THE USE OF GENETIC ALGORITHMS IN SEVERAL APPLICATIONS

Hilda Deborah*, Anita Kosasih**, Aniati Murni Arymurthy***

ABSTRACT

THE USE OF GENETIC ALGORITHMS IN SEVERAL APPLICATIONS. Genetic Algorithms (GAs) as evolutionary algorithms have been successfully applied to solve many problems on broad range of problem domains, such as remote sensing, biomedical, cultural heritage, physics, and many more. This paper summarizes the basic concept of GAs and its typical implementation in three different applications with different objectives. They include the applications on old document restoration, cell identification in Pap smear image, and a potential application in nuclear reactor core optimization. In the application of old document image restoration, GAs’ task is to do a process similar to contrast stretching by changing the input image’s gray level value following certain rules that will be described in this paper. The fitness function used for that problem is number of edges having length of 1 pixel. In the application of Pap smear cell identification, the task of GAs is to find objects with round shape in the input image and the objective criteria used here is the number of detected round objects. The last one is the use of GAs in a potential application for optimizing nuclear reactor core cost. The task of GAs in this particular application is to find solution(s) from the problem of minimizing nuclear reactor...
core cost and the amount of material used to build it, with of course considering safety as the constraint. These studies have reported that GAs is a potential method and has a good performance.

**Keywords**: Genetic algorithm, optimization, image restoration, pattern recognition, old document image, Pap smear cell, and nuclear reactor core.

**INTRODUCTION**

Genetic Algorithms (GAs) are evolutionary algorithms based on the mechanism of natural evolution on biological entities. It was found and developed by John Holland in the 60s and is originally developed to learn about evolution mechanism in nature so that it can be digitally simulated with computer [1]. GAs is usually used in optimization problems, where there is a set of solutions and the task of searching algorithms is to find the optimal solution.

The next sections are organized as follows. Section 2 contains an explanation about GAs concept and its implementation on digital image processing. Three examples of the use of GAs, in the application of (i) identifying cells in Pap smear images, (ii) image enhancement and restoration of old document images, and (iii) optimizing the cost of nuclear reactor core are discussed in Section 3, Section 4, and Section 5, consecutively. This paper is closed by a summary in Section 6.

**GENETIC ALGORITHMS**

GAs are search algorithms that try to model natural evolution on biological entities. It combines survival of the fittest among string structures with a structured yet randomized information exchange to form a search algorithm with some of the innovative flair of human search [2]. Some biological terminologies are still used in the implementation of GAs in digital computation, such as individual, population, generation, crossover, mutation, etc.

An individual or a chromosome is represented in a string of bits. Different characteristics of individual will be represented in different strings of bits. GAs take a group of individual, i.e. a population as an input and then, this population will be forwarded into some processes that together build up a complete GAs. Those processes are crossover, mutation, and selection, each of them will be explained in the next sub-sections. At the end of iteration in GAs, the obtained solutions will be copied to the next iteration as an input. The basic concept of GAs is shown in Fig. 1.

Because GAs is a randomize process, a criteria to measure the goodness of obtained solutions is a key factor. This criterion is what we call as fitness function. Each chromosome will have a fitness value representing its goodness as a solution. Intuitively, we want to maximize the fitness value [2], and that is simply means, a chromosome with a higher fitness value will have a higher chance of being selected as an optimal solution to the problem.
1. Crossover

Crossover is a process where a pair of chromosomes is chosen from the current population as parent chromosomes and exchanges the elements between them to create more chromosomes called offspring. The offspring is then added to the current population. After passing crossover process, the current population, i.e. the population after adding resulting offspring from crossover process, is forwarded to mutation process. See an example in Fig. 2.

![Crossover diagram]

Figure 1. Basic concept of Genetic Algorithm.

![Crossover example]

Figure 2. An example of a crossover process.
In Fig. 2 we can see that there are 2 parent chromosomes for crossover process input. The 3rd and 4th elements of each parent chromosomes are elements to be exchanged. After the crossover process, we’ll have 2 new chromosomes, i.e. offspring-1 and offspring-2 that will be added to the current population.

2. Mutation

In a mutation process, a small amount of individuals is selected from the current population to be mutated. Each of the selected individuals will get some of its element’s value changed to another value. The new individual is then added into current population. A further explanation is illustrated in Fig. 3.

![Mutation Example](image)

Figure 3. An example of mutation process.

3. Selection

Selection process is a mechanism of selecting an optimal solution from the current population. As what has been said before, each chromosome have its own fitness score that represents its profit or utility or goodness as a solution. Selection process will select solution(s) from the whole individuals in the population according to its fitness value.

