A DFT based rotation and scale invariant Gabor texture descriptor and its application to gastroenterology

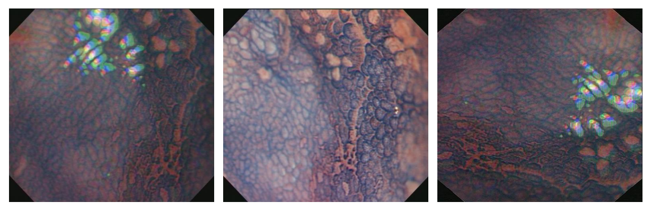
Farhan Riaz[1], Mario Dinis Ribeiro[3], Pedro Pimentel-Nunes[3], Miguel Tavares Coimbra[2]

***Abstract*— Classification of texture images, especially in cases where the images are subjected to arbitrary rotation and scale changes due to dynamic imaging conditions is a challenging problem in computer vision. This paper proposes a novel methodology to obtain rotation and scale invariant texture features from the images. The feature extraction for a given image involves the calculation of the averages of Gabor filter responses at various scales and orientations. For rotation and scaling of images, these averages indicate the respective shifts in the features. These shifts are normalized by doing summations of Gabor responses across scales and then taking the magnitude of Discrete Fourier Transforms across the resulting features and vice versa thus giving us scale and rotation invariant texture features. The proposed features are used for identifying cancer in the vital stained magnification endoscopy images. Experiments demonstrate the superiority of the proposed feature set over several other state-of-the-art texture feature extraction methods with around 90% classification accuracy for identifying cancer in gastroenterology images.**

I. INTRODUCTION

Texture recognition is an active area of research, that plays an important role in a wide variety of computer vision problems. Over the past three decades, many attempts have been made to propose various texture descriptors but the subjectivity of their design is based on parameters which are essential for specific applications. An important criteria for the design of texture descriptors for some applications is their invariance to rotation, scaling and illumination changes in the images. An example of such an application is feature extraction from endoscopic images for the diagnosis of gastrointestinal (GI) cancer, where the need for invariance originates from the decreased low level control of the camera due to inability to maintain a fixed distance between the endoscopic probe and the visualized tissue (affecting scale and illumination), and varying perspectives of visualizing the internal walls of the GI tract. A reliable Computer Assisted Decision (CAD) system is expected to be able to cope up with these variations in the images, thus making the diagnosis robust to varying imaging conditions.

Conventionally in gastroenterology (GE), cancer diagnosis is done by visually finding some specific patterns in the images which lay the foundation of diagnosis. Currently, this involves a visual inspection of the internal GI organs and is therefore a function of human factors such as expertise and perception. CAD systems can be useful in improving the screening ability of the physicians by providing them with second opinion or assisting them in situations where their confidence level for diagnosis is low. Traditional texture analysis methods used for more conventional scenarios are not expected to behave satisfactorily given the imaging dynamics in GE images. Texture feature extraction has an extended history and a wide spectrum of methods are available in the literature.



(a) Original image (b) Magnified image (c) Rotated image

Fig. 1: Dynamic imaging conditions in endoscopy. (a). A CH image captured during a live exam, (b). CH image captured by bringing endoscopic probe closer to the tissue wall (scale and homogeneous illumination), (c). Artificially rotated image of same pattern (CH - Chromoendoscopy).

In this paper, we will explore the use of Gabor filters for the extraction of scale and rotation invariant texture features. The empirical demonstration of invariance characteristics of the proposed descriptor is carried out. The proposed descriptor is then used for the detection of cancer in gastroenterology images. The performance of the proposed descriptor is done with several other state-of-the-art methods. The rest of the paper is organized as follows: we will briefly explain the Gabor filters and their invariance poperties (Section II and III), followed by the proposal of the novel feature set (Section IV). We will later explain our dataset (Section V) followed by the presentation of our results (Section VI) and discuss them (Section VII).

