

Performance Evaluation of Geometric Feature Descriptors With Application to Classification of Small-Size Lung Nodules in Low Dose CT

Amal A. Farag

Department of Radiology
National Institute of Health, Bethesda, Maryland, USA
E-mail: amal.aly1@gmail.com

Abstract: Object modeling is a multifaceted and active area of research in image analysis, manufacturing, and computational geometry. In image analysis, the focus of this paper, objects may be described in terms of their shape and appearance. Geometric features that are invariant to scale, rotation and translation are essential for proper modeling of shape. Deformable random objects, common in biomedical imaging, do not carry distinct features and suffer from multiple sources of uncertainty, including resolution, size, location and occlusion. Registration involves generating the transformation parameters for mapping a source to a target, while categorization involves matching entities into classes. Both processes, involve mapping attributes of objects with respect to each other, and robust features together with proper similarity measures are keys for good performance. This article examines the issues of features, similarity measures for registration and categorization of small-size deformable random objects; a case study of lung nodules in chest CT is used as an example. We show that feature descriptors capturing textural contents are more effective for these types of objects. The paper includes survey and evaluations of several modern feature descriptors that have been introduced in the computer vision and image analysis literature.

Keywords: feature descriptors, SIFT, SURF, LBP, deformable objects, object classifications, lung nodules

1. Introduction

Objects may be represented by various forms, however, robust representations must maintain the features that describe the objects and enable analysis and decision making. As objects may not have a specific geometric description, features may not be as easy to specify. Likewise, known features about an object may be altered in the imaging process. Furthermore, if an object is to be compared with similar ones in a database, it is important that features involved in this comparison be robust (ideally invariant) to changes in scale, rotation and translation. This paper deals with feature definitions and characterization through feature descriptors. Good surveys of feature interest point detectors and feature descriptors exist in the computer vision literature (e.g., [1]-[5]). Biomedical objects are usually marred with uncertainty of the imaging process, randomness of shape, size, location and occlusion. Only a few biomedical objects possess specific shape characteristics (e.g., the kidney, chest bones, and the corpus callosum); hence, precise feature definitions for biomedical objects is not a straight forward matter. This creates difficulties for such applications as object detection and recognitions; and complicates fundamental image analysis tasks such as segmentation and registration. In this paper we focus on biomedical objects that the authors have worked on during the past several years (e.g., [6] [7]).

A local feature is a point or pattern in an image that differs from its immediate neighborhood, and is associated with a change of an image property or a number of properties simultaneously. The image properties commonly used include intensity, color, and texture. Local features can be points, edges or small image patches. A feature descriptor is a representation of a feature, or a region around a feature, computed from attributes such as intensity, color or texture of the feature or the region. Global features describe objects as a whole using a single descriptor. Fig. 1 shows test images with distinct features; with corners (left) and textures/intensities (right). These images will be used to measure effectiveness of various corner detectors and descriptors with respect to rotation and noise.

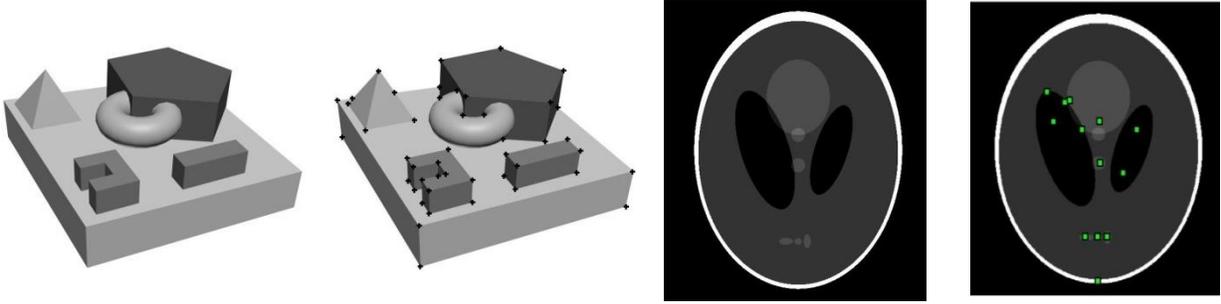


Fig. 1: Test images with features identified by a corner detector.

Graylevel histograms, color histograms, texture, shape context, and object geometric attributes such as area or perimeter, are examples of commonly used global features. Global features may not be as robust as local features if the objects have complicated topologies. Indeed, various objects may have similar graylevel histograms, for example; thus some global features may not provide satisfactory results in image analysis. In contrast to global features, local features describe the object of interest using a set of vectors representing patches around detected features or points of interest. For every detected interest point, a numerical vector called "feature descriptor" is built to describe that point, or a patch (local neighborhood), around it. Several approaches have been developed for this purpose; e.g., the SIFT, SURF and LBP algorithms, which will be studied in this paper. Feature matching may be conducted by various search methods based on a particular similarity or dissimilarity measure. Image registration is a form of matching to find the optimum rotation, translation and scale among a set of objects to re-align together. In this paper we examine some feature descriptors that have efficient algorithmic implementation to enable incorporation into such image analysis tasks as registration and recognition.

