

Appearance-based Salient Features Extraction in Medical Images Using Sparse Contourlet-based Representation

Rami Zewail¹, Ahmed Hag-ElSafi²

University of Alberta, Edmonton, Alberta, Canada

Smart Empower Innovation Labs Inc., Edmonton, Alberta, Canada

Email: rami.zewail@smartempower.ca¹, ahmed.hagelsafi@smartempower.ca²

Received: 12 May 2017; Accepted: 05 July 2017; Published: 08 September 2017

Abstract—Medical experts often examine hundreds of x-ray images searching for salient features that are used to detect pathological abnormalities. Inspired by our understanding of the human visual system, automated salient features detection methods have drawn much attention in the medical imaging research community. However, despite the efforts, detecting robust and stable salient features in medical images continues to constitute a challenging task. This is mainly attributed to the complexity of the anatomical structures of interest which usually undergo a wide range of rigid and non-rigid variations.

In this paper, we present a novel appearance-based salient feature extraction and matching method based on sparse Contourlet-based representation. The multi-scale and directional capabilities of the Contourlets is utilized to extract salient points that are robust to noise, rigid and non-rigid deformations. Moreover, we also include prior knowledge about local appearance of the salient points of the structure of interest. This allows for extraction of robust stable salient points that are most relevant to the anatomical structure of interest.

Index Terms—Salient features, multiscale, appearance, sparse, contourlet

I. INTRODUCTION

Within the field of medical imaging, salient feature extraction has drawn much attention in a wide range of applications. Examples include: computer-aided diagnosis, segmentation, registration, and image retrieval. However, detecting robust and stable salient features in medical images continues to constitute a challenging task. This is mainly attributed to the high resolution of medical images, and complexity of the anatomical structures of interest which usually undergo a wide range of rigid and non-rigid variations.

In the era of Big Data and mobile-based healthcare, these challenges have even escalated where there is a growing demand for analyzing of high resolution medical images and large scale biomedical data in general. Over

the last decade, concepts of sparsity of representation and its applications in computer vision has been gaining an increased interest from the research community. In the medical imaging paradigm, sparsity techniques have been used in applications such as image enhancement, segmentation, quantification of diseases. In this paper, we present a novel appearance-based salient feature extraction and matching method based upon a sparse Contourlet-based representation. The multi-scale and directional capabilities of the Contourlets is utilized to extract salient points that are robust to noise, rigid and non-rigid deformations. Moreover, we also include prior knowledge about local appearance of the salient points of the object of interest. This allows for detection a reduced set of robust and stable salient points that is most relevant to the structure of interest.

The rest of the paper is organized as follows: Section II covers some of the related work in literature. Section III presents the details of the proposed Appearance-based salient feature extraction method. Experiments and results are presented in section IV. Finally, conclusions are drawn in section V.

II. RELATED WORK

Inspired by the Human Visual System, salient (interest) point detection methods have been used in many computer vision applications such as registration, tracking, robot navigation, and recently in medical imaging. Salient points need to be well defined and stable under various transformations. Various salient point detection algorithms are often evaluated in terms of its Repeatability; that is the ability of the method to detect the same salient points under various transformations such as: scaling, rotation, and change in view.

Several methods have been proposed in the literature for salient point detection. In [1], Schmidt et. al. classified salient detection methods into three classes: edge-based, intensity-based, and biologically-inspired methods. The Harris corned detector, [2], is among the most famous methods in the literature. It detects corners in an image based upon the autocorrelation matrix.

Despite being useful in simple images, the Harris detector suffers from the following limitations: it is based on single scale analysis, and has poor localization ability. Another famous salient point detection method is the SIFT algorithm in [3]. The SIFT algorithm extracts salient points by finding the local maxima of the response of the Difference-of-Gaussian (DOG) operator. Biologically-inspired salient point detectors have focused onto directional and multi-scale analysis of the images. Q. Tian [4] and Loupais et. al. [5] proposed wavelet-based salient point detectors. The wavelet transform has several appealing properties such as the multi-scale representation and the spatial-frequency localization. Nevertheless, wavelets do suffer from lack of shift invariance and limited orientation selectivity. These are both important properties for efficient detection of salient points in medical images with complex anatomical structures. In [6], Itti and Koch proposed a biologically inspired bottom-up saliency detector based upon Gabor filters. In [7], the authors presented a salient point detector based upon the phase congruency of Log-Gabor filter response. In [8], Xinting Gao et. al. proposed a multi-scale corner detection method based upon Log-Gabor transform.

In [11], Perazzi et al. used contrast-based filtering for detection of salient regions. In [12], the authors used Harris method to build a saliency map that is used in segmentation. In [13], Dual tree complex wavelet transform is used in a content-based image retrieval application. In [14], the authors used a deep network architecture to build a linear invariant feature transform (LIFT). In [15], Mark Brow et al. presented a generalized framework for saliency feature detection using histogram.

In [17], the authors proposed a method for salient point detection based on steerable filters and Harris point detector. In [18], Omprakash et. al. proposed a method for fast multi-scale visual saliency detection using scale space interpolation.

III. METHODOLOGY

We next present the details of the proposed salient feature extraction and matching method. The new method is suited for extraction of robust salient features of interest in medical images. The proposed method attempts to incorporate the following characteristics:

- Biologically inspired: Since there no strict definition of what constitutes a salient point in an image, it is logical to use biologically-inspired filters in the detection process.
- Localized: The localization property of the detector is also an important aspect that tends to affect other higher-level tasks such as feature extraction, matching, and registration.
- Able to detect higher order salient structures in images (e.g. contours).
- Incorporate prior knowledge about the pathology of interest: In most medical image analysis applications, we are interested in a structure; hence

it is beneficial for the salient point detector to be target driven. Thus, it is desirable to integrate top-level prior knowledge within the salient feature extraction process.

