way to overcome the problem of the imbalanced classes. The idea of the classification is to model as “target” the asymmetric cases and consider as “outliers” all the other cases, leading to an one-class scheme. The symmetric cases are not specifically modelled, but simply considered as non-asymmetric. In addition, note that our method is computationally much simpler and, more importantly, it is based on feature values that have already been computed and used. Thus, our method addresses the tasks of mammographic breast density estimation and asymmetry detection in an automatic, unified and generic way.

All the previously reported techniques can be combined and integrated to a clinical-level CAD system. All the algorithms are fully-automated and there is no need for external assistance. In addition, the processing time is not large enough, so each mammogram can be analyzed online; that is, on the fly as it is inserted the system. Moreover, the proposed scheme is considered to be robust against noise, as it has been verified by its application to the miniMIAS mammographic images database, in which the noise levels are very high and of varying nature.

3 Shape-based tumor retrieval in mammograms using relevance feedback techniques

3.1 Overview of work

As a second CAD system, a real application of a content-based medical image retrieval system is presented, while Relevance Feedback (RF) techniques are employed, in order to incorporate the radiologist to the retrieval process and further improve the results [24]. The system retrieves mammograms containing masses of the same morphology as the query image. The adopted features for the shape description are first-order statistics (mean value, standard deviation, mass circularity, entropy, area ratio parameter, zero-crossing count, roughness index) of the Normalized Radial Distance [25], extracted from the mass boundary. For the classification of the masses at step 0 of the RF procedure, a simple Euclidean minimum distance classifier [26] is used. On the next steps of the process, an SVM classifier is trained according to the feedback of the user. Note that we examine the performance of two different systems, in order to investigate the importance of the type of the patterns that the search engine returns to the user for labelling. In the simple SVM case [27], the system returns the most "confident" relevant patterns for labelling, e.g., the furthest patterns to the positive (relevant) side of the classifier. This can be easier for the user, but gives no useful information to the system, leading to slow convergence. However, in the active SVM case [28], the system returns the most "ambiguous" patterns for labelling, e.g. the patterns closest to the decision boundary, in order to improve the speed of the convergence. In order to evaluate the performance of the retrieval results at each round of the RF, the precision curve [29] is used. The precision at each round is defined as $pr = \frac{R}{N}$, where $N = 10$ is the total number of returned images to the user and R are the relevant images among them.

3.2 Results and discussion

For the evaluation of the Relevance Feedback scheme, a dataset of 2274 masses of the DDSM database [30], are used. Note that apart from the detailed boundary of each mass -used for the feature extraction step-, a classification of the shape of the masses in the following 7 distinct classes is also available: Irregular, Lobulated, Lymph Node, Oval, Round, Tubular or Other.

The experiments were carried out according to the following scenario:

- The user chooses a mass from the database as query image
- Repeat for steps 0 (no feedback yet) to 10 (user gave feedback 10 times)
  - The system returns to the user 10 images for evaluation and the precision is estimated
  - The system returns to the user 10 images to label
  - The user labels a subset of the images, as "relevant" or "non-relevant"
  - The system is re-trained, using the feedback of the user as new information

The above scenario is repeated for all the images of the database, in order to achieve more focused results. The system uses the simple SVM scheme [27], or the active SVM scheme [28]. In addition, the user is modeled as follows:

- The 'patient' user, that marks all the patterns returned by the system at each step as relevant or non-relevant, that can lead to imbalanced training sets.
- The less 'patient' user, that marks up to four relevant and four non-relevant patterns, among the patterns that the system returns at each step.
- The 'impatient' user, that marks up to three relevant and three non-relevant patterns, among the patterns that the system returns at each step.
- The 'lazy' user, that marks up to two relevant and two non-relevant patterns, among the patterns that the system returns at each step.

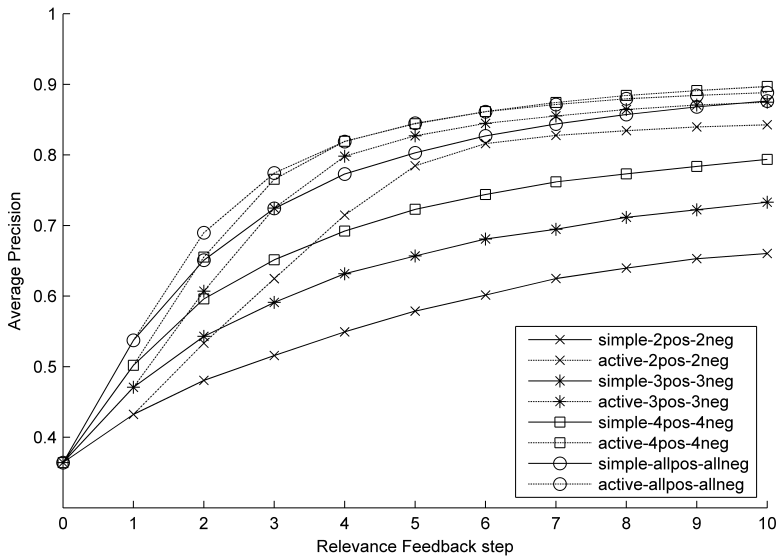

Fig. 1. Average precision at different steps of the RF procedure.

