A system on chip for automatic karyotyping system☆

Faroudja Abid*, Latifa Hamami, Faiza Badache, Houssem Derdour

Département d'électronique, Ecole Nationale Polytechnique, Algiers, Algeria

A B S T R A C T

The Automatic karyotyping System is a computer-aided tool that automates the chromosome analysis and karyotyping processes, manually performed in most cytogenetic laboratories. Artificial neural networks (ANNs) have been widely used in chromosome classification due to their parallelism that reduces the computational complexity and time. However, existing classifiers are software-based, running on a computer that transforms the parallelism features of the ANNs into serial operations, thus significantly reducing their computing power. To efficiently address the above issue due to software implementation, we propose a Field-Programmable Gate Array-based System on Chip (SoC) architecture for human chromosome classification. The hardware implementation of such system can achieve the parallelism inherent to ANNs while reducing the power consumption and circuit size, thus the cost of designing such a system is reduced. The achieved part concerns the classification subsystem based on Kohonen neural network, which has been successfully tested on FPGA platform.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

The correlation between human diseases and chromosome abnormalities has been recognized since 1956, when Tjio and Levan [1] discovered that the number of human chromosomes was 46. Thus, chromosome analysis and karyotyping (a standard representation of the 23 pairs of chromosomes) are becoming key procedures for efficient genetic diseases diagnosis. However, traditional chromosome karyotyping processes are performed manually in most cytogenetic laboratories. In fact, the cytogeneticist must cut the individual chromosomes in the metaphase images (where the chromosomes appear as a succession of dark and light bands) and visually determine their centromere location in order to evaluate the arm length of each chromosome. The chromosomes of each class are paired and pasted together in decreasing order of size in the karyotype. The preliminary ordering is by length and centromere position for numerical abnormalities detection. Furthermore, additional time and effort are needed for structural abnormalities detection; in this case the cytogeneticist should revise the resulting karyotype according to the banding patterns (typically 400–800 visible bands).

The steps of centromere location, chromosome isolation, chromosome arm length evaluation, chromosome counting, banding observation, and pairing are performed manually. All these procedures require meticulous attention to details, which makes the process of manual karyotype repetitive and time-consuming. Computer-aided systems for chromosome classification are therefore highly needed to automate the chromosome analysis and help cytogeneticists to efficiently perform this time-consuming task.

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Associate Editor Dr. A. Isazadeh.

* Corresponding author.
  E-mail address: abidfaroudja@gmail.com (F. Abid).

Various computer-aided systems [2–6] have been developed to automate the chromosome classification procedure and karyotype establishment routine as well as the chromosomes abnormalities detection. Artificial neural networks have been widely adopted for chromosome classification, due to their processing capacity inherent to their parallel architectures. However, most of the works done in this area consist of software implementations running on a conventional computer that transforms the parallelism features of the ANNs to serial operations, thus reducing their computing power. On the other hand, commercially available automated systems for karyotyping are semi-automatic and still expensive [5,7].

To the best of the authors' knowledge, there are no hardware implementations for chromosome classification systems in the available literature; existing Automatic Karyotyping Systems (AKS) are software-based. Since a hardware implementation of such system can recreate the parallelism inherent to neural networks, we designed a system on chip for human chromosome classification, while reducing power consumption, circuit size, and the design cost.

Fully parallel modules can be achieved by Application Specific Integrated Circuits (ASICs) and Very-Large-Scale Integration (VLSI) circuits but it is expensive and time consuming to design such chips. We therefore targeted a Field-Programmable Gate Array (FPGA) platform to validate the functionality of the proposed architecture. This choice is justified by its processing capabilities and relatively reduced development cycle and cost. Also, the proposed architectural model is technology-independent and can be implemented on another technology. FPGA-based implementation of ANN, mainly for classification applications, allows parallelism and easy storage of input data and the synaptic weights. The motivations behind an FPGA-based SoC for automatic chromosome classification are mainly:

1. Parallelism: The software-based implementation of these systems running on a conventional computer transforms the parallelism features of the ANN into serial operations. Because the FPGA combines parallel computing, function integration of the system and low power dissipation, the hardware implementation of such system can achieve the parallelism required by ANN and significantly increase its speed.
2. Reduced processing time: One major improvement with the hardware implementation of such chromosomes classifier is in the training processing time, i.e., in the enhancement of the convergence time, which is an important issue. The network convergence time of software-based ANN classifiers is in the range of 20 s–50 s, whereas the convergence time of the proposed hardware ANN classifier is about 409 μs.
3. Reduced power consumption: FPGA-based SoC for human chromosome classification is indeed less power consuming than a software-based one.
4. Miniaturization: The SoC based on FPGA implementation allows a miniaturization of the chromosomes classifier thus, considerably reducing the space used by conventional computers or any related equipment.

This paper is organized as follows: In Section 2 we present a state-of-art-of ANN-based chromosome classification systems and feature extraction techniques. The Kohonen self-organizing map neural network is described in Section 3. An insight into the adopted methodology for the SoC-AKS elaboration is given in Section 4. Section 5 describes the proposed hardware architecture, mainly the Kohonen ANN-based classifier. The system prototyping results are exposed in Section 6. Finally, discussions and conclusions are given in Section 7 and Section 8, respectively.

2. Related works

This section undertakes a review of previous works related to ANN-based classifiers and feature extraction methods. Various ANN models have been used for the chromosome classification [8], mainly the Multi-Layer Perceptron (MLP) [4,9–12]. For example, in [12], Eskiizmirliler et al. proposed a hybrid structure that combines a supervised ANN (MLP) for numerical abnormalities detection and unsupervised ANN (Kohonen) for structural abnormalities detection. Besides the MLP, numerous types of ANNs have been adopted for chromosome classification namely, the probabilistic ANN [13,14], the wavelet neural network (WNN) [15], and the Bayesian classifier. A combined use of fuzzy logic and the Hopfield ANN has been also described in [16].

The MLP is the most widely adopted model and gives a maximum classification rate as shown in Table 1 (that summarizes some studies concerning the ANN-based chromosome classification systems). The classification rates obtained with this type of ANN range from 78% to 98% while the classification rates reported for other ANNs range from 61.30% to 96.98%. For example, a rate of about 68.18% for the probabilistic NN is reported in [14] and about 94% for the WNN in [15]. Apart from the ANN-based chromosome classifiers, other types of classifiers have been considered such as the support vector machine (SVM) [3]. This latter is relatively new compared to the ANN classifier; nevertheless, the comparison presented in [3] shows that the SVM and the ANN give comparable classification accuracy.

Accurate classification results are linked to a good chromosome features vector. These features are usually extracted from the medial axis. Therefore, the first step in extracting these features is the chromosome medial axis estimation. The Medial Axis Transform (MAT) [9] is the most used technique for medial axis detection and feature extraction. The mainly used algorithms for MAT are skeletonization [17] and thinning algorithm. The MAT technique is widely used because it preserves the shape properties of the chromosome. Medial axis estimation is particularly used for the chromosome length computation (calculated in the majority of studies as a sum of the medial axis pixels), the centromere (the thinnest region of the chromosome structure) identification, and the computation of the density profile.

Table 2 summarizes a set of studies on feature extraction methods, principally those related to the centromere localization, which is a widely used feature for classification. In addition to the MAT technique, other techniques have been applied