
(Invited Paper)

Chao-Hui HUANG**,†¶, Daniel RACOCEANU*||†, Ludovic ROUX*,‡ and Thomas C. PUTTI§

*Image & Pervasive Access Lab (IPAL, UMI2955, CNRS), France
**Bioinformatics Institute, Agency for Science, Technology and Research, Singapore
||Centre National de la Recherche Scientifique (CNRS), France
†School of Computing, National University of Singapore, Singapore
‡Université Joseph Fourier, Grenoble, France
§Department of Pathology of the National University Hospital, Singapore
¶EMail: ch@mail2000.com.tw

Abstract—Bio-inspired computer vision is an emerging field. It aims to reproduce the capabilities of biological vision systems, eventually to simulate the visual functions for various purposes. In this paper, we propose a bio-inspired computer visual system using Graphical Processing Unit (GPU), and its application on breast cancer prognosis. The system simulates the major biological mechanisms of human visual system, such as encoding the edges, textures, and shapes of natural scenes. The system extracts visual features from an input image using a mechanism which is similar to human visual system. Then those visual features are converted to artificial neural activity signals which can be classified by a machine learning algorithm. As a result, the elements of the input image, which might be related to particular knowledge concepts, can be identified.

In order to operate the bio-inspired computer visual system, in this work, a new computer language, named Visual Pattern Assessment Language (ViPAL), is proposed. The ViPAL provides a shell between the bio-inspired computer visual system and the users. The complexity of the low-level image feature assessment can be managed by the shell. As a result, the workload of the users can be reduced.

Based on the bio-inspired computer visual system and the ViPAL, we propose an application for breast cancer prognostics. Currently, there is no computer-aided system to help pathologists making their final decision on breast cancer prognosis. Thus prognosis depends only on the medical human doctors. A well-trained pathologist usually can complete the analysis of a biopsy in 10 to 20 seconds. However, in a big medical institute, sometimes a pathologist need to analyse more than one hundred biopsies per day. So, for the sake of improving the reliability of breast cancer prognosis, we implemented the proposed bio-inspired computer visual system using GPU for detection of breast cancer invasive areas. As a result, object recognition and scene understanding of digitized breast biopsies are performed. Consequently, the system can roughly simulate the decision-making as the procedures of the breast cancer prognosis.

I. INTRODUCTION

Researchers have been interested for years in trying to mimic biological vision systems. The traditional way to perform it on image processing and computer vision usually follows a top-down approach: an engineer evaluates the requirements, creates an algorithm and imbues it with its intelligence. Bio-inspired computer vision, on the other hand, takes a more bottom-up, decentralised approach. Most of the bio-inspired systems are starting from the simulations of simple organisms by computing a set of simple rules. Those simple rules collaborate with each other and produce information for a higher level application. As a result, some forms of complex behaviour arise [1].

In this study, we aim to create a bio-inspired computer visual system which simulates some mechanisms of human visual system in a biological way. We simulate neural activity of the retina and of the primary visual cortex, in order to mimic some visual functions such as encoding the edges, textures, and shapes of natural scenes. The system extracts visual features from an input image by using mechanisms similar to human visual system. Those visual features are converted to artificial neural activity signals which are then classified by a machine learning algorithm. As a result, the contents of the input image, which might be related to particular knowledge concepts, can be identified.

By using the bio-inspired computer visual system, it is possible to make connections between image features and visual knowledge concepts. However, a visual knowledge concept (for example: the invasive area on a digitized biopsy) is rather complex and usually it is not possible to represented it by a single image feature. Thus, in this paper, we propose a new computer language, named Visual Pattern Assessment Language (ViPAL), in order to provide to the users a high-level user interface for the application.