2. Local Descriptors

Image matching algorithms consist of three major parts: feature detector, feature descriptor, and feature matching. This section describes some of the feature detectors and descriptors common in image analysis.

2.1 Scale Invariant Feature Transform (SIFT)

As detailed in Lowe, 2004 [1], SIFT consists of four main steps: (1) scale-space peak selection, (2) keypoint localization, (3) orientation assignment, (4) keypoint descriptor.

Scale space selection: The scale space $\mathbf{L}(\mathbf{x}, \sigma_s)$ is constructed by the linear convolution of the image $\mathbf{I}(\mathbf{x})$ with a cylindrical Gaussian kernel $\mathbf{G}(\mathbf{x}, \sigma_s)$ which can be viewed as a stack of 2D Gaussians one for each band. The scale is discretized as $\sigma_s \in \{k^s\}$ where $k = 2^{1/3}$ and $s = \{-1, 0, 1, 2, \dots, \frac{\log(s_{max})}{1/3 \log 2}\}$. Scale-space extrema detection is performed through searching over all scales σ_s and image locations $\mathbf{x} = \{(x, y)\}$, in order to identify potential interest points which are invariant to scale and orientation. This can be efficiently implemented using Difference-of-Gaussians $\mathbf{D}(\mathbf{x}, \sigma_s)$ which takes the difference between consecutive scales, i.e. $\mathbf{D}(\mathbf{x}, \sigma_s) = \mathbf{L}(\mathbf{x}, \sigma_s) - \mathbf{L}(\mathbf{x}, \sigma_{s-1})$, where for a spectral band b , a point \mathbf{x} is selected to be a candidate interest point if it is larger or smaller than its $3 \times 3 \times 3$ neighborhood system defined on $\{D(\mathbf{x}, \sigma_{s-1}; b), D(\mathbf{x}, \sigma_s; b), D(\mathbf{x}, \sigma_{s+1}; b)\}$, where σ_s is marked to be the scale of the point \mathbf{x} . This process leads to too many points some of which are unstable (sensitive to noise); hence removal of points with low contrast and points that are localized along edges is accomplished. Examples of keypoints detection using the SIFT detector is shown in Fig. 2 for the Shepp and Logan model under moderate rotation and noise. Fig. 3 shows the keypoints detection using the SIFT detector for test image showing the effects of rotation and blur.

Keypoint localization: In order to obtain a point descriptor which is invariant to orientation, a consistent orientation should be assigned to each detected interest point based on the gradient of its local image patch. Considering a small window surrounding x , the gradient magnitude and orientation can be computed using finite differences. Local image patch orientation is then weighted by the corresponding magnitude and Gaussian window. Eventually the orientation is selected to be the peak of the weighted orientation histogram.

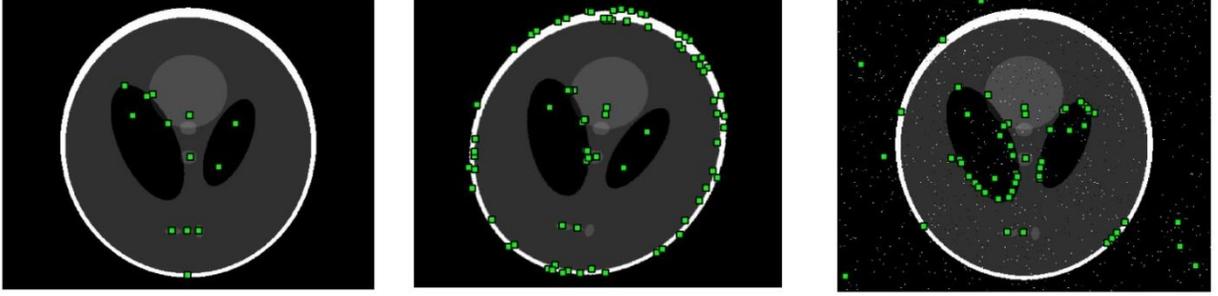


Fig. 2: Examples of keypoints detection using the SIFT detector with moderate rotation and noise.