The method is based upon Non Sub-sampled Contourlet Transform (NSCT). The multi-scale and directional capabilities of the NSCT transform is utilized to extract salient points that are robust to noise, rigid and nonrigid deformations. We also include prior knowledge about the local appearance of the salient points of the pathology of interest. This enables us to extract robust and stable salient points that are most relevant to the pathology of interest.

First, salient features are extracted in a bottom-up approach using Non-subsampled Contourlet-based representation. Next, local appearance profile for the pathology of interest is built using Gaussian Mixture Model (GMM). The details of these two steps are presented in sections III.A and III.B consequently.

A. Contourlet-based Bottom-up Salient Feature Extraction

The bottom-up multi-scale salient feature extraction step is achieved using Non-subsampled Contourlet representation(NSCT) and second moment matrix. Using the ability of the NSCT transform to capture higher order image structures, we are able to extract robust salient features in the medical images. We will refer to the new method as: NSCT-SMM. Fig.1 shows a block diagram for the proposed method.

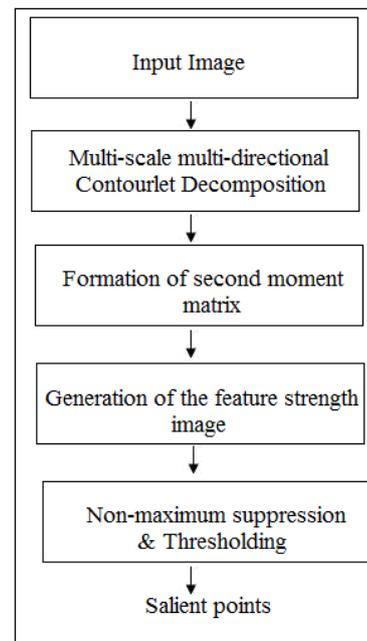


Fig.1. Flowchart of the proposed Contourlet-based bottom-up salient feature extraction step (NSCT-SMM)

A multi-scale directional decomposition is achieved using the Non Sub-sampled Contourlet transform. The input image $I(x, y)$ is decomposed into 4 scales. In the first scale level, we use a 4-level Directional Filter Bank

(DFB) to get 16 directional frequency partitioning. In the second and third scales, eight directional frequency partitions are obtained using a 3-level Directional Filter bank. In the final scale level, 2-level directional filter-banks are used. Let $\{B_j\}$, $j = 0, 1, 2, 3$ denotes the set of response images that results from the non sub-sampled pyramid decomposition step, where j is the scale level. At each scale j , the response image B_j is further decomposed using l_j -level non sub-sampled Directional Filter Bank to produce a set of multi-scale multi-directional contourlet sub-band images $\{W_{j,k}\}$, where $k = 0, 1, \dots, 2^{l_j-1}$, and $j = 0, 1, 2, 3$.

Fig.2 shows the oriented contourlet basis functions for two different scales. Due to the flexibility of the transform, we can decompose the image using different angular resolutions at every scale.

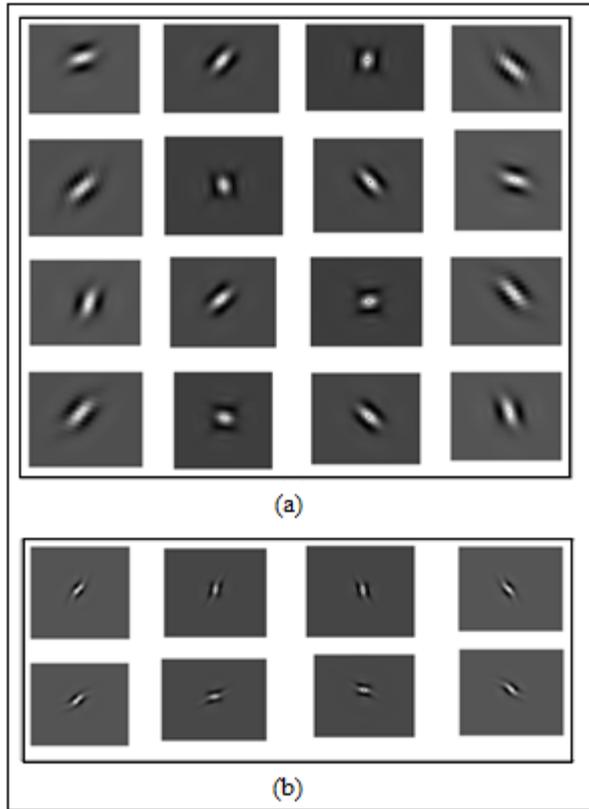


Fig.2. NSCT basis functions for two scales with different angular resolutions at each scale/(a) 16 directional basis functions at scale1, (b) 8 directional basis functions at scale 2.

Next, following classical moment analysis equations, [9,16], we compute the following:

$$a = \sum_{j=0}^M \sum_{k=0}^{N_j} (|W_{j,k}| \cos(\theta_k))^2 \quad (1)$$

$$b = \sum_{j=0}^M \sum_{k=0}^{N_j} |W_{j,k}| \cos(\theta_k) \sin(\theta_k) \quad (2)$$

$$c = \sum_{j=0}^M \sum_{k=0}^{N_j} (|W_{j,k}| \sin(\theta_k))^2 \quad (3)$$

Where j is the scale of the decomposition, M is the maximum number of scales, k is the orientation at scale j , N_j is the number of decomposition orientations at scale j , $|W_{j,k}|$ is the magnitude of the non-subsampled contourlet coefficients at scale j and orientation k , and $\theta_{j,k} = \frac{(k-1)\pi}{N_j}$ is the angle of the k^{th} orientation at scale level j .