The application we propose is a computer-aided invasive area detection system for breast cancer prognosis. Breast cancer is the second most common type of cancer and the fifth most common cause of death from cancer. Early signs of breast cancer are often hidden within the breast tissues. Thus, breast cancer prognosis relies highly on the medical human...
During breast cancer prognosis, the pathologists need to identify the tumours and to analyse their histological appearance. Usually they first isolate the most interesting regions, the so-called Regions of Interest (ROI), at lower magnification. Then, they analyse the ROI at higher magnification and make their prognosis. A well-trained pathologist usually can complete the analysis of a biopsy in 10 to 20 seconds. However a doctor in a big medical institute sometimes need to evaluate more than one hundred biopsies per day. Thus, to support their works for breast cancer prognosis becomes an important issue.

Although many medical modalities are used to support the prognosis of breast cancer, histopathology is the golden standard to make the final decision for the prognosis and the following treatments. Thus, some researchers proposed their studies based on histopathology [3], [4], [5]. However, to the authors' knowledge, not many works have been published about computer-aided breast cancer prognosis on histopathology. This is probably because of the difficulty of the problem itself. One of the key issues is that, since the image size of the digitized breast biopsies is quite huge, the implementations in those researches usually requires a lot of computation time [6], often several hours, sometimes even several days. Some researchers even suggested to use specific integrated circuits to implement the algorithms [6].

In this paper, we introduce an invasive area detection algorithm based on bio-inspired visual simulation. A part of the human visual system is simulated in order to approximately mimic the behaviours of the visual system of the pathologists. Thus, the number of parallelizable components can be maximized. Consequently, the Graphic Processing Unit (GPU) can be used. In this study, NVIDIA's GPU and its Compute Unified Device Architecture (CUDA) were used to accelerate the algorithm.

We have applied our method on 15 breast biopsies. All these biopsies have been annotated locally and globally by pathologists of the Department of Pathology of the National University Hospital (NUH), Singapore.

II. METHODS

The proposed model is based on the human visual system. Human visual system includes two major parts: the retina and the primary visual cortex. In the retina, there are two major photoreceptors: the *cone-* and the *rod-cells*. The cone-cells can be classified into three types: *L-, M-,*, and *S-cells*. These kinds of photoreceptors react to various light wavelengths. A set of opponent photoreceptors form a Receptive Field (*RF*) (see Fig. 1). These photoreceptors form a field which is called ganglion RF since they collect visual information and send neural spikes to a ganglion cell. The ganglion RF can be considered as the fundamental element of the human visual system. Eventually, the ganglion cells produce various stimulations which are sent to the primary visual cortex [7].

This biological mechanism results in the color space of human visual sensory. Hering’s color opponent theory [8] is a popular way to describe the color space [7]. This theory interprets the results of color mixing by proposing the existence of three opponent processes in the eye, the brain, or both. These three processes are the red-green, the yellow-blue, and the black-white sensations. These color stimulations are transferred to the primary visual cortex [7].

In the primary visual cortex, there are two major kinds of cells: the *simple-* and the *complex-cells* [7]. Generally speaking, these cells produce two kinds of visual features: the *first-* and the *second-order* features [9], [7]. The first-
order feature contains the information of intensities of various color channels, and the second-order feature includes spatial variance of visual signal [10], [11].

This knowledge of human visual system can help us to discover the strategies of decision-making in cognitive process [12] when the pathologists are analysing a breast biopsy for the prognosis of breast cancer. In this study, we simulate some mechanisms of human visual system, generate the first- and second-order features, and classify these features by Support Vector Machine (SVM). After the SVM has been trained, the system is capable to detect the ROIs. In order to accelerate the computing, the GPU technology is used in some major components of the system. Consequently, the result of each input image can be obtained in a reasonable period of time.

In the experiments, the training patterns were given by the researchers under the supervision of the pathologist. Those patterns came from a set of golden samples which have been identified by the pathologist. These golden samples contain typical features of invasive area. During the learning phase, the system was sometimes unable to identify some positive cases. After we discussed with the pathologist, some of these cases were considered as the training patterns and the system was retrained in order to improve its performance. The flow of breast cancer invasive areas detection is shown on Fig. 2.