II. GABOR FILTERS

A Gabor filter is a sinsuoidal plane of a particular frequency and orientation modulated by a Gaussian envelope [1]. These filters have been shown to possess good localization characteristics in both space and frequency and have been used very successfully for several applications. A two

This work was supported by National University of Sciences and Technology (NUST), Islamabad, Pakistan and Fundação para a Ciência e a Tecnologia, Portugal.

[1]F. Riaz is with the National University of Sciences and Technology, Pakistan (e-mail: farhan.riaz@ceme.nust.edu.pk).

[2]M. Coimbra is with the Instituto de Telecomunicaçoes, Department of Computer Science, Faculdade de Ciências da Universidade do Porto, 4169-007 Porto, Portugal (e-mail: mcoimbra@dcc.fc.up.pt).

[3]M. Dinis-Ribeiro and Pedro Pimentel-Nunes are with the Instituto Português de Oncologia, Porto, Portugal.

dimensional Gabor function $g(x,y)$ and its Fourier transform $G(u,v)$ can be written as [2]:

$$g(x,y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jWx\right) \tag{1}$$

$$G(u,v) = \exp\left(\frac{1}{2}\left[\frac{(u-W)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right]\right) \tag{2}$$

Where $\sigma_u = \frac{1}{2\pi\sigma_x}$, $\sigma_v = \frac{1}{2\pi\sigma_y}$ and W is a constant representing the center frequency of the high frequency. Equation 1 is the product of a Gaussian function with complex sinusoid. This forms a bandpass filter in the frequency domain, where the bandwidth and center frequency of the filter are controlled by the standard deviation of the Gaussian function and the frequency of complex sinusoid respectively. A Gabor filter bank having a number of bandpass filters, with varying center frequencies, bandwidths and orientations is controlled by the parameters of Gabor wavelets. An input image, $\xi(x,y)$ when filtered by the set of Gabor wavelets $g(x,y)$ is given as:

$$R_{mn}(x,y) = \int \xi(x,y)g^*_{mn}(x-x_1,y-y_1)dx_1dy_1 \tag{3}$$

Where m corresponds to the m^{th} scale and n corresponds to the n^{th} orientation of the Gabor filter.

III. Invariance Properties of Gabor filters

Let us consider an image $\xi_1(x,y)$ filtered by Gabor filters $g(x,y,f,\theta)$ to give $r_{\xi_1}(x,y,f,\theta)$ where x and y represent the spatial location, f and θ are center frequency and orientation of the filter respectively. For $\xi_2(x,y)$, which is the rotated version of $\xi_1(x,y)$ by an angle α it can be shown that the resulting GF response is [3]:

$$r_{\xi_2}(x,y,f,\theta) = r_{\xi_1}(x,y,f,\theta-\alpha) \tag{4}$$

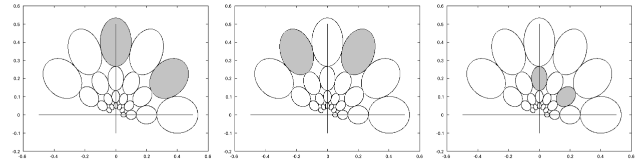
This property shows that the Gabor response for an image which has been rotated by some angle α is the same as the response of a correspondingly rotated filter for the original image. Let us now assume that the image $\xi_1(x,y)$ has been scaled homogeneously by a constant factor c giving us $\xi_3(x,y) = \xi_1(cx,cy)$. It is easy to show that the filter response of ξ_3 can be given as [3]:

$$r_{\xi_3}(x,y,f,\theta) = \frac{1}{c^2}r_{\xi_1}(cx,cy,\frac{f}{c},\theta) \tag{5}$$

The scale property of Gabor filters shows that the response for a scaled image by a constant factor 'c' is equal to the attenuated response of a correspondingly scaled Gabor filter used for filtering the original image. Let us now assume that $\xi_1(x,y)$ undergoes a uniform illumination change. This change can be modeled by multiplying the image with a constant i.e., $\xi_4(x,y) = c\xi_1(x,y)$. Where 'c' is a constant factor. The filter response for ξ_4 can be derived as

$$r_{\xi_4}(x,y,f,\theta) = cr_{\xi_1}(x,y,f,\theta) \tag{6}$$

which represents simple scaling of the feature vector.



