Chapter 7
A Reconfigurable Hardware for Subtractive Clustering*

Abstract. This chapter presents the development of a reconfigurable hardware for classification system of radioactive elements with a fast and efficient response. To achieve this goal is proposed the hardware implementation of subtractive clustering algorithm. The proposed hardware is generic, so it can be used in many problems of data classification, omnipresent in identification systems.

7.1 Introduction

Radioactive sources have radionuclides. A radionuclide is an atom with an unstable nucleus, i.e. a nucleus characterized by excess of energy, which is available to be imparted. In this process, the radionuclide undergoes radioactive decay and emits gamma rays and subatomic particles, constituting the ionizing radiation. Radionuclides may occur naturally but can also be artificially produced [1]. So, radioactivity is the spontaneous emission of energy from unstable atoms.

Correct radionuclide identification can be crucial to planning protective measures, especially in emergency situations, by defining the type of radiation source and its radiological hazard [2]. The gamma ray energy of a radionuclide is a characteristic of the atomic structure of the material.

When these emissions are collected and analyzed with a gamma ray spectroscopy system, a gamma ray energy spectrum can be produced. A detailed analysis of this spectrum is typically used to determine the identity of gamma emitters present in the source. The gamma spectrum is characteristic of the gamma-emitting radionuclides contained in the source [3].

This chapter introduces the development of a reconfigurable hardware for a classification system of radioactive elements that allow a rapid and efficient to be implemented in portable systems. our intention is to run the clustering algorithms in a portable equipment to perform the radionuclides identification. The clustering algorithms consume high processing time when implemented in software, mainly on processors of portable use, such as micro-controllers. Thus, a custom

* This chapter was developed in collaboration with Marcos Santana Farias.

N. Nedjah and L. de Macedo Mourelle, *Hardware for Soft Computing and Soft Computing for Hardware*, Studies in Computational Intelligence 529,
DOI: 10.1007/978-3-319-03110-1_7,

implementation suitable for reconfigurable hardware is a good choice in embedded systems, which require real-time execution as well as low power consumption.

The rest of this chapter is organized as follows: first, in Section 7.2, is demonstrated the principles of nuclear radiation detection. Later, in Section 7.3, we review briefly existing clustering algorithms and we concentrate on the subtractive clustering algorithm. In Section 9.2, we describe the proposed architecture for cluster centers calculator using the subtractive clustering algorithm. Thereafter, in Section 9.8, we present some performance figures to assess the efficiency of the proposed implementation. Last but not least, in Section 7.6, we draw some conclusions and point out some directions for future work.

7.2 Radiation Detection

The radioactivity and ionizing radiation are not naturally perceived by the sense organs of human beings and can not be measured directly. Therefore, the detection is performed by analysis of the effects produced by radiation as it interacts with a material.

There are three main types of ionizing radiation emitted by radioactive atoms: alpha, beta and gamma. The alpha and beta are particles that have mass and are electrically charged, while the gamma rays and x-rays are electromagnetic waves. The emission of alpha and beta radiation is always accompanied by the emission of gamma radiation. So most of the detectors is to gamma radiation. Gamma energy emitted by a radionuclide is a characteristic of the atomic structure of the material. The energy is measured in electronvolts (eV). One electronvolt is an extremely small amount of energy so it is common to use kiloelectronvolts (keV) and megaelectronvolt (MeV).

Consider, for instance, Cesium-137 (Cs137) and Cobalt-60 (Co60), which are two common gamma ray sources. These radionuclides emit radiation in one or two discreet wavelengths. Cesium-137 emits 0.662 MeV gamma rays and Cobalt-60 1.33 and 1.17 MeV gamma rays. These energy are known as decay energy and define the decay scheme of the radionuclide. Each radionuclide, among many others, has a unique decay scheme by which it is identified [1].

When these emissions are collected and analyzed with a gamma ray spectroscopy system, a gamma ray energy spectrum can be produced. A detailed analysis of this spectrum is typically used to determine the identity of gamma emitters present in the source. The gamma spectrum is characteristic of the gamma-emitting radionuclides contained in the source [3].