GA IMPLEMENTATION ON OLD DOCUMENT IMAGE ENHANCEMENT & RESTORATION

The implementation of GAs in our previous work on an old document image problem domain [3] gives an evidence of GAs being a good search algorithm. Besides implementing GAs, the work has also tried to see the effect of using median filter to remove noises in the input image. The results show relatively effective image enhancement (92.9% of the data get more than 90% success rate), and in general the effect of using median filter preprocessing will decrease the sharpness of the images (only 59.5% of the data get success rate above 90%). Fig. 3 and Fig. 4 show a pair of
input images and their output obtained from GA, given that the input image was either preprocessed by a median filter (Fig. 5) or not (Fig. 4).

Figure 4. Result of restoration without applying median filter;
(a) Original image; (b) Restored image [3]

Figure 5. Result of restoration with applying median filter;
(a) Original image; (b) Restored image [3]
The application has used old document images coming from hundred years ago that has been degraded throughout the time as an input. The objective is to digitize, enhance and restore the input image into another image with a better quality that is as the approximation of its original condition. GAs’ role in this particular application is in doing a process similar to contrast stretching. Chromosome representation for the input image is its gray (or intensity) values. The fitness function which will determine whether a chromosome is selected or not is in the form of the number of edges containing only one pixel and the nearer a solution’s fitness value to input image fitness value, will has the higher chance to be selected as a solution or forwarded as parent chromosome for the next generation.

**GAS IMPLEMENTATION ON PAP SMEAR CELL IDENTIFICATION**

This section presents the application of GAs as a searching algorithm on the identification of the cells’ location in a single Pap smear image containing many cells [4]. The input to this application is a Pap smear test image which contains many objects and probably tampered by noises. Even when the image is free from noises, we’ll still have to deal with detecting the right Pap smear cell because in that single image we have not only Pap smear cells, but also non-Pap smear cells, for example hemoglobin (Hb), etc. The Pap smear cell is considered as having round shape, and so is Hb cells. The difference between them is that Pap smear cell has cytoplasm and not only having a nucleus like in Hb cells. So at first, we have to extract the cells with both nucleus and cytoplasm, while the others are ignored. The image is then inputted to GAs. The complete steps of processing an input image are as follows:

- **Pre-processing step:**
  - Crop the image input;
  - Convert the image into a gray-scale image;
  - Apply median filtering;
  - Segment the image using Fuzzy C-Means Clustering (FCM);
  - Extract the nucleus and cytoplasm objects;
  - Segment the nucleus object and reduce noises;
  - Segment the cytoplasm object and reduce noises;
  - Extract nucleus and cytoplasm features (maximum gray-level value, minimum and maximum width, minimum and maximum radius);

- **Processing step:**
  - Apply GAs;

- **Post-processing:**
  - Reduce unnecessary circles;
  - Detect wrong Pap smear cell;
  - Cut image into several single Pap smear cell.
The task of GAs in this application is in detecting circle area that is considered as Pap smear cell. An element in a population is represented by chromosome. In this application, the chromosome is defined by 3 parts or variables (x, y, r). The point (x,y) represents the center of a circle and the r is its radius. The maximum value of x is the number of pixels in each image row, and the maximum value of y is the number of pixels in each image column, and the maximum value of r is the maximum radius of a nucleus object. x is encoded in 9 bits, y is encoded in 10 bits, and r is encoded in r_bit (the value of r_bit is obtained from a preprocessing before the execution of the GAs method). The fitness function that is used in this GAs represents the number of nucleus-object pixels in a neighborhood window from each obtained circle. The neighborhood window is defined by the information (x, y, r). A fitness value is assigned to be 0 if the obtained neighborhood window is over the border of image size, or larger than the maximum number of nucleus-object pixels, or less than the minimum number of nucleus-object pixels, so that it will not be one of the solutions.

![Figure 6. An input image.](image1.png)

![Figure 7. A gray-scale image [3].](image2.png)

![Figure 8. Segmented image with nucleus objects [3].](image3.png)

![Figure 9. Output image with detected Pap smear cell using Gas [3].](image4.png)
Fig. 6 to Fig. 9 show typical processes done to an input image. The experimental results show that GAs is capable to identify the Pap smear cells (indicated by red circles in Fig. 9), even when the images contain much noises.

GAS IMPLEMENTATION ON NUCLEAR REACTOR CORE COST OPTIMIZATION

In this section, an implementation of GAs for cost-based optimization of a nuclear reactor core is presented [5]. In an optimization problem of a nuclear reactor core design, the objective is to minimize the radial power peaking factor in a three-enrichment zone reactor, considering restrictions on the average thermal flux, critical and sub-moderated. Instead of minimizing power-peaking factor, the literature has addressed the same problem using the minimization of the fuel and cladding material costs as the objective function, and of the average peak-factor as an operational constraint. The fitness function is given by the material costs of a single reactor cell:

\[
\text{Fitness} = \pi h \left( \rho_f s_f R_f^2 + \rho_c s_c (2R_f \Delta c + \Delta c^2) \right),
\]

where:
- \( h \) = cell height (cm), which is fixed and equal to 163 cm;
- \( \rho_f, \rho_c \) = fuel and cladding material densities (g/cm\(^3\));
- \( s_f, s_c \) = fuel and cladding material prices (US$/g);
- \( R_f \) = fuel radius (cm);
- \( \Delta c \) = cladding thickness (cm).