The second moment matrix is then constructed as:

$$M = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad (4)$$

The eigenvalues of the second moment matrix in equation 4 is then calculated as:

$$E_1 = \frac{1}{2}[a+c] - \frac{1}{2}\sqrt{4b^2 + (a-c)^2} \quad (5)$$

$$E_2 = \frac{1}{2}[a+c] + \frac{1}{2}\sqrt{4b^2 + (a-c)^2} \quad (6)$$

Next, we define a contourlet-based feature strength image, F , as a weighted sum of the two eigenvalues of the second moment matrix. The smaller eigenvalue, E_1 , is a measure of the corners in the image and the larger eigenvalue, E_2 , is a measure of the contours and edges in the image [9]. Fig.3 shows an example of x-ray image of human spine and the corresponding Contourlet Feature strength map.

$$F = \alpha E_1 + (1-\alpha)E_2 \quad (7)$$

Salient features are finally extracted from the Contourlet-based feature image through non-maximum suppression and thresholding. This step selects local maxima points in the Feature Image as salient points.

B. Incorporation of prior pathology-related knowledge

Finally, we incorporate prior knowledge about the structure of interest to bias the bottom-up salient point detection procedure detailed in III.A. This allows us to extract a more stable and robust set of salient features that are most relevant to the pathology of interest. The new method benefits from the rich multi-scale directional information offered by the Non-Subsampled Contourlet decomposition to construct localized descriptors for salient points of the structure of interest. The local descriptors are then modeled using Mixture of Gaussians

to construct a local appearance model of the target salient points.

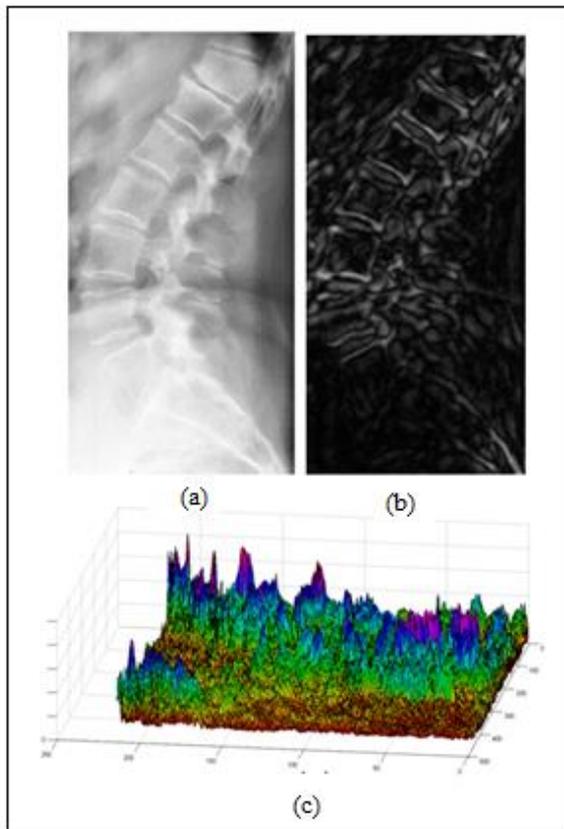


Fig.3. Example x-ray image and corresponding Contourlet-based Feature Strength Image. (a) original x-ray image, (b) contourlet-based feature strength image, (c) 3D view of the contourlet-based feature image.

Given a set of M training images with a ground truth segmentation for the structure of interest, we extract a set of N ground truth salient points. First, we use the method described in section III.A to extract salient features in the image. Next, the ground truth salient points that correspond to the structure of interest are selected using the ground truth segmentation. A ground truth salient point is defined as: a salient point that has a point-to-curve distance less than 5 pixels. Fig.4 shows two x-ray images with the ground truth salient points that define the border of the lumbar vertebrae L1-L5.

For every ground truth, salient point (P_i), we extract a local descriptor based upon Contourlet-based representation. We use a square grid of $N \times N$ ($N=7$) points around each ground truth salient point. Variance-based features are then extracted from the NSCT directional

sub-bands at every scale, excluding the low-pass response. The average absolute deviation of the gray values of the square grid for each sub-band is calculated as follows:

$$E_j^i(a,b) = \frac{1}{N^2} \sum_{l=a-\frac{N}{2}}^{a+\frac{N}{2}} \sum_{k=b-\frac{N}{2}}^{b+\frac{N}{2}} |x_j^i(l,k) - \mu_j^i| \quad (8)$$

Where suppose that x_j denotes the j^{th} NSCT-based Feature image. $x_j^i(l,k)$ is the NSCT coefficient of the x_j NSCT Feature image located in (l,k) . μ_j^i is the mean of the coefficients in the $N \times N$ square grid whose center is the characteristic salient point P_i . The contourlet-based local descriptor for salient point P_i is then defined as:

$$F_i = [E_1, E_2, E_3, \dots, E_D]^T \quad (9)$$

Where E_i is the average absolute deviation in a square grid for the i^{th} NSCT sub-band. D is the total number of NSCT sub-bands.

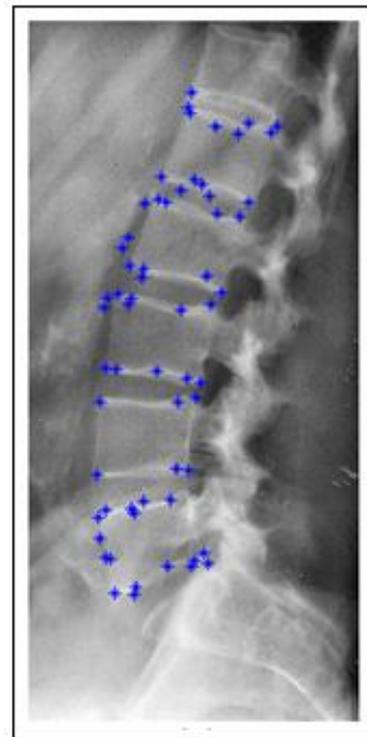


Fig.4. Example of ground truth salient points that defines the human lumbar spine in x-ray images.