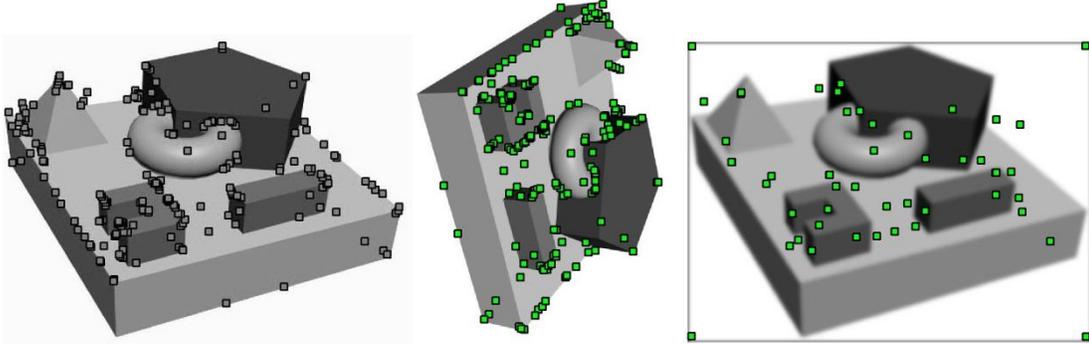


Fig. 3: Examples of keypoints detection using the SIFT detector with moderate rotation and blurring.

Building a point descriptor: The process of building a descriptor around a key point is similar to orientation assignment. A 16×16 image window surrounding the interest point \mathbf{x} is divided into sixteen 4×4 sub-window, an 8-bin weighted orientation histogram is computed for each sub-window, hence we end up with $16 \times 8 = 128$ descriptors for each interest point. Thus each detected interest point can now be defined at location, specific scale, certain orientation θ and a descriptor vector as $\mathbf{x} = \{x, y, \sigma, \theta, \mathbf{d}\}$.

Fig. 4 shows the plot of the 128 values of the SIFT descriptor under different blur, noise, rotation and scale levels at the same selected point on transformed images in Fig. 3.

Interest point matching: Interest point matching is performed to provide correspondences between the given images. Two points \mathbf{x}_i^t and \mathbf{x}_j^{t+1} with SIFT descriptors \mathbf{d}_i^t and \mathbf{d}_j^{t+1} are said to be in correspondence, if: $d_{L_2}(\mathbf{x}_i^t, \mathbf{x}_j^{t+1}) = \sqrt{\|\mathbf{d}_i^t - \mathbf{d}_j^{t+1}\|^2}$ is minimum. This measure is computed as by: $d_{L_2}(\mathbf{x}_i^t, \mathbf{x}_j^{t+1}) = \left(\sum_{k=1}^{128} |d_{ik}^t - d_{jk}^{t+1}|^2 \right)^{1/2}$.

Examples of matched feature pairs using SIFT are shown in Fig. 5 on the Shepp and Logan head phantom.

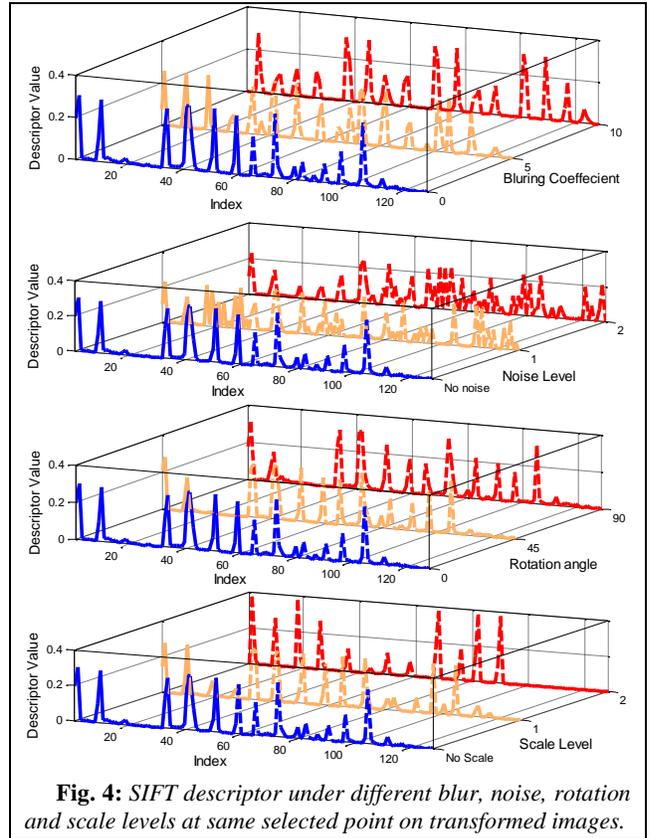


Fig. 4: SIFT descriptor under different blur, noise, rotation and scale levels at same selected point on transformed images.