A. Color Representation

In this study, we aim to discover the relationship between a breast biopsy image displayed on a screen, and the decision-making procedure of a pathologist who is analysing the image on the screen.

In order to simulate the related visual reactions of human visual system, we have to obtain few parameters. First, the sensitivities of the photoreceptors in visible spectrum. A well-known cone-cell sensitivity factor in visible spectrum is presented in Fig. 3(a) [13], [14]. Second, the radiance factor of the screen in the visible spectrum is required in order to evaluate the energy which might be received by the human visual system. Fig. 3(b) shows a radiance factor in the visible spectrum of a typical CRT screen. Most of the modern screens provide Gamma correction. An example of Gamma Display Coefficient (GDC) functions is shown in Fig. 3(c). The GDC functions of the screen are also required. These coefficients can be used to evaluate the radiance which is emitting from the screen.

Based on what we introduced above, a transform operator can be obtained:

\[
\begin{bmatrix}
  l \\
  m \\
  s
\end{bmatrix} = s \cdot T \cdot \begin{bmatrix}
  \Gamma_{\text{red}}(r) \\
  \Gamma_{\text{green}}(g) \\
  \Gamma_{\text{blue}}(b)
\end{bmatrix},
\]

where \( r, g, b \) are the colors of the pixel, \( \Gamma_{\text{red}}(\cdot), \Gamma_{\text{green}}(\cdot), \) and \( \Gamma_{\text{blue}}(\cdot) \), are GDC functions used to reconstruct the Gamma correction, \( s \) is a necessary scale in order to normalize the input values, and \( T \) is a \( 3 \times 3 \) matrix which is the linear combination of the cone-cell sensitivity factor and the radiance factor in visible spectrum. Further details can be found in [13], [14], [15].

In order to represent a color image from a psychological point of view, we propose the use of Ewald Hering’s opponent color theory [8]:

\[
\begin{align*}
    u_{\text{red}} &= \log(l),
    u_{\text{green}} &= \log(m),
    u_{\text{blue}} &= \log(s),
    u_{\text{yellow}} &= \log\left(\frac{1}{2}(l + m)\right).
\end{align*}
\]

Note that logarithm is used to approximate the human visual system.

Here \( (u_{\text{red}}, u_{\text{green}}) \) and \( (u_{\text{blue}}, u_{\text{yellow}}) \) are used to describe the Hering’s opponent color, Red-Green (RG), Blue-Yellow (BY), and Luminance (L):

\[
\begin{align*}
    v_{\text{RG}} &= u_{\text{red}} - u_{\text{green}}, \\
    v_{\text{BY}} &= u_{\text{blue}} - u_{\text{yellow}}, \\
    v_{L} &= \frac{1}{2}(u_{\text{red}} + u_{\text{green}} + u_{\text{blue}}) - 1.
\end{align*}
\]

Those color pairs and the luminance information can be used to describe the visual signal in the human visual system [7], [16], [15].

Based on the color information, the human visual system is able to extract features from the image. In our method, these feature extraction algorithms include intensity, colors and texture perception.

B. Computer Fovea Model

In order to model the retina, various models are suggested for various purposes [17], [18], [19]. Huang et al. proposed a computer fovea model with various applications [15]. The computer fovea model aims to model a simplified version of the full retina system. First, a general assumption of center/surround receptive field (RF) of ganglion can be considered as a reference. Some physiological experiments indicated that the RF of the ganglion exhibits a center/surround characteristic. Furthermore, various publications stated that the RF of ganglion can be modeled as follows [20], [7], [21], [17], [18], [19]:

\[
h_{G}(\delta(\alpha)) \triangleq \nabla^{2}(G_{\sigma_{G}}(\alpha)),
\]

Fig. 4. The computer fovea model, proposed by Huang et al. [15], including the photoreceptor cells, the horizontal cells, the bipolar cells and the ganglion cells.
where $∇^2(·)$ denotes a Laplace filter and $G_{σ_R}(·)$ is a Gaussian filter with standard deviation $σ_R$.