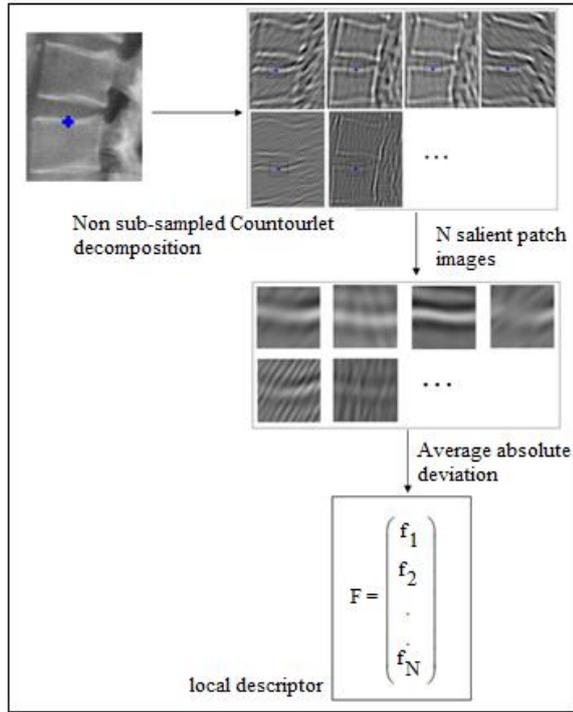


Fig.5. Extraction of contourlet-based local descriptor for each ground truth salient point

The outcome of this stage is a set of $K=N \times M$ contourlet-based feature vectors extracted from the ground truth salient points in the training set. Fig.5 summarizes the procedure for construction of the contourlet-based local descriptor.

We then use the matrix of local descriptors to construct a local appearance model for the structure of interest using Mixture of Gaussians.

Let $\{F_i\}_{i=1..K}$ be the set of contourlet-based local descriptors for the ground truth salient points in the training set. K is the total number of ground truth salient points, and F_i is the local descriptor of the i^{th} salient point. We model the local descriptors of the ground truth salient points in the training set as a mixture of N_m Gaussians. The Gaussian Mixture model is a flexible parametric distribution [10].

The Gaussian mixture model for the local descriptors is given by:

$$p(F_k / \theta) = \sum_{n=1}^{N_m} \alpha_n f_n(F_k) \quad (10)$$

Where N_m : Number of Gaussian functions in the model.

$f_n(F_k)$: Multivariate Gaussian density functions characterized by the mean (μ_n) and covariance matrix (Σ_n).

$$f_n(F) = \frac{1}{(2\pi)^D (\Sigma_n)^{0.5}} \exp\left\{-\frac{1}{2}(F - \mu_n)^T (\Sigma_n)^{-1} (F - \mu_n)\right\} \quad (11)$$

α_n : Mixture weights, where $\sum_{n=1}^{N_m} \alpha_n = 1$

The parameters of the Gaussian mixture model are as follow:

$$\theta = \{\alpha_n, \mu_n, \Sigma_n\}, n = 1..N_m \quad (12)$$

Using the training set $\{F_k\}_{k=1..K}$, we estimate the parameters of the Gaussian mixture model (GMM) by maximizing the likelihood function. The GMM likelihood function is given by:

$$p(\{F_k\}_{k=1..K} / \theta) = \prod_{k=1}^K p(F_k / \theta) \quad (13)$$

We use the Expectation-Maximization algorithm (EM-algorithm) to estimate the model parameters that maximize the likelihood function given in equation 13.

$$\hat{\theta} = \arg \max_{\theta} \{p(\{F_k\}_{k=1..K} / \theta)\} \quad (14)$$

The prior local appearance model is then used to bias the outcome of the bottom-up salient feature extraction step described in section III.A. For every detected low-level salient point, we extract a contourlet-based local descriptor in the same manner as explained in Fig.5. Next, we estimate the confidence level that the detected salient point belongs to the model learnt during the training phase. To do so, following Bayes decision rule [10], we use the posterior probability, as a measure for the confidence level.

IV. EXPERIMENTS AND RESULTS

We evaluate the performance of the proposed salient feature extraction method using various test images. Fig.6 shows examples of the images used in the experiments. Fig.6(a) and Fig.6(b) are commonly used in the literature to evaluate salient point detection algorithms. Fig.6(b) is an example of the x-ray images of the human spine in the evaluation process. The synthetic image in Fig.6(a) exhibits different boundaries with varying strength levels. It is used to evaluate the localization capabilities of the new method. The ‘‘Van Gogh’’ image shown in Fig.6(b) is a famous painting that contains a lot of texture. It is used to evaluate the robustness of the proposed method against various geometric and photographic transformations; these include: rotation, illumination, and noise. Fig.6(c) shows example of x-ray image of human spine that were used to test the performance of the proposed method when applied to medical images.

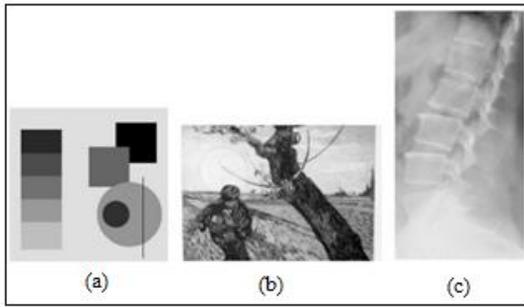


Fig.6. Images used for evaluation of NSCT-based salient point detector. (a) Synthetic image, (b) VanGogh Image, (c) X-ray image of human spine

We evaluate the robustness of the proposed salient point detector using the Repeatability Rate criterion (rr) introduced by Schmid et. al. [1]. Given two images that are related together by some transformation, the repeatability rate is the ratio between the number of corresponding points that are detected in both images to the total number of points in the common area. In our experiments, we consider the following relationships between the reference and test images:

- Rotation transformation.
- Thin Plate Spline transformation (as an example of non-rigid transformation).
- Variations in illumination levels.
- Variations in noise levels.