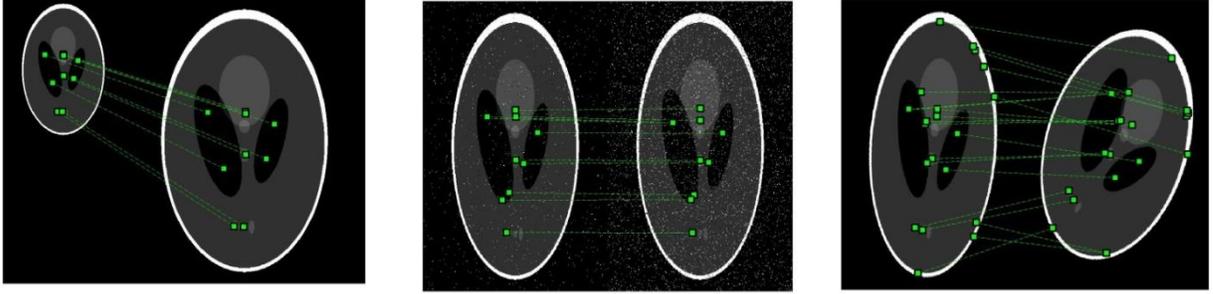


Fig. 5: Matched feature pairs using SIFT on the Shepp and Logan head phantom. Left (scale); middle (noise); and right (rotation) effects.

2.2 The speeded-up robust features (SURF)

The (SURF) descriptor (Bay et al., 2008 [4]) is a distribution of Haar-wavelet responses within the neighborhood of interest. The SURF descriptor consists of several steps; a square region is constructed around the interest point and oriented either in a rotation invariant method, where the Haar-wavelet response in the x- and y- directions are computed and weighted with a Gaussian centered at the interest point, or a non-rotation invariant method. The wavelet responses in both directions are then summed-up over each sub-region. The total number of descriptors for each point is 64. SURF uses mainly the texture information concentrated around interest points. Principle component analysis (PCA) and linear discriminate analysis (LDA) are used to project the extracted SURF descriptors to a low-dimensional subspace where noise is filtered out. We show the plot of the 64 values of the SURF descriptor under different blur, noise, rotation and scale levels at the same selected point on transformed images in Fig. 6.

2.3 Multi-Resolution Local Binary Pattern (LBP)

The Local Binary Pattern (Ojala et al., 2002 [5]) is an operator invariant to monotonic changes in grayscale and can resist illumination variations as long as the absolute gray-level value differences are not badly affected. The original operator labeled the pixels of an image by thresholding the 3×3 neighborhood of each pixel with the center value and considered the result as a binary number. At a given pixel position (x_c, y_c) , the decimal form of the resulting 8-bit word is given by the following equation: $LBP(x_c, y_c) = \sum_{i=0}^7 s(I_i - I_c)2^i$; where, I_c corresponds to the center pixel (x_c, y_c) , I_i to gray level values of the eight surrounding pixels and function $s(\cdot)$ is a unit-step function. The LBP operator was extended to a circular neighborhood of different radius size to overcome the limitation of the small original 3×3 neighborhood size failing to capture large-scale structures. Each instance is denoted as (P, R) , where P refers to the equally spaced pixels on a circle of radius R . The parameter P controls the quantization of the angular space and R determines the spatial resolution of the operator. An LBP pattern is considered uniform if it contains at most two bitwise transitions from 0 to 1 and vice-versa, when the binary string is circular. The reason for using uniform patterns is that they contain most of the texture information and mainly represent texture primitives. The operator is derived on a circularly symmetric neighbor set of P members on a circle of radius R denoting the operator as LBP_{PR}^{u2} . In the multi-resolution analysis the responses of multiple operators realized with different (P, R) are combined together and an aggregate dissimilarity is defined as the sum of individual log-likelihoods computed from the responses of individual operators. The notation LBP_{PR}^{u2} used in this paper refers to the extended LBP operator in a (P, R) neighborhood, with only uniform patterns considered.

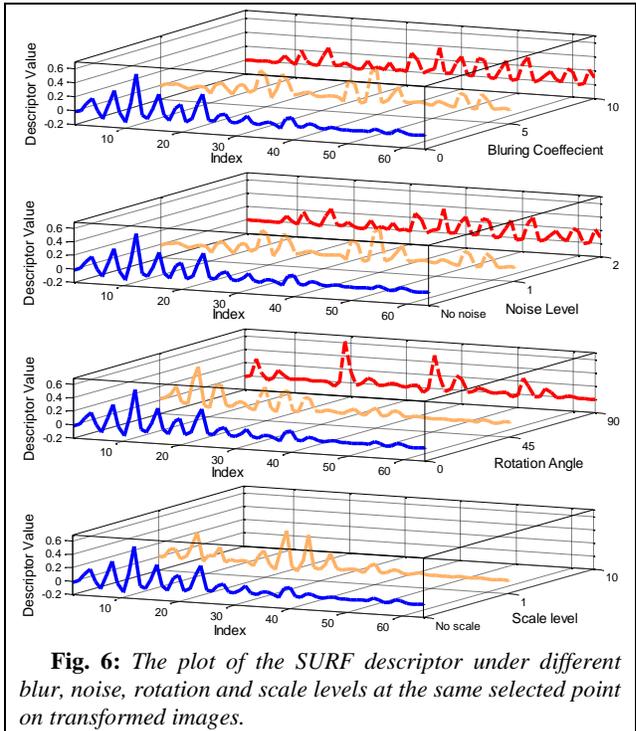


Fig. 6: The plot of the SURF descriptor under different blur, noise, rotation and scale levels at the same selected point on transformed images.