The average precision achieved at each iteration step for all the above configurations is shown in figure 1. Note that all the curves start from the same point at step 0, as no information is given from the user. At step 1, the simple and active techniques of the same type of user achieve equal precision rate, as the available images at step 0 for each user type are the same for these two scenarios. However, at step 1 the user of the active scenario provides more informative feedback than the one of the simple scenario, leading to a quicker convergence of the classifier. This is the reason for the fact that active SVM outperforms the simple SVM at steps greater than or equal to 2, always for the same type of user. The maximum precision rate of 89.7% is observed for the case of active scenario that the user marks up to 4 relevant and 4 non-relevant patterns and not for the 'patient' user, because probably the latter one creates sometimes imbalanced training sets.

In this part of the thesis, Relevance Feedback has been employed as a complementary tool to a Computer Aided Diagnosis system, that retrieves masses with similar shape as the query one. The judgement of the radiologist is considered to be of high importance to such a sensitive system as a medical application, where the errors should be eliminated and therefore it is suggested to be taken into consideration. The results, which almost reach 90% precision rate, show that the retrieval process can be improved significantly, when the radiologist is incorporated in the retrieval process, even for a hard classification task of 7 classes, using features of first-order statistics.

The system converges much faster when the user is more actively involved in the process, by labeling more samples as "relevant" or "non-relevant". In addition, the active technique converges faster to better results than the simple one, while the average precision for each class follows the rules of the Relevance Feedback scheme. The mammographic dataset used for the evaluation is rather extensive, consisting of the large number of 2274 masses, categorized in 7 distinct classes; these facts ensure that the results presented are very useful, reliable and consistent.

The system is also available online for any user at [31].

4 Conclusions

The current thesis provides two CAD systems for the retrieval of mammographic images. The first one performs a -fully automated- segmentation of a mammogram. In addition, it estimated the breast density and detects possible asymmetry between a pair of images, corresponding to a pair of breasts. The system presented achieves high success rates, often higher than the works of the bibliography, although it uses simple -and inexpensive to compute- features. For this reason, the average processing time of an image is not large enough, so each mammogram can be analyzed online.

The other proposed CAD system examines the improvement achieved when RF techniques are used. More specifically, a "morphological analysis" retrieval system for mammograms is presented. It is evaluated on an extensive dataset of masses belonging to 7 distinct classes. The results, which almost reach 90% precision rate, prove that the retrieval process can be improved significantly, when the radiologist is incorporated to the retrieval process, even for such a hard classification task.

References

1. R. Nishikawa, "Current status and future directions of computer-aided diagnosis in mammography," *Computerized Medical Imaging and Graphics*, vol. 31, no. 4-5, pp. 224–235, 2007.
2. M. Siddiqui, M. Anand, P. Mehrotra, R. Sarangi, and N. Mathur, "Biomonitoring of organochlorines in women with benign and malignant breast disease," *Environmental Research*, vol. 98, no. 2, pp. 250–257, 2005.
3. A. Wroblewska, P. Boninski, A. Przelaskowski, and M. Kazubek, "Segmentation and feature extraction for reliable classification of microcalcifications in digital mammograms," *Opto-Electronics Review*, vol. 11, no. 3, pp. 227–235, 2003.
4. R. Rangayyan, F. Ayres, and J. Leo Desautels, "A review of computer-aided diagnosis of breast cancer: toward the detection of subtle signs," *Journal of the Franklin Institute*, vol. 344, no. 3-4, pp. 312–348, 2007.
5. T. Gevers and A. Smeulders, "Content-based image retrieval: An overview." Prentice Hall, 2004.
6. R. Datta, J. Li, and J. Wang, "Content-based image retrieval: approaches and trends of the new age," in *Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval, November.* Citeseer, 2005, pp. 10–11.
7. X. Zhou and T. Huang, "Relevance feedback in image retrieval: A comprehensive review," *Multimedia systems*, vol. 8, no. 6, pp. 536–544, 2003.
8. W. Song, C. Huangshan, and T. Hua, "Analytic Implementation for Medical Image Retrieval Based On FCM Using Feature Fusion With Relevance Feedback," in *Proceedings of the 2nd International Conference on Bioinformatics and Biomedical Engineering, 2008. ICBBE 2008.*, 2008, pp. 2590 – 2595.
9. I. El-Naqa, Y. Yang, N. Galatsanos, R. Nishikawa, and M. Wernick, "A similarity learning approach to content-based image retrieval: application to digital mammography," *IEEE Transactions on Medical Imaging*, vol. 23, no. 10, pp. 1233–1244, 2004.
10. S. Kwok, R. Chandrasekhar, Y. Attikiouzel, and M. Rickard, "Automatic pectoral muscle segmentation on mediolateral oblique view mammograms," *Medical Imaging, IEEE Transactions on*, vol. 23, no. 9, pp. 1129–1140, 2004.
11. V. Andolina, S. Lillé, and K. Willison, *Mammographic imaging: a practical guide.* Lippincott Williams & Wilkins, 2001.
12. R. Chandrasekhar and Y. Attikiouzel, "A simple method for automatically locating the nipple on mammograms," *Medical Imaging, IEEE Transactions on*, vol. 16, no. 5, pp. 483–494, 1997.
13. J. Wolfe, "Risk for breast cancer development determined by mammographic parenchymal pattern." *Cancer*, vol. 37, no. 5, pp. 2486–92, 1976.
14. M. Masek, E. University of Western Australia School of Electrical, C. Engineering, and U. of Western Australia Centre for Intelligent Information Processing Systems, "Hierarchical Segmentation of Mammograms Based on Pixel Intensity," Ph.D. dissertation, 2004, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.128.1245.