Let $\phi^{(r)} = \{X_i^{(r)}\}$, and $\phi^{(t)} = \{X_j^{(t)}\}$ be two sets of interest points detected in the reference and test images respectively. Let N_c be the number of salient points in the common area. Let $C\{X_i^{(r)}, X_j^{(t)}\}$ be the number of corresponding pairs of salient points; with a localization error of 1.5 pixels. Then, the Repeatability rate (rr) between the two images is defined by:

$$rr = \frac{C\{X_i^{(r)}, X_j^{(t)}\}}{N_c} \quad (15)$$

We evaluate the proposed method (*NSCT-based*) against the following famous salient point detectors from the literature:

1. Harris detector (*Harris*).
2. The detector proposed by Kovesei et. al. in [9] that is based upon Log-Gabor decomposition and phase congruency (*PC-based*).
3. A salient point detector that is based upon steerable pyramid decomposition instead of Non sub-sampled contourlet decomposition (*Steerable-based*).

A. Localization

The localization capabilities of the proposed method are evaluated using the synthetic test image shown in Fig.6 (a). Fig.7 shows salient points extracted using various salient point detectors. The NSCT-based detector

is able to reveal most of the salient points in the image. These include corners, strong and weak edges, and smooth contours. On the other side, the Harris detector and the Laplacian-of-Gaussian (LOG) detector seem to perform good only in case of strong salient locations (corners and edges), while both detectors show limited localization capabilities in case of weak edges and smooth contours. The phase congruency method (PC-based) demonstrates better performance compared to the Harris and LOG detectors. However, the proposed NSCT-based detector appears to be superior in localizing salient points along the smooth circular contour. This is attributed to the fact that Non-subsampled Contourlet Transform (NSCT) offers more flexible directional analysis compared to Log-Gabor filters. Similar observations are obtained when the salient detectors were applied to the x-ray image of human spine in Fig.8. Among all detectors investigated, the proposed detector seems to better detect anatomical structure with poor edges and smooth contours.

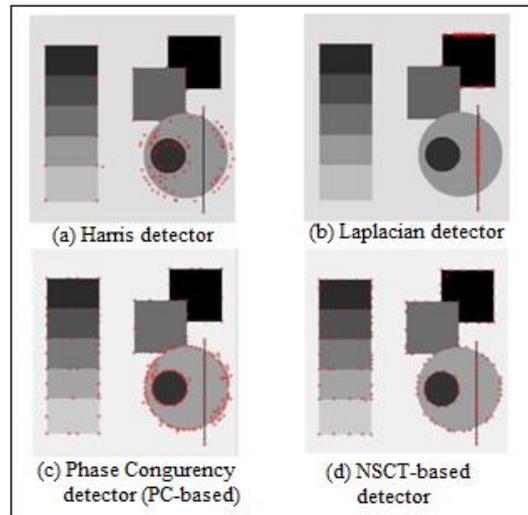


Fig.7. Localization of various salient point detectors. (a) Harris detector, (b) Laplacian detector, (c) Phase Congruency detector (PC-based), and (d) NSCT-based detector.

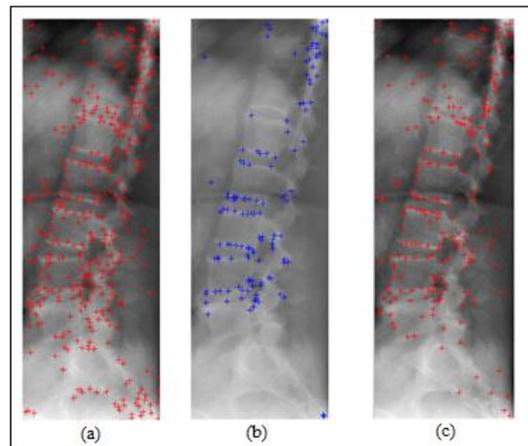


Fig.8. Localization of various salient point detectors in medical images. (a) proposed NSCT-based detector, (b) Harris detector, and (c) Phase Congruency detector.

B. Robustness to rotation

Image rotations are obtained by rotating the image around its central axis. Test images for evaluation of the interest point detectors can be found in [10]. We use a set of rotated VanGogh images. The repeatability rate is calculated as a function of the rotating angle 10-180. Fig.9 shows the performance of the proposed salient point detector compared to several famous salient point detectors including Harris detector, Cottier detector, and Heitger detector [1]. The Harris detector and Cottier detector are based upon auto-correlation matrix. Heitger detector is based upon Gabor-like filters. The proposed NSCT-based detector appears to outperform all three detectors at various rotation angles.

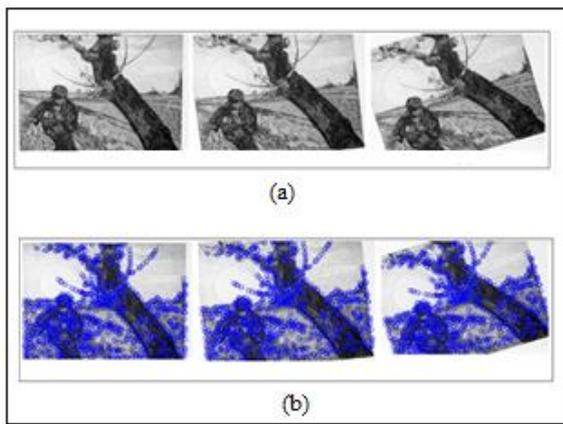


Fig.9. Performance of the proposed salient feature extraction method under various rotations (a) VanGogh test image at different rotations, (b) Corresponding salient features extracted at each rotation position.