As we did for the SIFT descriptor, we show in Fig. 7 readings of the LBP for some keypoints on the test image in Fig. 1 under the effect of blur, noise, rotation and scale. In general, the LBP descriptor works well when the neighborhood around the keypoints have reasonable texture content. The comparison of the three descriptors is shown in the following subsections. In general, the LBP works better with high textural contents, whereas the SIFT provides better performance with robust definition of keypoints.

2.4 Performance Evaluation on Test Images

In order to test the performance of the three descriptor (SIFT, SURF, and LBP) we used the test images in Fig. 1, we manually selected a set of 46 points on the original image. Fig. 1(second column) shows the manually selected feature points. The location ground truth location of these points are calculated on every transformed image based on the transformation applied to generate this image. The descriptors are calculated at these points for all the images. The number of correct matched points are used as an evaluation criteria.

The results are shown in Fig. 8. The LBP showed a more robust performance with respect to noise, while the SIFT was more robust to rotation.

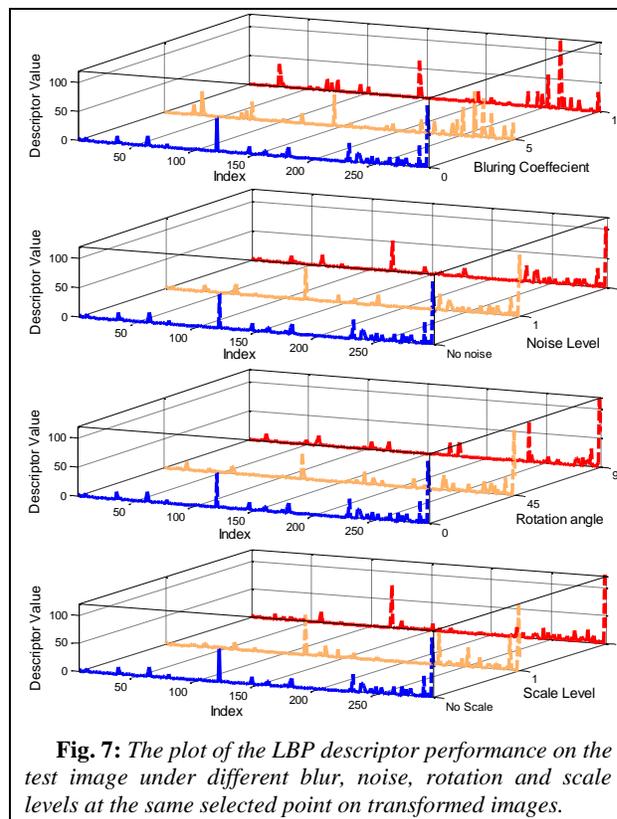
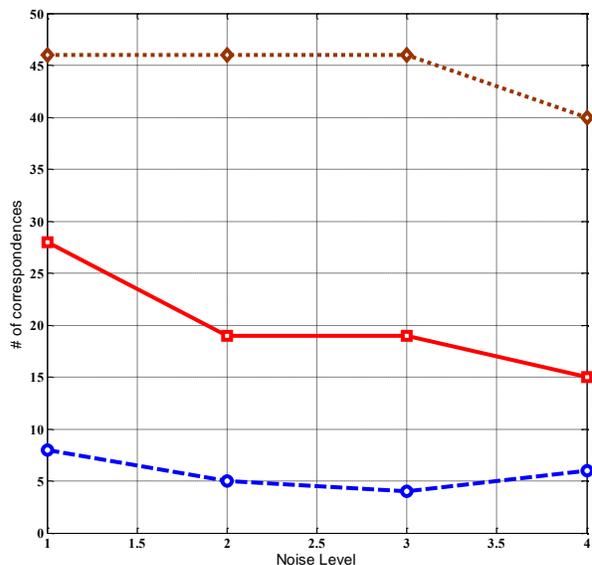
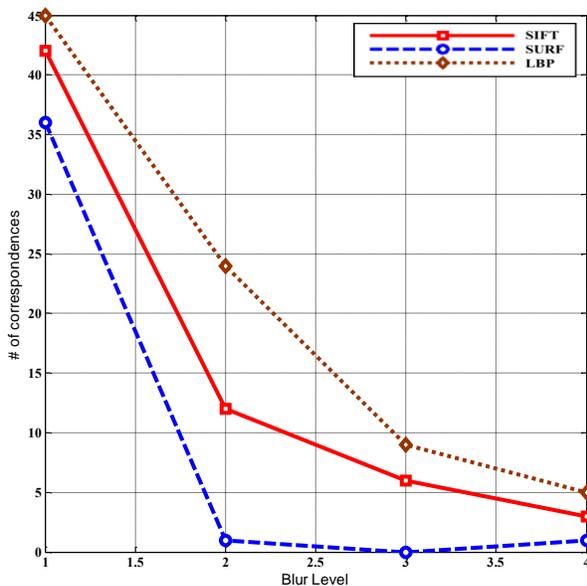


Fig. 7: The plot of the LBP descriptor performance on the test image under different blur, noise, rotation and scale levels at the same selected point on transformed images.