(a) Original image (b) Rotated image (c) Scaled image

Fig. 2: Effect of rotation and scaling of image on Gabor responses: transfer of energy to other filter banks, dilated and rotated by the same amount as that of the image.

IV. Proposed Methodology

Assuming that the local texture in an image is spatially homogeneous, the Gabor filter responses for an image region can aggregated by taking the mean of Gabor filter responses across various scales and orientations giving us:

$$G_\mu = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1N} & \mu_{21} & \cdots & \mu_{MN} \end{bmatrix} \tag{7}$$

which represents a collection of means of Gabor filter responses in the form of a vector. M is the total number of scales and N is the total number of orientations for which Gabor filter responses have been calculated. Given that Gabor responses for a transformed images will be represented by a filter that is transformed by the same amount as that of the image, this vector will undergo shifts when an image undergoes rotation and scale changes. If these shifts are catered for in the feature vectors, rotation and scale invariant features can be constructed. Previously, a specific arrangement of such features as a matrix has been utilized such that the rotation and scale changes are interpreted as row and column shifts. Autocorrelation of the resulting matrix has led to the foundation of invariant features as it gives shift invariant representation of the matrix [4]. However, this increases the size of the feature vector making it difficult to use these features for applications with limited data as it results in sparse feature spaces. Let us now define:

$$G^{(S)}_{\mu_n} = \sum_{k=1}^{M} G_{\mu_{kn}} \tag{8}$$

where $G^{(S)}_{\mu_n}$ represents the summation of the aggregated filter responses at n^{th} orientation. It is important to note that $G^{(S)}_{\mu_n}$ is a vector which is scale invariant as the responses across all scales have been summed up. However, if an image is rotated by a certain angle, this will be represented by $G'^{(S)}_{\mu_n}$ that is a shifted representation of $G^{(S)}_{\mu_n}$. It should be noted that the application of a shift invariant transform on $G'^{(S)}_{\mu_n}$ can give us rotation invariant features for the image. It is well known that the magnitude of the Discrete Fourier Transform (DFT) is shift invariant [5] and thus the effect of rotation on the feature vector can be curtailed using DFT of $G^{(S)}_{\mu_n}$:

$$F_{\mu}^{SR} = |\sum_{n=1}^{N} G_{\mu_n}^{(S)}.\exp(-i2\pi\frac{k}{N}n)| \tag{9}$$

It is important to note that F_{μ}^{SR} is invariant to rotation and scale changes in the images. Let us now define:

$$G_{\mu_k}^{(R)} = \sum_{n=1}^{N} G_{\mu_{nk}} \tag{10}$$

where $G_{\mu_k}^{(R)}$ represents the summation of the aggregated filter responses at k^{th} scale. Note that $G_{\mu_k}^{(R)}$ is a vector which is rotation invariant as the responses across all orientations have been summed up. However, if an image is scaled by a certain amount, this will be represented by $G_{\mu_k}^{'(R)}$ which is a shifted representation of $G_{\mu_k}^{(R)}$. The magnitude of the Discrete Fourier Transform (DFT) is shift invariant and thus the effect of scale changes on the feature vector can be curtailed using DFT of $G_{\mu_k}^{(R)}$.

$$F_{\mu}^{RS} = |\sum_{k=1}^{M} G_{\mu_k}^{(R)}.\exp(-i2\pi\frac{k}{M}n)| \tag{11}$$

where F_{μ}^{RS} is invariant to rotation and scale changes in the images. Concatenation of features obtained using equations 9 and 11 gives us

$$F = [\ F_{\mu}^{SR} \quad F_{\mu}^{RS}\ ] \tag{12}$$

where F is a novel feature vector which is invariant to rotation and scale changes in the images. This feature set can now be used for classifying the texture images using standard machine learning algorithms.