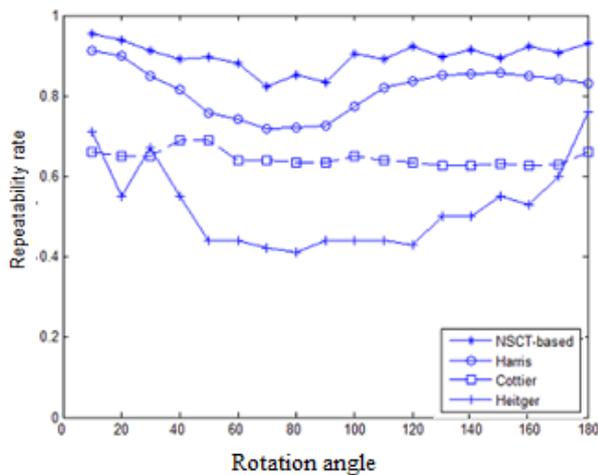


Fig.10. Robustness to rotation of various salient point detectors

C. Robustness to noise

The robustness to noise is evaluated using the VanGogh test image with added Gaussian noise. Fig.11 shows the effect of the increased noise level on the repeatability rate of the salient points for both the proposed NSCT-based method, and Harris detector. From Fig.11, we observe the following: an average improvement in the repeatability rate of 15.23 % is

achieved for noise variance levels below 0.02. At higher noise levels, the proposed method shows much more robustness to increased noise levels with an average improvement in the repeatability rate of 70.40 %.

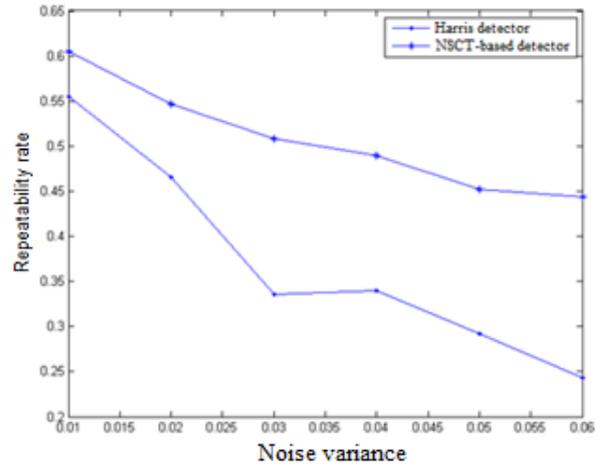


Fig.11. Effect of variations in noise levels on the repeatability rate for various salient point detectors

D. Robustness to change in illumination

The robustness of the proposed method to uniform change in illumination is evaluated using the VanGogh test image with changing levels of brightness. This is done by changing the illumination factor (β). Fig.12 shows how the repeatability rate is affected by altering the brightness level of an image a positive value of β increases the brightness level, and a negative value of β decreases the brightness level. Compared to Harris detector, the proposed NSCT-based detector has achieved an average improvement in repeatability rate of 48.89%. Compared to the Phase congruency-based detector (PC-based), [9], our method achieves an improvement of 13.124 % in the repeatability rate.

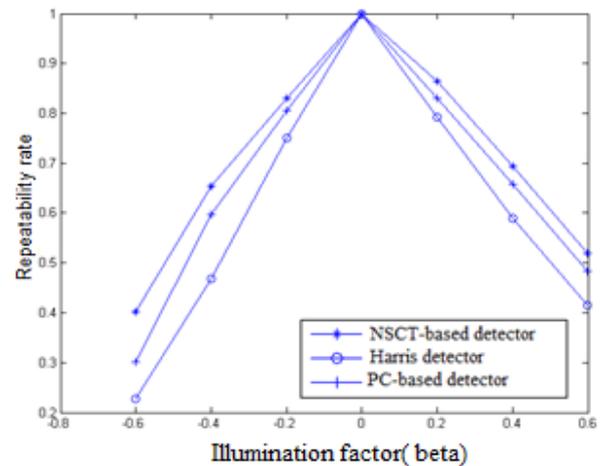


Fig.12. Effect of brightness level on the repeatability rate for various salient point detectors

E. Robustness to non-rigid transformations

In medical images, anatomical structures undergo a range of non-rigid deformations. Hence, a good salient point detector for medical images needs to be robust to non-rigid deformations in the object of interest. We next evaluate the robustness of the new method to non-rigid deformations. Experiments are conducted using a set of x-ray images of lumbar spine with different non-rigid deformations. The structure of interest in every image pair is related by a non-rigid transformation estimated using the Thin Plate Spline model. For each method, we calculate the repeatability rate as the percentage of correct matching to the total number of common points between the two images. Table 1 compares the achieved average repeatability rate for the proposed NSCT-based detector, Harris detector, PC-based detector, and the steerable-based detector.

The proposed NSCT-based detector achieves improvements of 64.224 %, 18.0 % in the average repeatability rate compared to Harris detector, and PC-based detector respectively. The NSCT-based detector achieves an improvement of 9.95% in the average repeatability rate compared to the Steerable-based detector.