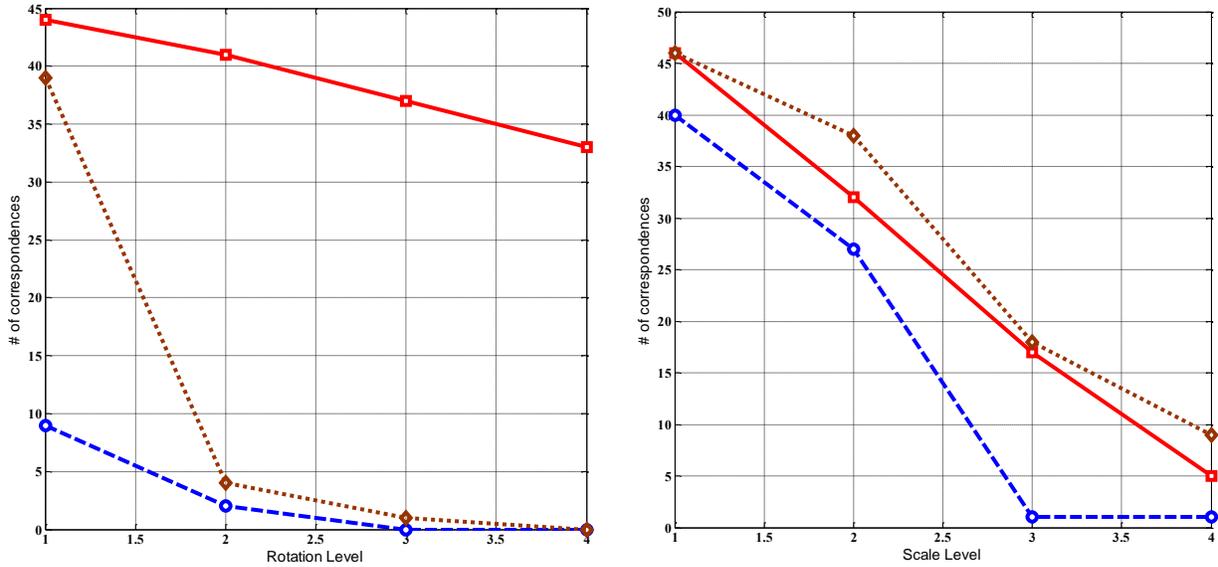
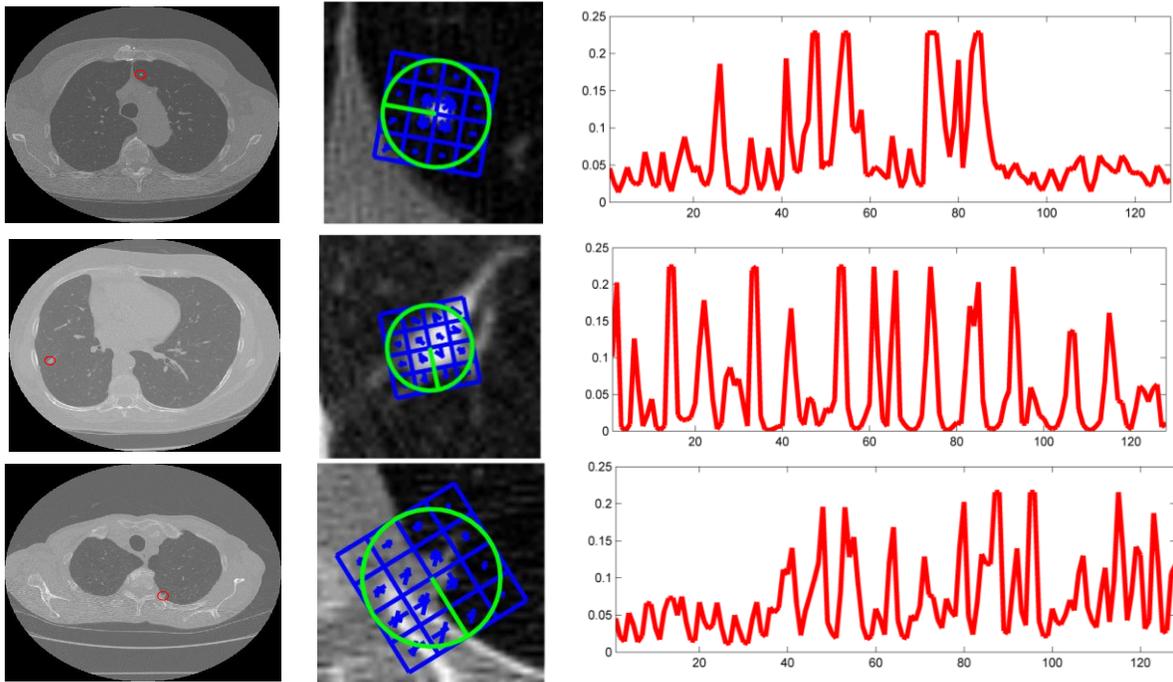