Solutions given by the GAs are then being sent to HAMMER Reactor Physics code [6] to obtain power-peaking, average thermal flux and the effective multiplication factor of that particular solution which is used as a preliminary fitness function.

SUMMARY

This paper summarizes two examples of the use of GAs in the application of old document enhancement and restoration [3] and in the application of cells identification in a Pap smear image [4], and one example of the potential use of GAs in cost-based optimization of a nuclear reactor core derived from the literatures [5,6]. These studies have reported that GAs is a potential method and has a good performance.

REFERENCES


DISKUSI

YUSTIKI KURNIATI

Mohon diterangkan tentang penerapan algoritma genetika dalam rancangan aplikasi pada optimisasi inti reaktor nuklir

HILDA DEBORAH

Penerapan algoritma genetika dalam rancangan aplikasi pada optimisasi inti reactor nuklir sebenarnya ada cukup banyak. Yang penulis sajikan pada makalah ini adalah salah satu contohnya dimana optimisasi dilakukan untuk meminimalisasi harga dan penggunaan material/bahan dengan tetap mempertimbangkan keamanan sebagai batasan dari usaha meminimalisasi harga pembuatan inti reaktor nuklir tersebut. Rumus yang digunakan untuk melakukan perhitungan optimisasi tersebut memasukkan harga dan jumlah material yang digunakan (dapat dilihat pada makalah). Solusi yang diperoleh kemudian diberikan kepada HAMMER Reactor Physics Code untuk dilihat apakah inti reactor nuklir dengan rancangan tersebut masih aman atau tidak.
TRI HAN.DOYO

Apa perbedaan GAs dengan software image processing yang ada dalam menyelamatkan dokumen tua? (Jika image processing software mampu melakukannya)

HILDA DEBORAH

GAs adalah sebuah algoritma yang pada kasus ini digunakan untuk menyelamatkan informasi dari gambar sebuah dokumen tua. Jika dilakukan perbandingan dengan software image processing yang sudah ada tentunya tidak relevan karena dua hal yang dibandingkan ini bukan berasal dari kategori yang sama. Pengembangan aplikasi yang telah dilakukan ini sendiri masih berada pada tahap awal, dimana kami juga masih mengeksplorasi banyak hal untuk kemudian baru bisa menjadikannya sebuah aplikasi yang khusus menangani restorasi dokumen tua.

ANTONIUS DARMA SETIAWAN

1. Apa justifikasi restorasi berhasil? Karena gambar dokumen aslinya tidak ada
2. Seberapa akurat penggunaan threshold untuk segmentasi?

HILDA DEBORAH

1. Menggunakan OCR (Optical Character Recognition). Dokumen tersebut sudah ada teks referensinya, kemudian citra hasil restorasi akan dibaca oleh OCR dan teks keluaran dari OCR akan dibandingkan dengan teks referensi dari citra yang bersangkutan untuk diukur nilai kemiripannya. Semakin tinggi nilai kemiripannya, restorasi dianggap berhasil.
2. Penggunaan threshold untuk segmentasi inti sel dan nukleus dipilih berdasarkan hasil observasi eksperimental terhadap beberapa gambar. Untuk gambar-gambar tersebut tentunya hasil threshold ini cukup akurat, akan tetapi untuk gambar lain yang mungkin belum dijadikan gambar uji, harus dilakukan pengamatan lagi.

DAFTAR RIWAYAT HIDUP

Nama : Hilda Deborah
Tempat & Tanggal Lahir : Jakarta, 13 September 1988
<table>
<thead>
<tr>
<th><strong>Pendidikan</strong></th>
<th>Sarjana S1 (Fakultas Ilmu Komputer, Universitas Indonesia, Class of 2006)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Riwayat Pekerjaan</strong></td>
<td>Research Assistant, Lab Pattern Recognition &amp; Image Processing, Fakultas Ilmu Komputer, Universitas Indonesia (Agustus 2010 - now)</td>
</tr>
<tr>
<td><strong>Keanggotaan</strong></td>
<td>-</td>
</tr>
<tr>
<td><strong>Kelompok</strong></td>
<td>-</td>
</tr>
<tr>
<td><strong>Makalah</strong></td>
<td>The Use of Genetic Algorithms in Several Applications</td>
</tr>
</tbody>
</table>