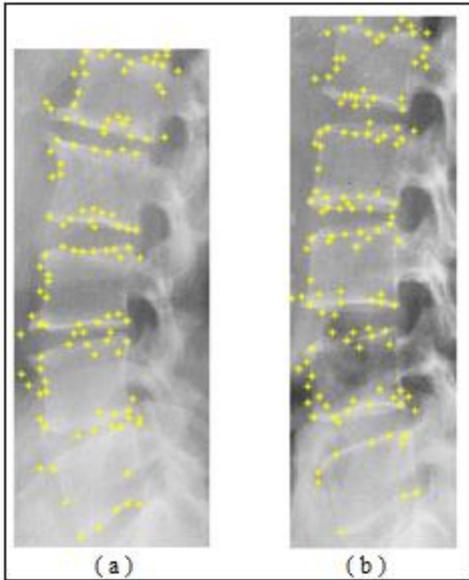


Fig.13. Example of two x-ray images where the object of interest is related by non-rigid deformations

Table 1. Evaluation of robustness of different salient point detectors to non-rigid deformations of the object of interest

Salient Point detector method	Average repeatability rate
Proposed method NSCT-based	0.7891
Steerable-based	0.7177
Harris detector	0.4805
PC-based detector [9]	0.6687

F. Appearance-based prior knowledge

Fig.14 demonstrates the benefit of incorporating prior pathology-related knowledge with the bottom-up salient feature extraction process. By applying the pathology-related prior knowledge, an average reduction ratio of 51.86% is achieved in the extracted salient features. This allows for extraction of features that are robust and most relevant to the pathology of interest.

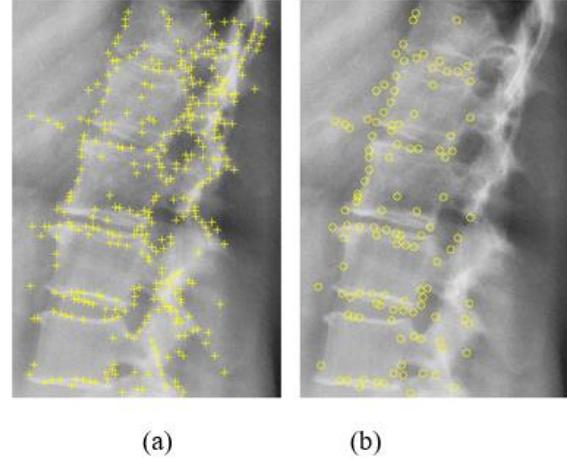


Fig.14. Example for salient point detection using NSCT-based method (a) with no top-level knowledge, (b) with top level appearance knowledge.

G. Performance in medical image retrieval applications

Finally, we evaluate the suitability of the proposed appearance-based salient feature extraction method in medical image retrieval applications. We use images from the IRMA database which is made publicly available as a part of the Image Retrieval for Medical Applications (IRMA) project. The database consists of radiograph images that were randomly collected from regular routines at the department of diagnostic Radiology, Aachen University of Technology (RWTH), Aachen, Germany. In our experiments, we use a subset of 1000 images, where 800 images used in training and 200 images left out for testing as query images.

We evaluate the performance of the new method using the following criteria:

- *Average retrieval rate (ARR)*: For a query image, the retrieval rate is the ratio between the number of relevant images retrieved to the total number of retrieved images. The Average retrieval rate, ARR , is calculated by averaging the retrieval rates over all query images.

- *Average Error Rate (ER)*: This is the percentage that the first retrieved image is not a relevant image averaged over all query images.

Fig.15 evaluates the performance of the proposed system using average retrieval rates (ARR) as a function of the number of top retrieved images. We compare the performance of the following systems:

- i. The proposed retrieval system using Contourlet-based salient point detectors (NSCT-local).

- ii. Retrieval system based upon Gabor-based salient detectors (Gabor-local).
- iii. Retrieval system based upon global features extracted from the response of the NSCT sub-bands (NSCT-global).

The proposed retrieval method (NSCT-local) tends to consistently outperform the other two methods. Over a range of 5-40 top retrieved images, the proposed method achieves an average improvement in the Average Retrieval Rate (ARR) of 9.574 % and 17.338 % over the Gabor-based system, and the NSCT-global system respectively.

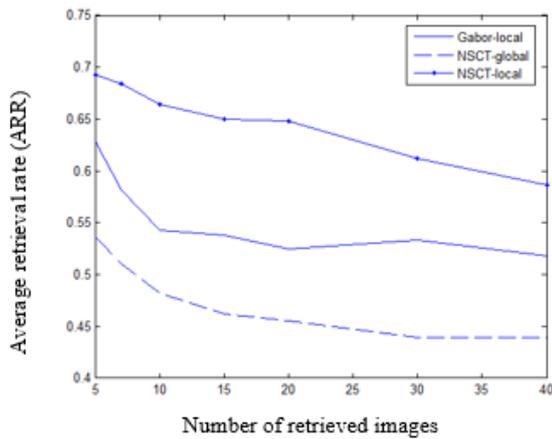


Fig.15. Average retrieval rates versus number of retrieved images for the following image retrieval methods: Gabor-based salient points (Gabor-local), NSCT-based salient points (NSCT-local), and NSCT-based global response (NSCT-global).

Table 2 evaluates the performance of the proposed method in terms of the average error rate (AERR). The proposed method (NSCT-local) achieves an average error rate of 21.43 % versus 35.71 % and 46.43 % for the Gabor-based system, and the NSCT-global system respectively.

Table 2 Average error rates for different content-based image retrieval methods

Method	Average Error Rate
NSCT-local	21.43 %
NSCT-global	46.43 %
Gabor-local	35.71

V. CONCLUSION

In this paper, we presented a novel appearance-based salient feature extraction that can extract robust and stable salient features in medical images. The salient feature extraction process is performed in a sparse contourlet-based domain. The proposed method is based upon Non Sub-sampled Contourlet Transform (NSCT), a truly 2D multi-scale and multidirectional image analysis tool. Rich discriminative information was extracted using local descriptors from the non-subsampled contourlet

sub-bands. The new method also integrates prior local appearance knowledge of the structure of interest in the detection process. Compared to related methods in the literature, our method has proven successful in detecting stable and relevant salient points in medical images. The proposed method opens door for pathology-related machine learning tasks such as image retrieval, automated detection of abnormalities and diseases.