Fig. 8: The number of correspondences under different blur (upper left), noise (upper right), rotation (lower left) and scale levels (lower right) for the SIFT (solid curves), SURF (dashed curves), and LBP (dotted curves) descriptor.

3. Descriptors of small-size lung nodules in chest CT

Farang, et al. 2010 [6] used the SIFT algorithm for classification of small-size nodules that appear in low dose CT (LDCT) scanning of the human chest for early detection of lung cancer were examined. Fig. 9 (left) shows four slices containing nodules < 1 cm in size. These nodules are magnified in Fig. 9 (middle) to show the components of the descriptor. The SIFT algorithm was applied to the four nodule types and the resulting descriptors were used to classify the nodules after a detection step, in order to reduce false positives. Small-size nodules lack textural distinction, but the shapes are distinct. Fig. 9 (right) shows the construction and values of the SIFT algorithm for the four nodule types. The values of the SIFT descriptor shows decent discrimination among the nodules.



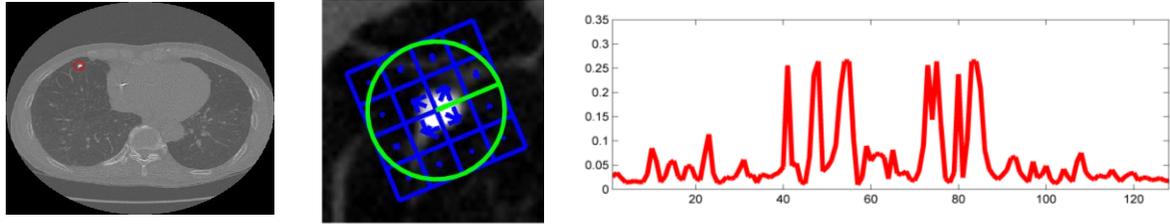


Fig.9: SIFT descriptor applied to small-size nodule types in LDCT of the human chest. From top to bottom: well-circumscribed, vascular, juxta-pleural and pleural-tails nodule types.

As an example of the performance of the LBP descriptor, Fig. 10 illustrates its formation on a juxta-pleural nodule in LDCT scans.

4. Feature-based Nodule classification

The classification stage using various feature descriptors was analyzed in the author's work (e.g., [7][8]). The most significant classification results were obtained when the shape based signed distance transform was combined to the texture based LBP approach. The results in Tables 1-3 illustrate the classification results of the signed distance transform versus the multi-resolution local binary pattern (LBP). A third feature descriptor using the combination of the methods is also shown. Higher true-positive rates can be seen from the LDA projection in tables 1 and 2 when more training is conducted using either the LBP or distance descriptors separately. When comparing the PCA results less training data resulted in better true-positive classification of nodules. In the non-nodule distance transform experimentations more training data was needed to obtain in some instances perfect results. This is understandable since the non-nodules do not have specific shape characteristics that can be defined or manipulated as in the nodules case. Overall, the PCA combinational shape and feature description of nodules improved true-positive rate increase in classification.

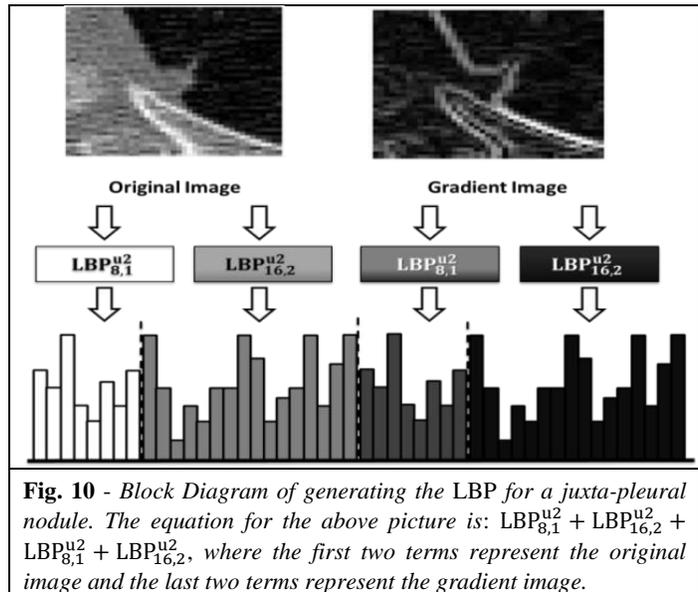


TABLE 1 - Classification Results for various nodules using Raw LBP, LDA LBP and PCA LBP with variable training percentages.

