SIFT texture description for understanding breast ultrasound images

Joan Massich[1 2 *], Fabrice Meriaudeau[2], Melcior Sentís[3], Sergi Ganau[3], Elsa Pérez[4], Domènec Puig[5], Robert Martí[1], Arnau Oliver[1], and Joan Martí[1]

[1] Computer Vision and Robotics Group, University of Girona, Spain.
jmassich@atc.udg.edu
[2] Laboratoire Le2i-UMR CNRS, University of Burgundy, Le Creusot, France.
[3] Department of Breast and Gynecological Radiology, UDIAT-Diagnostic Center, Parc Taulí Corporation, Sabadell, Spain.
[4] Department of Radiology, Hospital Josep Trueta of Girona, Spain.
[5] Department of Computer Engineering and Mathematics, Rovira i Virgili University, Tarragona, Spain.

Abstract. Ultra-Sound (US) imaging is the most common adjunct image modality to assess breast cancer. In order to interpret such images, texture interpretation is of major importance. This paper proposes to use Self-Invariant Feature Transform (SIFT) descriptor as texture and provides a feasible interpretation of its behavior related to the underlying depicted tissue type. The proposal has been evaluated using a set of breast images with accompanying expert-provided Ground Truth (GT) which describes all the tissues present within the images.

Key words: breast cancer, ultrasound, texture, SIFT

1 Introduction

Breast cancer is the second most common cancer (1.4 million cases per year, 10.9% of diagnosed cancers) after lung cancer, followed by colorectal, stomach, prostate and liver cancers. In terms of mortality, breast cancer is the fifth most common cause of cancer death. However, it places as the leading cause of cancer death among females both in western countries and in economically developing countries [3].

Medical imaging plays an important role in breast cancer mortality reduction, contributing to its early detection through screening, diagnosis, image-guided biopsy, treatment follow-up and suchlike procedures [5]. Despite Digital Mammography (DM) still remains as the image modality of reference for diagnose purposes, US offers useful complementary diagnose information due to its capabilities for differentiating between solid lesions that are benign or malignant [6].

---

* This work was partially supported by the Spanish Science and Innovation grant nb. TIN2012-37171-C02-01 and TTIN2012-37171-C02-02, the Regional Council of Burgundy and the University of Girona BR grant nb. 09/22

It is estimated that between 65 $\sim$ 85% of the biopsies prescribed using only mammography imaging could be avoided if US information had been taken into account while issuing the diagnose [7].

In US images, texture is a major characteristic to distinguish between different breast tissues, which also allows assessing of the lesion's pathology [6]. Therefore, the importance of incorporating texture data from US images into Computer Aided Diagnosis (CAD) systems. A comprehensive list of texture descriptors used for detection, segmentation or diagnose tasks applied to US breast images is given in Cheng *et al.* [1]. Where most of the descriptors are ad-hock descriptors or based on well-known texture descriptors such as co-occurrence matrices, wavelet coefficients or Gray-Level Difference Method (GLDM).

In this article explore the usage of SIFT descriptors for encoding the US characteristic texture produced by the speckle noise present within the images, we evaluate its performance in annotated dataset.

2 Material and methods

In order to develop segmentation methodologies applied to delineate breast lesions in US data, a set of 700 US images was acquired at the *UDIAT Diagnostic Centre of Parc Taulí* in Sabadell (Catalunya), between 2010 and 2012. All with accompanying GT delineation of the lesions present in the image. From this image database, a reduced dataset of 16 images corresponding to different patients was selected and complemented with multi-label GT in order to evaluate the texture description of the observable tissues in the breast.

Figure 1 illustrates a breast image from the dataset with its associated GT.

3 Using SIFT as a low-level texture descriptor in order to differentiate the tissues present in breast US images

Self-Invariant Feature Transform (SIFT) [4] transforms key-points into scale and rotation invariant coordinates relative to local features by sampling the magnitude and orientation of the gradients surrounding the key-point to generating a 128 element feature. In order to generate a texture descriptor out of SIFT descriptors, a regular grid is used to generate evenly sparse SIFT descriptors. At