ACKNOWLEDGEMENT

The authors wish to acknowledge the US. National Library of Medicine for making the digital x-ray images available.

REFERENCES

- [1] C. Schmid, R. Mohr and C. Bauckhage, "Evaluation of Interest Point Detectors," International Journal of Computer Vision, vol. 37(2), pp.151-172, 2000.
- [2] C. Harris, M. Stephens, "A combined corner and edge detector," Proceedings of Alvey Vision Conference, 1988.
- [3] D. Lowe, "Object Recognition from Local Scale-Invariant Features," Proceedings of 7th IEEE International Conference on Computer Vision, vol. 2, pp. 1150-1157, 1999.
- [4] Q. Tian, N. Sebe, M. S. Lew, E. Loupias, and T. S. Huang, "Image retrieval using wavelet-based salient points," Journal of Electronic Imaging, vol. 10(4), pp. 835-849, 2001.
- [5] E. Loupias, N. Sebe, "Wavelet-based Salient Points: Applications to Image Retrieval Using Color and Texture Features," Proceedings of 4th Int. Conf. on Visual Information Systems, pp. 223-232, 2000.
- [6] L. Itti, C. Koch and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 20(11), pp. 1254-1259, 1998.
- [7] C. Wu, Q. Wang, "A Novel Approach for Interest Point Detection Based on Phase Congruency," Proceedings of TENCON 2005, pp. 1-6, 2005.
- [8] X. Gao, F. Sattar, R. Venkateswarlu, "Multiscale Corner Detection of Gray Level Images Based on Log-Gabor Wavelet Transform," IEEE Transactions on Circuits Systems and Video Technology, vol. 17(7), pp. 868-875, 2007.
- [9] P. Kovesei, "Phase Congruency Detects Corners and Edges," Australian Pattern Recognition Society Conference: DICTA 2003, pp. 309-318, 2003.
- [10] A. Webb, Statistical Pattern Recognition. 2nd edition, 2002.
- [11] Perazzi, F., Krahenbuhl, P., Pritch, Y., Hornung, A.: Saliency filters: contrast based filtering for salient region detection. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 733-740. IEEE (2012).
- [12] Hangbing Gao, Yunyang Yan, Youdong Zhang, Jingbo Zhou, Suqun Cao, and Jianxun Xue "Automatic Segmentation of Nature Object Using Salient Edge Points Based Active Contour," Mathematical Problems in Engineering, Volume 2015 (2015), Article ID 174709.
- [13] Stella Vetova, Ivan Ivanov, Content-based Image Retrieval Using the Dual-Tree Complex Wavelet Transform," Proceedings of MCSI '14 Proceedings of the 2014 International Conference on Mathematics and Computers in Sciences and in Industry, pp. 165-170.

- [14] K. M. Yi, E. Trulls, V. Lepetit and P. Fua, "LIFT: Learned Invariant Feature Transform," European Conference on Computer Vision (ECCV) 2016.
- [15] Mark Brown, David Windridge, Jean-Yves Guillemaut, "A generalized framework for saliency-based point feature detection," Computer Vision and Image Understanding, Volume 157, April 2017, pp. 117–137.
- [16] M. Nixon, A. Aguado. Feature Extraction and Image Processing. 2nd edition, Elsevier, 2008.
- [17] Mahesh, Subramanyam M. V, "Feature Based Image Mosaic Using Steerable Filters and Harris Corner Detector", IJIGSP, vol.5, no.6, pp.9-15, 2013.DOI: 10.5815/ijigsp.2013.06.02.
- [18] Omprakash S. Rajankar, Uttam D. Kolekar, "Scale Space Reduction with Interpolation to Speed up Visual Saliency Detection", IJIGSP, vol.7, no.8, pp.58-65, 2015.DOI: 10.5815/ijigsp.2015.08.07

Authors' Profiles



Ahmed S. Hag-ElSafi (Khartoum,1978) holds a BSc and a MSc of Electronics and Communications Engineering, Arab Academy for Science and Technology, 2002 and 2004 respectively.

He has 15 years of experience of research and industrial development in the areas of embedded systems and machine learning. He has published more than fifteen papers in the areas of IOT security, biometrics, machine

learning and predictive analysis, and medical image processing. He is currently the Staff Researcher-Machine Learning at Smart Empower innovation labs Inc., Alberta, Canada.

Mr. Hag-ElSafi is a member the Smart City Alliance in Alberta, Canada and the Association of Professional Engineers and Geoscientists of Alberta (APEGA).



Rami Zewail received a BSc. and MSc. In Electronics & Communications Engineering, Arab Academy for Science and Technology, Egypt, in 2002 and 2004 respectively. And a PhD degree in Electrical and Computer Engineering, University of Alberta, Canada, in 2010.

He has over 15 years of academic and industrial R&D experience in areas of machine learning, digital signal processing, and embedded computing. He has contributed to the scientific community with over 15 publications in areas of image processing, statistical modeling, and embedded computing. Currently, he is the Principal Researcher for Machine Learning at Smart Empower Innovations Labs Inc., Edmonton, Alberta, Canada.

Dr. Zewail is a member of the Institute of Electrical and Electronics Engineers (IEEE) and the Association of Professional Engineers & Geoscientists (APEGA). He served as a reviewer for the Journal of Electronics Imaging, and Journal of Optical Engineering for the SPIE society, USA.

How to cite this paper: Rami Zewail, Ahmed Hag-ElSafi, "Appearance-based Salient Features Extraction in Medical Images Using Sparse Contourlet-based Representation", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.9, No.9, pp.1-10, 2017.DOI: 10.5815/ijigsp.2017.09.01