V. Materials

For our experiments, we have used chromoendoscopy (CH) images that were obtained using an Olympus GIF-H180 endoscope at the Portuguese Institute of Oncology (IPO) Porto, Portugal during routine clinical work. Optical characteristics of this endoscope include 140° field of view and four way angulation (210° up, 90° down and 100° right/left). The endoscopic videos were recorded on tapes using a Digital Video (DV) recorder while performing real endoscopic examinations. Around 4 hours of video (360000 frames) were analyzed and 176 images were initially selected given their clinical relevance. Two clinicians classified these images into three groups, following Dinis-Ribeiro's classification proposal [6], manually segmenting the image region that led them to this conclusion and labeling their choice with a confidence value (high or low confidence). Finally, a careful analysis of images was carried out to remove the images belonging to the same patients, giving us a dataset of 130 images distributed as 31% (40 images) belonging to Group I, 57% (75 images) belonging to Group II and 12% (15 images) to Group III.

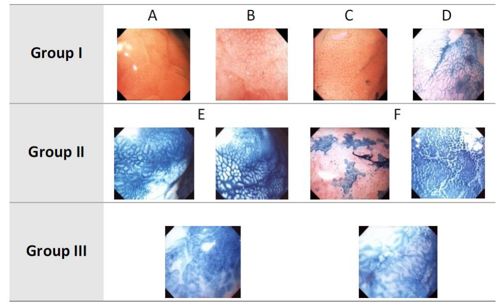


Fig. 3: Dinis-Ribeiro classification proposal for classification of vital stained magnification endoscopy images [6]

VI. Experimental Results

Gabor filtering in our experiments was done using 6 orientations and 8 scales (i.e. N=6 and M=8). The selection of these parameters was done empirically by analyzing the performance of filters over a range of possible parameters and then selecting the suitable ones for our datasets. Nonetheless, the relative results concerning both invariance and classification are consistent irrespective of the filter parameters. The [1]Weka data mining tool was used in the classification experiments presented here [7]. Classification was done using Support Vector Machines [8].

A. Invariance properties of proposed descriptor

For assessing the invariance characteristics of the proposed descriptor we have used matlab routines to obtain rotated (*imrotate* Matlab function) and scaled (*resample* Matlab function) versions of real endoscopic images. The original images from the dataset are rotated (30, 60, 90, 120, and 150 degrees) and split into two equal halves: One (original images, 30, 60 degrees rotated) is used for training while the other (90, 120, 150 degrees rotated) is used for testing the classifier. Similar experiments are repeated to assess the effect of scaling (1.2x, 1.4x and 1.6x) the GE images on the overall classification performance of the descriptor. Experiments demonstrate the stability of the proposed descriptor to image transformations hinting at the efficiency of the proposed descriptor in dynamic imaging conditions.

TABLE I: Invariance of proposed features as compared to state-of-the-art feature extraction method.

	Rotation	Scaling	Original
Novel	**87.50%**	**72%**	**88.40%**
Manjunath et al. [2]	83%	37%	88%

B. Classification Experiments

In our classification experiments, we have used our novel descriptor for the classification of Gastrointestinal (GI) images in one of the two possible classes: normal (Group A) or potentially cancerous (Group B). Consequently, the images in

[1]http://www.cs.waikato.ac.nz/ml/weka/

the DR proposal belonging to Group I should be classified as Normal (Group A) whereas those belonging to Group II or III should be classified as potentially cancerous (Group B). For classification of images, we used only manually annotated image patches i.e., we assumed perfect segmentation of images based on clinician's observations. All results were obtained using 10-fold cross validation. Classification was done using Support Vector Machines (SVM) [8]. Performance comparison was done using overall classification accuracy and Area under Receiver Operating Characteristics (ROC) curves.

TABLE II: Comparison of performance obtained using novel features and state-of-the-art feature extraction methods (AUC - Area under ROC curve).