A typical gamma-ray spectrometry system (fig. 7.1) consists of a scintillator detector device and a measure system . The interaction of radiation with the system occurs in the scintillator detector and the measurement system interprets this interaction. The scintillator detector is capable of emitting light when gamma radiation transfers to him all or part of its energy. This light is detected by a photomultiplier optically coupled to the scintillator, which provides output to an electrical signal whose amplitude is proportional to energy deposited. For gamma radiation, the

most widely used scintillator is the Sodium Iodide crystal activated with thallium, NaI (Tl).

The property of these detectors provide an electrical signal proportional to the deposited energy spectrum allows the creation of the gamma energy spectrum by a radioactive element (histogram). To obtain this spectrum is used a multichannel analyzer or MCA. The MCA consists of an ADC (Analog to Digital Converter) which converts the amplitude of analog input in a number or channel. Each channel is associated with a counter that accumulates the number of pulses with a given amplitude, forming a histogram. These data form the energy spectrum of gamma radiation. As said, since different radionuclides emit radiation at different energy distributions, analyzing the spectrum can provide information on the composition of the radioactive source found and allow the identification.

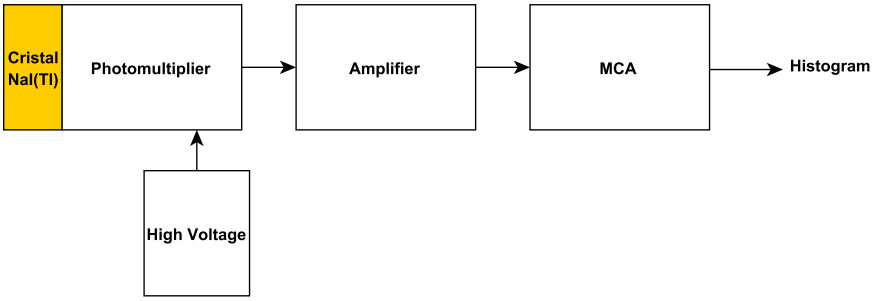

Fig. 7.1 Gama Spectrometry System - main components

Figure 7.2 shows a spectrum generated by simulation, to a radioactive source with of Cs137 and Co60. The x-axis represents the channels for a 12-bit ADC. In such a representation, 4096 channels correspond to 2.048 MeV in the energy spectrum. The first peak in channel 1324 is characteristic of Cs137 (0.662 MeV). The second and third peaks are energies of Co60.

The components and characteristics of a gamma spectrometry system (the type of detector, the time of detection , the noise of the high-voltage source, the number of channels, the stability of the ADC, temperature changes) can affect the formation of spectrum and quality of the result. For this reason it is difficult to establish a system for automatic identification of radionuclides, especially for a wide variety of these. Equipment that are in the market, using different algorithms of identification and number of radionuclides identifiable, do not have a good performance [2].

7.3 Clustering Algorithms

Clustering algorithms partition a collection of data into a certain number of clusters, groups or subsets. The aim of the clustering task is to group these data into clusters

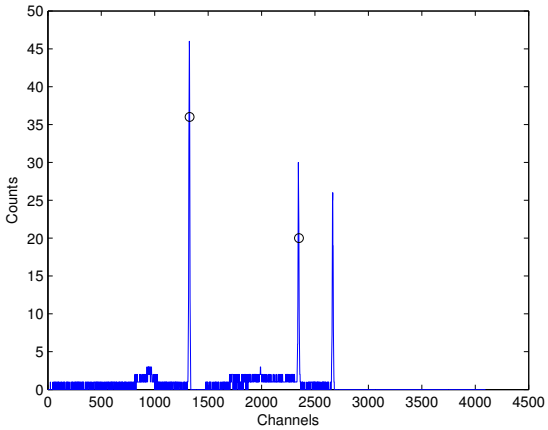

Fig. 7.2 Energy spectrum simulated by a source with Cs137 and Co60

in such a way that similarity between members of the same cluster is higher than that between members of different clusters. Clustering of numerical data forms the basis of many classification algorithms.