1) Photoreceptor: Photoreceptor includes various cone and rod cells. There cells react to different wavelength of visible light. Cone cells can be roughly classified into three types: short wavelength (S/blue cell), middle wavelength (M/green cell), and long wavelength (L/red cell). Generally speaking, a bipolar cell collects the signal from a set of cone cells and form a diffuse pathway. The use of a Gaussian function to model the diffusion is suggested in various publications [17], [19]. Thus,

$$h_R(δ(α)) ≜ G_{σ_R}(α).$$  

where $G_{σ_R}(·)$ represents a Gaussian filter. As mentioned in [17], $σ_R$ represents the standard deviation with a range from 1.5 to 12 (cell space).

2) Bipolar Cell: Bipolar cells collect the signal from a number of cone cells and transmit the spiking to ganglion cells. Although there are various types of bipolar cells, in this model the simplest one is chosen. That is, the bipolar cell maps to one cone cell and its opponent channel of cone cell with a bias $-b$.

3) Horizontal Cell: Horizontal cells have been considered as a set of cells to contribute to the surround response of bipolar cells. The horizontal cells have been shown to be color opponent in response.

$$h_H(α) ≜ \frac{1}{b}(δ(α) - \frac{1}{g}h_G(α) \otimes h_R^{-1}(α)).$$  

4) Ganglion: In most cases, a ganglion cell collects signal from only one bipolar cell. Thus, in this model, only a bias $g$ is used to represent the function of a ganglion cell.

C. First Order Feature

1) Receptive Field: In human visual system, the Receptive Field (RF) of ganglion is considered as the fundamental element [7]. In the first order extraction, the RF is defined as a set of pixels which are sampled on a particular area. Generally speaking, we need to compute all of the pixels in this area. However, in order to reduce the computational costs, we suggest to pick up only some pixels in this area according to a sampling distribution. The sampling distribution $p_s(·)$ is based on Gaussian distribution such that

$$p_s(·) \overset{d}{=} N(μ_s, Σ_s),$$  

where $l = (l_1, l_2, ..., l_m), l_i \in \mathbb{R}^2$ is a set of the locations of the data points, $μ_s = [μ_1, μ_2]^T$ is the mean, and $Σ_s$ is the covariance matrix:

$$Σ_s = \begin{bmatrix} σ_{11} & σ_{12} \\ σ_{21} & σ_{22} \end{bmatrix}.$$  

In the experiments, $σ_{11} = σ_{22} = 1000$ pixels and $σ_{12} = σ_{21} = 0$ (see Fig. 5).

2) First Order Feature Extraction: When a RF is chosen, a set of pixel locations $l = (l_1, l_2, ..., l_m), l_i \in \mathbb{R}^2$ is generated. We obtain the so-called Hering’s opponent color: Red-Green (RG), Blue-Yellow (BY), and the Luminance (L) information as follows:

$$v_i = [v_{RG}(l_i), v_{BY}(l_i), v_L(l_i)], \ i = 1 ... m.$$  

According to Geusebroek, et al., the opponent color theory can be applied to computer vision and implemented as the Gaussian color model [16], [22], [23], [24]. Thus, all $v_i$ can be used to generate a multivariate Gaussian distribution:

$$p_1(v) \overset{d}{=} N(μ_1, Σ_1).$$  

The mean $μ_1$ and the covariance matrix $Σ_1$ can be obtained by Expectation Conditional Maximization (ECM) [25]. In our study, $μ_1$ and $Σ_1$ are defined as the first order features.

D. Second Order Feature

Second-order features cannot be detected by mechanisms sensitive only to intensity changes. It is ubiquitous in visual scenes, but the neuronal mechanisms mediating perception of such stimuli are not well understood [26], [27].