Nodule Type	Raw LBP				LDA LBP				PCA LBP			
	100%	75%	50%	25%	100%	75%	50%	25%	100%	75%	50%	25%
Juxta Pleural	52	50	47	38	100	86	65	50	64	64	59	67
Well-Circumscribed	40	41	40	26	65	80	63	36	64	60	66	82
Vascular	22	29	32	10	32	76	56	32	20	22	37	56
Pleural Tail	22	20	17	11	100	76	52	39	33	17	33	46
Non Nodule	78	77	74	68	100	88	60	44	86	87	83	96

TABLE 2 - Classification Results for various nodules using Raw Distance Transform, LDA LBP and PCA Distance Transform with variable training percentages.

Nodule Type	Raw Distance Transform				LDA Distance Transform				PCA Distance Transform			
	100%	75%	50%	25%	100%	75%	50%	25%	100%	75%	50%	25%
Juxta Pleural	38	39	35	34	100	88	61	45	62	54	60	68
Well-Circumscribed	33	33	36	34	74	83	63	45	46	59	48	55
Vascular	12	12	15	15	29	76	54	29	37	22	61	63
Pleural Tail	17	17	17	15	100	85	54	33	17	24	35	52
Non Nodule	63	68	68	49	100	87	65	49	83	89	85	79

Overall, the PCA combinational shape and feature description of nodules resulted in a drastic true-positive rate increase in classification. All of the results depicted in Tables 1 and 2 allow the conclusion to be made that non-nodules do in-fact contain descriptor variations that allow them to be correctly classified. Also, combination of shape and texture feature information allows for better object representation to be obtained, thus improved results in classification. Table 3 depicts impressive results when the LBP was obtained from the distance transform images. A 20% true-positive rate increase was found, in the PCA 25% training combinational vascular nodule case when comparing it to the PCA LBP results obtained when only the texture information was used for classification, and a 13% increase over the distance transform results alone. Variations of percentage increases were seen for each nodule category.

TABLE 3- Classification Results obtained from Raw Combinational Feature Transform and PCA Combinational Feature Transform with variable training percentages.

Nodule Type	Raw Combinational Feature Descriptor				PCA on Combinational Feature Descriptor			
	100%	75%	50%	25%	100%	75%	50%	25%
Juxta Pleural	40	41	39	37	78	76	76	79
Well-Circumscribed	40	37	36	34	73	68	71	68
Vascular	24	20	22	12	51	54	44	76
Pleural Tail	22	26	22	20	33	35	41	54
Non Nodule	63	57	58	49	100	99	100	98

5. Summary

This paper studied feature definition and descriptors as basic ingredients for object modeling. Many of these concepts are the ingredients of modern approaches of segmentation and registration. Biomedical objects rarely have defined topology, therefore, robustness of various image analysis methods depends on accurate feature detection and representation (description). The computer vision literature has various surveys and algorithms on the subject of feature extraction and description. Unfortunately, in small-size random objects, one cannot evaluate the accuracy of the approach with confidence, as the objects' features are not well-defined in the first place.

Acknowledgments: Portion of this work was conducted while the author was with the CVIP Lab, University of Louisville. Special thanks are to Mr. Mostafa Abdelrahman, Ahmed Shalaby and Dr. Aly Farag.

6. References

- 1 David G. Lowe (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91-110.
- 2 Cordelia Schmid and Roger Mohr and Christian Bauckhage (2000), Evaluation of Interest Point Detectors. *International Journal of Computer Vision*, 37(2):151-172.
- 3 Krystian Mikolajczyk and Cordelia Schmid (2005). A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, 27(10):1615-1630.
- 4 H. Bay, A. Ess, T. Tuytelaars and L. Van Gool, (2008). Speeded-Up Robust Features (SURF), *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359.
- 5 T. Ojala, M. Pietikainen, and T. Maenpaa, (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (24): 971-987.
- 6 Amal Farag, Shireen Elhabian, James Graham, Aly Farag, and Robert Falk (2010). Toward Precise Pulmonary Nodule Descriptors for Nodule Type Classification. *Proc. of the 13th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pp. 626-633.
- 7 Amal Farag, Asem Ali, James Graham, Aly Farag, Salwa Elshazly and Robert Falk, "Evaluation of geometric feature descriptors for detection and classification of lung nodules in low dose ct scans of the chest," *International Symposium on Biomedical Imaging (ISBI-11)*, pp. 169 - 172, March 2011.
- 8 Amal Farag (2012). Modeling Small Size Objects under Uncertainty: Novel Algorithms and Applications. PhD Dissertation, CVIP Lab., University of Louisville, May 2012.