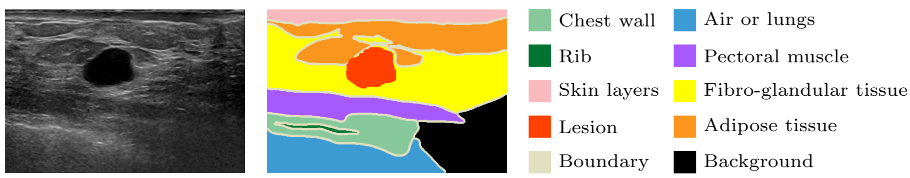
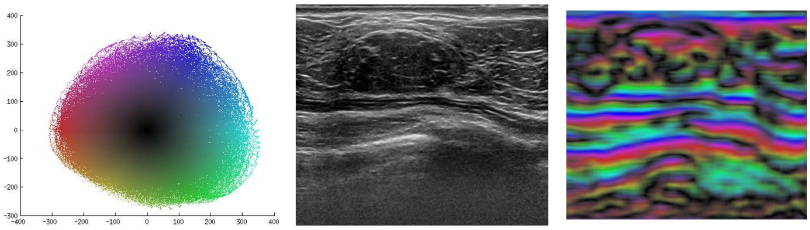


Fig. 1: Dataset sample. From left to right: image sample, accompanying multi-label GT, tissue label GT color-coding.



(a) (b) (c)

Fig. 2: Low level SIFT descriptor example. (a) Arbitrary coloring of the projected SIFT space. (b) Original image. (c) Recoding of the extracted SIFT descriptors using the color coding in (a).

the finest resolution every single pixel of an image can be used as a key-point to extract a SIFT descriptor out of every possible position.

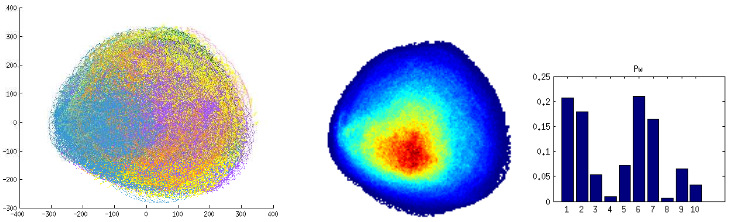
The usage of SIFT descriptor brings invariability to scale, rotation and minor affine transformations along with robustness to illumination changes [4], which allows to differentiate the tissues despite the variability from US adquisition.

Figures 3 to 7 are used to analyze both qualitatively and quantitatively the usage of SIFT as a low-level descriptor to encode US texture. In order to study the image in terms of SIFT descriptors the images are mapped to this SIFT space by extracting a SIFT descriptor at every pixel position. In order to visually keep track of this SIFT, their 128 dimensions are projected into a two dimensional spce using Principal Component Analysis (PCA). When combining features using PCA is convenient to know the amount of data variability captured in the new projected space. For the current experiment, the projected space describes $21,6\%$ of the data variability present in the original space. Figure 2 offers a visual interpretation of a breast US image in terms of low-level SIFT descriptors. On it, fig. 2a shows the scatter plot resulting from projecting all the images within the dataset, at the SIFT space and further project them at the space described by their two principal components. In this projected space, every sample has been arbitrary colored in such a manner that two close samples share similar color. In such manner, this arbitrary coloring can be used to color code the SIFT descriptors and visually asses the images in terms of SIFT (see fig. 2c).

In order to analyze the tissue distribution in this texture space of SIFT descriptors, the analysis of the Maximum A Posteriori (MAP) estimator has been chosen. The model is presented in equation 1 and figures 3 to **??** visually interprets its terms.

$$P(\omega|\bar{x}_a) = \frac{P(\bar{x}_a|\omega) \cdot P(\omega)}{P(\bar{x}_a)} \tag{1}$$

Where $P(\omega|\bar{x}_a)$ is the probability that the sample a belongs to class $\omega \in W$ (see fig. 1b for all the possible classes) based on its position in the projected space. $P(\bar{x}_a|\omega)$ corresponds to the Maximum Likelihood (ML) of the feature



(a) (b) (c)

Fig. 3: 2D visualization of the SIFT hyper space. (a) Projected space colored according to GT tissue labeling. (b) Occurrence density.

distribution for a particular class ω and, $P(\omega)$ and $P(\bar{x}_a)$ are the priors for the class and feature respectively.

Figure 3 uses the entire dataset to illustrate the underlying problem and the priors extracted from the same dataset. Fig. 3a shows a scatter plot where every sample has been colored according to its GT. Fig. 3b shows an occurrence study of the samples carried out in a discretization of the space in fig. 3a that illustrates the distribution of the SIFT descriptors present within the images, which corresponds to the $P(\bar{x}_a)$ term in eq. 1. Fig. 3c illustrates the class prior corresponding to the proportion of samples present in the dataset for each class.