Most of the researchers agree that the second-order feature includes spatial variance of visual signal [10], [11]. In order to extract the basis, various methods were proposed. Generally speaking, most of them invoked one or several basis analysing algorithms. Those algorithms include Gabor filtering, Principal Component Analysis (PCA), Independent Component Analysis (ICA), Sparse Coding, etc. Each of them has various
can be generated. Finally, the parameters $\mu$ obtained. Third, the Gaussian distribution of these coefficients of these basis vectors for various images can be obtained using a training set $x_1, \ldots, x_m$. Thus, an optimization problem for $b$ and $s$ corresponds to the minimization of:

$$\sum_{i=1}^{m} \frac{1}{2\sigma^2} \|x_i - \sum_{j=1}^{n} b_j s_{j,i}\|^2 + \beta \sum_{i=1}^{m} \sum_{j=1}^{n} \phi(s_{j,i}),$$

subject to $\|b_j\| \leq c, j = 1 \ldots n$, (12)

where $\sigma^2$ is the variance of the noise, $\phi(\cdot)$ is a sparsity function, and $\beta$ is a constant [28], [31].

There are many solutions able to solve this typical optimization problem. In this study, we consider GPUs to implement projective gradient descent [32]. As a result, the computation time has consistently been reduced.

1) Receptive Field: The receptive field of the second order feature extraction is more complex than the first order feature extraction. First, following (7), a set of data point locations $l_1, \ldots, l_m \in \mathbb{R}^2$ is generated. Let $l_i$ be the center of an image patch, from which a set of patches $x_1, \ldots, x_m$ can be obtained such that:

$$x_i = I(l_i - d), d = [n_1, n_2]^T,$$

where $I$ represents a channel which is obtained from the input image, $-N \leq n_1 < N$, and $-N \leq n_2 < N$, where $N$ is the size of the patches. In the experiments, $N = 10$.

Since all of the patches $x_1, \ldots, x_m$ are captured from the same region, they are sharing the same features. Thus, the texture can be encoded by coefficients $s_{i,j}$ in (12) with a set of basis vectors $b_j$, which need to be obtained previously. All of $s_{i,j}$ of RG, BY, and L channels are obtained, such as:

$$s_{i,j} = [s_{RG,j}(l_i), s_{BY,j}(l_i), s_{L,j}(l_i)],
\quad i = 1 \ldots m \text{ and } j = 1 \ldots n.$$ (14)

All $s_{i,j}$ generate a multivariate Gaussian probability distribution:

$$p_2 \triangleq \mathcal{N}(\mu_2, \Sigma_2).$$ (15)

Like the first order feature, the mean $\mu_2$ and the covariance matrix $\Sigma_2$ are obtained by Expectation Conditional Maximization (ECM) [25]. In our study, $\mu_2$ and $\Sigma_2$ are defined as the second order features.
Fig. 6. 400 basis vectors on 20×20 pixels image patches, extracted from digitized breast biopsies (virtual microscope).

Fig. 8. Performance of our system. The AUC is 0.9186.

III. EXPERIMENTAL RESULTS AND VALIDATION

In the testing phase, more than 25 images are obtained from 15 digitized breast biopsy slides. Two of the images are used as the training patterns. The global Receiver Operating Characteristic (ROC) curve and the Clinical Test Performance are shown in Fig. 8. The global Area Under the Curve (AUC) is 0.9186. The optimal cut point is −1.1348. We also computed the difference of ROIs between the results of our ROI detection and the ground truth by:

\[ c = \frac{R \cap G}{R \cup G}, \]

where \( R \) is our result, \( G \) is the ground truth provided by the experts, and \( c \in (0, 1) \) is the covering rate. The average of all of the testing results is 0.7077. Some of the results are presented in Table III.

Our results also present an interesting phenomenon. In the design of our system, the results are determined by the sign of the hyper plane \( \mathcal{H}(\cdot) \) which has been produced by SVM classification. However, the test performance suggested that the optimal cut point is −1.1348. This is because the pathologists were asked to give three regions for each image, the invasive area, the normal area, and unknown area. However, some suspect regions which are difficult to determine, tend to be considered as invasive areas by the pathologists. This results as a higher false-positive rate. In other words, the results reflect the exact behaviors of the pathologists who provided the evaluation patterns.