15. A. Bosch, X. Munoz, A. Oliver, and J. Martı, "Modeling and Classifying Breast Tissue Density in Mammograms," in *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Volume 2*. IEEE Computer Society Washington, DC, USA, 2006, pp. 1552–1558.
16. A. Oliver, J. Freixenet, A. Bosch, D. Raba, and R. Zwiggelaar, "Automatic classification of breast tissue," in *Iberian Conference on Pattern Recognition and Image Analysis*. Springer, 2005, pp. 431–438.
17. F. Yin, M. Giger, K. Doi, C. Vyborny, and R. Schmidt, "Computerized detection of masses in digital mammograms: Automated alignment of breast images and its effect on bilateral-subtraction technique," *Medical Physics*, vol. 21, p. 445, 1994.
18. R. Ferrari, R. Rangayyan, J. Desautels, and A. Frere, "Analysis of asymmetry in mammograms via directional filtering withGabor wavelets," *Medical Imaging, IEEE Transactions on*, vol. 20, no. 9, pp. 953–964, 2001.
19. R. Rangayyan, R. Ferrari, and A. Frere, "Analysis of bilateral asymmetry in mammograms using directional, morphological, and density features," *Journal of Electronic Imaging*, vol. 16, no. 1, pp. 13 003–13 003, 2007.
20. S. Tzikopoulos, H. Georgiou, M. Mavroforakis, N. Dimitropoulos, and S. Theodoridis, "A fully automated complete segmentation scheme for mammograms," in *Digital Signal Processing, 2009 16th International Conference on*. IEEE, 2009, pp. 1–6.
21. S. Tzikopoulos, H. Georgiou, M. Mavroforakis, and S. Theodoridis, "A fully automated scheme for breast density estimation and asymmetry detection of mammograms," in *17th European Signal Processing Conference (EUSIPCO)*, 2009.
22. S. Tzikopoulos, M. Mavroforakis, H. Georgiou, N. Dimitropoulos, and S. Theodoridis, "A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry," *Computer Methods and Programs in Biomedicine*, 2011.
23. J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok *et al.*, "The Mammographic Image Analysis Society Digital Mammogram Database," in *Exerpta Medica. International Congress Series*, 1994, pp. 375–378.
24. S. Tzikopoulos, H. Georgiou, M. Mavroforakis, and S. Theodoridis, "Shape-Based Tumor Retrieval in Mammograms Using Relevance-Feedback Techniques," *Artificial Neural Networks–ICANN 2010*, pp. 251–260, 2010.
25. J. Kilday, F. Palmieri, and M. Fox, "Classifying mammographic lesions using computerized image analysis," *IEEE Transactions on Medical Imaging*, vol. 12, no. 4, pp. 664–669, 1993.
26. S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed. Academic Press, 2009.
27. H. Drucker, B. Shahrary, and D. Gibbon, "Relevance feedback using support vector machines," in *MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE-*. Citeseer, 2001, pp. 122–129.
28. S. Tong and E. Chang, "Support vector machine active learning for image retrieval," in *Proceedings of the ninth ACM international conference on Multimedia*. ACM, 2001, pp. 107–118.
29. J. Luo and M. Nascimento, "Content-based sub-image retrieval using relevance feedback," p. 9, 2004.
30. M. Heath, K. Bowyer, D. Kopans, R. Moore, and P. Kegelmeyer, "The digital database for screening mammography," pp. 212–218, 2000.
31. Image Processing Techniques for Mammographic Images , http://mammo.di.uoa.gr, 2010.