Figure 4 shows the occurrence study for every class, corresponding to the $P(\bar{x}_a|\omega)$ in eq. 1 where every different tissue type is represent by different SIFT descriptors.

fa falta una comparacio dels models que son similars referince a l image de la figura 1 per illustrar que el background normalment correspon a zones que ja no es visualitza res, que ja no s entenen i que per tant son molt similars a aire, o diferencies entre greix i fibroglandular, etc... En definitiva l overlap de les classes

Using eq. 1 a MAP estimation the probability for each class can be calculated at every position based on the data shown in fig. 3 and fig. 4. Figure 5a colors the space based on the most probable class. As it can be observed, the coloring of the space is mostly congruent indicating that each tissue is grouped within the SIFT which facilitates class separability. This coloring of the space can be used as a Bayesian classification of the sample. Figure 5c shows the result of this classification from an unseen image (see fig. 5b). Since no multi-label GT is present for this image only qualitative assessment of the results can be made.

falta la imatge

In order to improve the results, MRF can be used to ensure spatial coherence for the labeling. Also instead of only using the map, the reward obtained from the other classes can also be taken into account. similar classes are easily mislabeled ej: background and lungs, fibro-glandular and fat

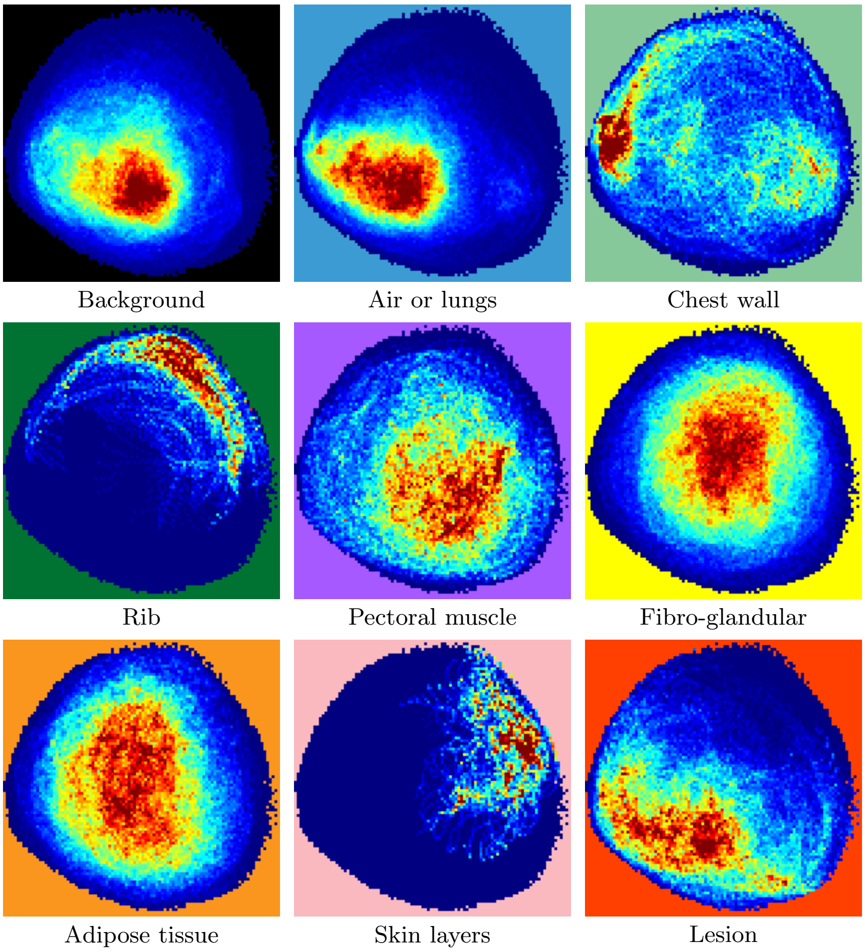


Fig. 4: Distribution of the SIFT descriptors for some of the tissue type.

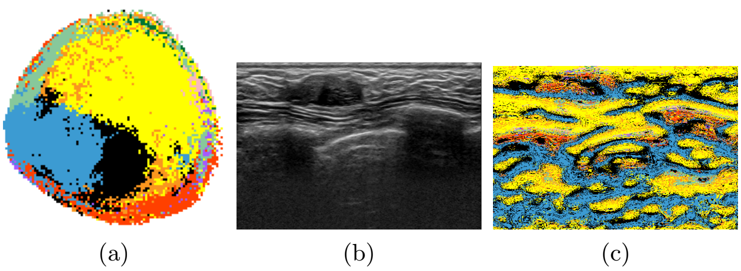


Fig. 5: Qualitative results. (a) MAP class label distribution of the 2D projection of the SIFT descriptors. (b) unseen image (c) image labeling results