The performance of GPU acceleration is presented in Table II. The GPU is a GeForce 9400M from NVIDIA. The computer is an Apple Macbook with Intel CPU Core 2 Duo with 4G memory. With a pre-trained kernel, the computation time of low resolution image ROI detection for one slide is about 120 seconds. It depends on the size of the slide.

IV. CONCLUSION

In this paper, a bio-inspired computer visual system using GPU is presented. It is implemented for breast cancer prognosis. Based on a predefined image database, the invasive

<table>
<thead>
<tr>
<th>Table II</th>
<th>THE PERFORMANCE OF GPU ACCELERATION.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>GPU</td>
</tr>
<tr>
<td>Sparse Coding</td>
<td>1406.759</td>
</tr>
<tr>
<td>SVM</td>
<td>92.265</td>
</tr>
<tr>
<td>EM1</td>
<td>26.314</td>
</tr>
</tbody>
</table>

E. Visual Pattern Assessment Language (ViPAL)

The Visual Pattern Assessment Language (ViPAL) is a new computer language which provides the users to access and retrieve the image features by using a script. The design of the ViPAL is based on the well-known database script language: SQL. The ViPAL provides an environment where the users can retrieve images in the same way as they would access to the tables of a database. As a result, this language serves as the interface for the system operation.

In order to reduce the learning curve of the language, the design of ViPAL was based on the famous database manipulating script named SQL. An example is shown in Table I. It can be understood easily. In the example, first, an image is loaded into the memory and represented by a variable named input_image, and a feature file is loaded into feature. Then, the SELECT command performs the feature selection. Finally the result is shown by the SHOW IMAGE command.

The ViPAL provides the functions of file operation, image preview, routine operation, etc. As a result, the complexity of the system is covered by a shell. So the workload on the users can be reduced.
cancerous areas can be identified on the digitized breast biopsies. This study came out with an interesting topic, the inter/intra observation source variability. For the same biopsy, different doctors can make different prognosis. Sometimes even the same doctor makes a different prognosis for the same biopsy. Based on this study, we may be able to evaluate the variability between the prognosis. The proposed algorithm can even be applied to different areas (for example: ROI detection on X-ray images). Those will be in our future works.

The ViPAL language provides a shell between the bio-inspired computer visual system and the users. The complexity of the low-level image feature assessment can be managed by the shell. As a result, the workload of the users can be reduced.

ACKNOWLEDGMENT

This research is supported by A*STAR SERC 052 101 0103 (NUS R-252-000-319-305).

The authors thank Dr. Yan Karklin from New York University for this support on this work.

REFERENCES


TABLE III
SOME ROIs WHICH HAVE BEEN DETECTED BY OUR SYSTEM. THE FULL IMAGE SET CAN BE OBTAINED FROM IPAL WEB SITE.

<table>
<thead>
<tr>
<th>Source</th>
<th>Result</th>
<th>ROI</th>
<th>ROC/AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="source1.png" alt="Image" /></td>
<td><img src="result1.png" alt="Image" /></td>
<td><img src="roi1.png" alt="Image" /></td>
<td>0.87476</td>
</tr>
<tr>
<td><img src="source2.png" alt="Image" /></td>
<td><img src="result2.png" alt="Image" /></td>
<td><img src="roi2.png" alt="Image" /></td>
<td>0.87203</td>
</tr>
<tr>
<td><img src="source3.png" alt="Image" /></td>
<td><img src="result3.png" alt="Image" /></td>
<td><img src="roi3.png" alt="Image" /></td>
<td>0.96631</td>
</tr>
<tr>
<td><img src="source4.png" alt="Image" /></td>
<td><img src="result4.png" alt="Image" /></td>
<td><img src="roi4.png" alt="Image" /></td>
<td>0.89279</td>
</tr>
</tbody>
</table>


[29] Y. Karklin, “Hierarchical statistical models of computation in the visual