	Overall Acc.	AUC
Novel	**89.5 %**	**0.87**
Riaz et al. [4]	87.70%	0.86
Manjunath et al. [2]	85.40%	0.85
T. Ojala [9]	80%	0.78

To demonstrate the performance of our proposed descriptor, we have compared its overall accuracy with other state-of-the-art descriptors which have been used previously for classifying gastroenterology (GE) images. The test bed used for all these methods is the same as that used for our proposed method. Our experiments demonstrate that the Local Binary Patterns (LBP) show the worst performance amongst state-of-the-art methods. We suspect that this is because of the lack of multiresolution analysis in traditional LBPs. Among descriptors based on Gabor filters, HT shows relatively low classification rates. Our novel descriptor shows better performance as compared to the other descriptors considered in this paper.

Given that the theme of the proposed work is to raise an alarm in case of suspicious visual pattern in the visualized tissue, a false positive for identifying potentially cancerous tissue is not a bad situation at all. Given this, we also analyze the performance of our proposed descriptor using the confusion matrix (Table III). A visual inspection of the table shows that out of the 14 misclassified instance, 8 images are originally normal which our descriptor is identifying as potentially cancerous. The more serious cases are false negatives (6 in number, only around 4% of all the classified images). The results show that the proposed feature set can be used for raising alarms by the assisted decision systems with high accuracy.

TABLE III: Confusion matrix using novel descriptor

		Automatic Classification	
		Group A	Group B
Manual Classification	Group A	32	8
	Group B	6	84

VII. Discussion

In this paper, a texture descriptor based on Gabor filters was proposed. Traditional features based on Gabor filters are sensitive to orientation and scale changes in the images and thus the features are not invariant to scale and rotation changes in the images. We have addressed this issue by summing filter responses across scales and applying DFT to the resulting features and vice versa giving us rotation and scale invariant texture features. The proposed features are used for classifying the gastroenterology (GE) images for the detection of cancer. Comparison with several other state-of-the-art methods was done. Overall classification accuracy and area under ROC curves was used for the quantitative comparisons. Our results show the superiority of the proposed descriptor as compared to other state-of-the-art feature extraction methods. Our results show a high detection rate of cancer implying that the proposed descriptor can be effectively used for raising alarms when classifying the GE images.

One of the problems typically encountered in GE imaging scenarios is the presence of illumination gradient in the images. This happens because of the the inability of the physicians to ensure that the full imaging site is orthogonal to the camera as the stomach wall is not rigid. In the future, we aim to extend our work to research on novel image descriptors which in addition to scale and rotation invariance can also deal with illumination gradients in the images. We hope that the use of such descriptors can further improve the results for the classification of gastroenterology images.

Acknowledgements

The authors would like to thank Andr Sousa of the Instituto de Telecomunicacoes and Miguel Areia from Portuguese Institute of Oncology, Coimbra for their support in carrying out this research. This work was financially supported by National University of Sciences and Technology (NUST), Islamabad, Pakistan and Portuguese FCT (Portuguese Science Foundation) grants SFRH/BD/45066/2008 and PTDC/EIA-CCO/109982/2009. We would like to thank NUST and FCT for their vital support for carrying out this research.

References

[1] D. Gabor, "Theory of communication," *IEE Journal of Radio and Communication Engineering*, vol. 93, no. 26, 1946.
[2] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, 1996.
[3] J.-K. Kamarainen, V. Kyrki, and H. Kmrinen, "Invariance properties of gabor filter based features - overview and applications," *IEEE Trans. on Image Process*, vol. 15, no. 5, 2006.
[4] F. Riaz, F. Silva, M. Ribeiro, and M. Coimbra, "Invariant gabor texture descriptors for classification of gastroenterology images," *Biomedical Engineering, IEEE Transactions on*, vol. 59, no. 10, pp. 2893–2904, 2012.
[5] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*. Prentice-Hall, 1999.
[6] M. D. Ribeiro, "Clinical, endoscopic and laboratorial assessment of patients with associated lesions to gastric adenocarcinoma," *Faculdade de Medicina da Universidade do Porto, PhD thesis*, 2005.
[7] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten, "The weka data mining software: an update," *ACM SIGKKD Explorations Newsletter*, vol. 11, no. 1, 2009.
[8] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer, 1995.
[9] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognition*, vol. 29, 1996.