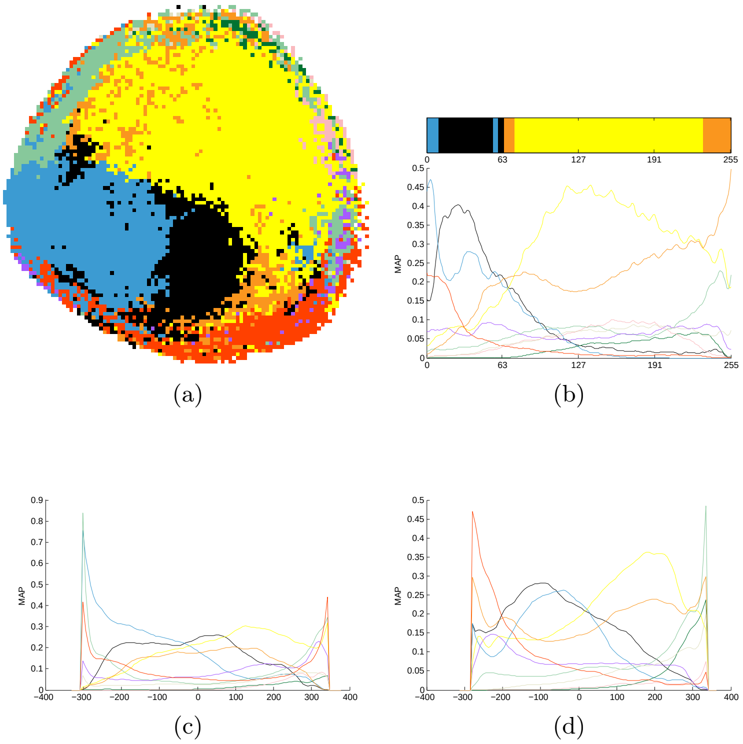
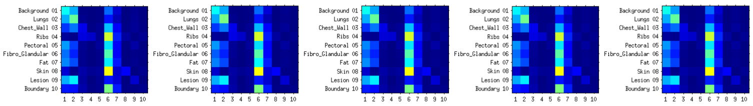


Fig. 6: Qualitative evaluation of the MAP labeling of the feature space. (a) SIFT (b) Intensity



(a)

Fig. 7

3.1 Quantitative assessing

In order to generate quantitative results using cross-validation, 5 folds have been generated by sampling extracting 10.000 samples for each class from the dataset. At each round 4 folds have been used for training the ML term in eq. 1 ($P(\bar{x}_a|\omega)$) and the classification has been tested using the remaining fold.

4 High-level texture descriptor using Bag-of-Features (BoF) and SIFT descriptors

Texture is an area property related to spatial repetition of structures, similar statistical properties of the area or both. A technique to embed statistical properties of a low level descriptor is BoF which analyses the occurrence of a set of keywords (or key-points) within a particular region [2].

In this proposal, the words or features representing the images are SIFT descriptors. In order to generate the dictionary that generates the BoF descriptor, the space of SIFT descriptors is clustered in order to produce a hard quantification of this space. In this case, a k-means procedure with $k = 36$ is used to generate the dictionary. To generate the BoF-SIFT feature, all the SIFT descriptors are substituted for the closest SIFT descriptor in the dictionary. Finally the texture description from a particular area is expressed as the keywords' occurrence in this area. The descriptor is normalized so that the sum of all the occurrences is 1. Figure 10 illustrates the BoF-SIFT feature for some example areas.

5 Quantitative experiment

In order to assess the usage of SIFT texture, the images are divided into superpixels, features are extracted to describe the superpixels and supervised ML procedure using Support Vector Machine (SVM) is used to infer the labeling of the superpixel. Incorporating BoF of SIFT descriptors in order to describe the superpixels improve the results of the labeling process for our preliminary experiments, however a more thorough study will be included in the final version of the document.

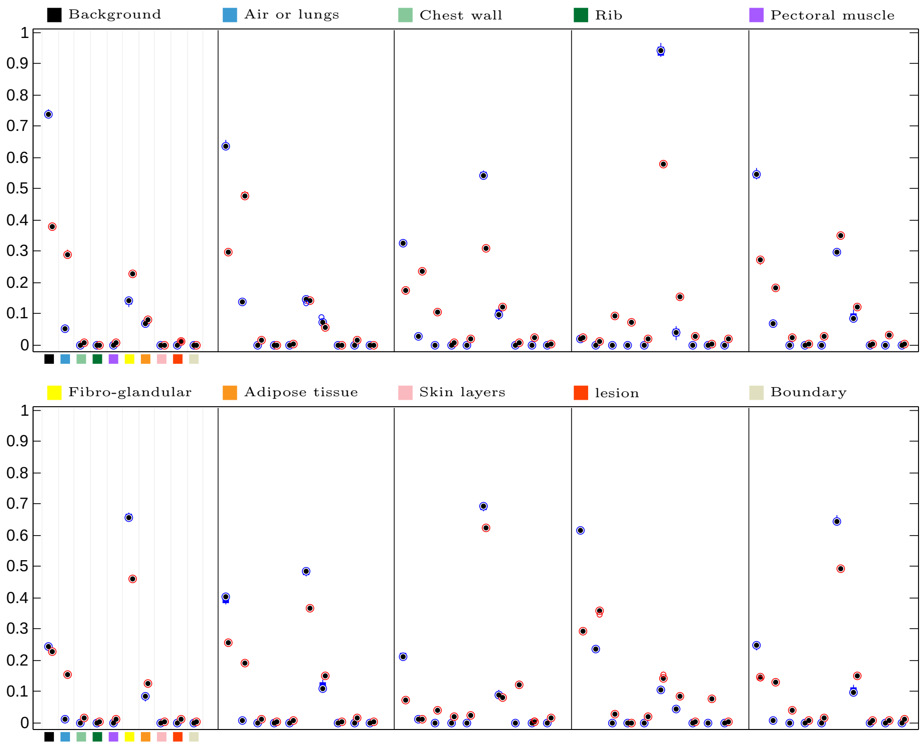


Fig. 8

(a)

(b)

Fig. 9: Two views of a generated codebook (dictionary) example. (a) Clusters mapped in the 2D projected SIFT space. (b) Simple concatenation of the codebook for further usage visualizing super pixel's signature.

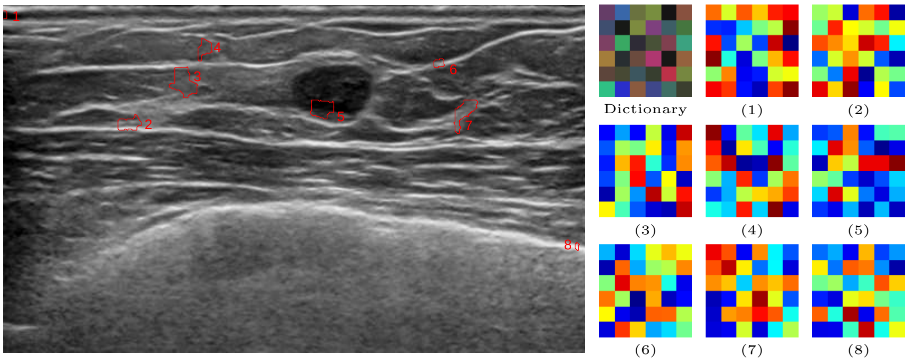


Fig. 10: SIFT-BoF descriptors qualitative analysis. (Left) image example. (Right) Dictionary representation colored using the location of the keypoint location in fig. 3a space. (1-8) Occurrence of the dictionary's key-points associated to each region highlighted in the original image.

6 Conclusion

Experiments show that SIFT is a worth candidate to incorporate in segmentation procedures, tissue charecterization and...

References

1. H D Cheng, Juan Shan, Wen Ju, Yanhui Guo, and Ling Zhang. Automated breast cancer detection and classification using ultrasound images: A survey. *Pattern Recognition*, 43(1):299–317, 2009.
2. Gabriella Csurka, Christopher Dance, Lixin Fan, Jutta Willamowski, and Cédric Bray. Visual categorization with bags of keypoints. In *Workshop on statistical learning in computer vision, ECCV*, volume 1, page 22, 2004.
3. Ahmedin Jemal, Freddie Bray, Melissa M. Center, Jacques Ferlay, Elizabeth Ward, and David Forman. Global cancer statistics. *CA: A Cancer Journal for Clinicians*, 61(2):69–90, 2011.
4. David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.
5. Robert A Smith, Debbie Saslow, Kimberly Andrews Sawyer, Wylie Burke, Mary E Costanza, WPIII Evans, Roger S Foster, Edward Hendrick, Harmon J Eyre, and Steven Sener. American cancer society guidelines for breast cancer screening: update 2003. *CA: a cancer journal for clinicians*, 53(3):141–169, 2003.
6. A Thomas Stavros. *Breast ultrasound*. Lippincott Williams & Wilkins, 2004.
7. Yading Yuan, Maryellen L Giger, Hui Li, Neha Bhooshan, and Charlene A Sennett. Multimodality computer-aided breast cancer diagnosis with ffdm and dce-mri. *Academic radiology*, 17(9):1158, 2010.