Various clustering algorithms have been developed. One of the first and most commonly used clustering algorithms is based on the Fuzzy C-means method (FCM). Fuzzy C-means is a method of clustering which allows one piece of data to belong to two or more clusters. This method was developed by Dunn [4] and improved by Hathaway [5]. It is commonly used in pattern recognition.

Yager and Filev [6] introduced the so-called *mountain function* as a measure of spatial density around vertices of a grid, showed in the function (7.1)

$$M(v_i) = \sum_{j=1}^{n} e^{-\alpha \|x_j - x_i\|^2}, \tag{7.1}$$

where $\alpha > 0$, M is the mountain function, calculated for the ith vertex v_i during the first step, N is the total number of data, which may be simple points or samples, that is assumed to be available before the algorithm is initiated. Norm $\|\times\|$ denotes the Euclidean distance between the points used as arguments and x_j is the current data point or sample. It is ensured that a vertex surrounded by many data points or samples will have a high value for this function and, conversely, a vertex with no neighboring data point or sample will have a low value for the same function. It should be noted that this is the function used only during the first step with all the set of available data. During the subsequent steps, the function is defined by subtracting a value proportional to the peak value of the mountain function. A very similar approach is the subtractive clustering (SC) proposed in [7]. It uses the so-called *potential* value defined as in (7.2).

$$P_i = \sum_{j=1}^{n} e^{-\alpha \|x_j - x_i\|^2}, \text{ where } \alpha = \frac{4}{r_a} \tag{7.2}$$

wherein, P_i is the potential-value i-data as a cluster center, x_i the data point and r_a a positive constant, called *cluster radius*.

The potential value associated with each data depends on its distance to all its neighborhoods. Considering (7.2), a data point or sample that has many points or samples in its neighborhood will have a high value of potential, while a remote data point or sample will have a low value of potential. After calculating potential for each point or sample, the one, say x_i^*, with the highest potential value, say P_i^*, will be selected as the first cluster center. Then the potential of each point is reduced as defined in (7.3). This is to avoid closely spaced clusters. Until the stopping criteria is satisfied, the algorithm continues selecting centers and revising potentials iteratively.

$$P_i = P_i - P_i^* e^{-\beta \|x_i - x_i^*\|^2}, \tag{7.3}$$

In (7.3), $\beta = 4/r_b^2$ represents the radius of the neighborhood for which significant potential revision will occur. The data points or samples, that are near the first cluster center, say x_i^*, will have a significantly reduced density measures. Thereby, making the points or samples unlikely to be selected as the next cluster center.

The subtractive clustering algorithm can be briefly described by the following 4 main steps:

- Step 1: Using (7.2), compute the potential P_i for each point or sample, $1 \leq i \leq n$;
- Step 2: Select the data point or sample, x_i^*, considering the highest potential value, P_i^*;
- Step 3: Revise the potential value of each data point or sample, according to (7.3);
- Step 4: If $maxP_i \leq \varepsilon P_i^*$, wherein ε is the reject ratio, terminate the algorithm computation; otherwise, find the next data point or sample that has the highest potential value and return to Step 3.

The main advantage of this method is that the number of clusters or groups is not predefined, as it is in the fuzzy C-means method, for instance. Therefore, this method becomes suitable for applications where one does not know or does not want to assign an expected number of clusters *á priori*. The cluster estimates obtained by the subtractive clustering can be used to initialize iterative optimization-based clustering methods and as well as the set of rules used in fuzzy clustering methods.

7.4 Proposed Architecture

This section provides an overview of the macro-architecture and contains information on the broad objectives of the proposed hardware. The hardware implements the subtractive clustering algorithm. The subtractive clustering algorithm was briefly explained in the previous section.

The implementation of this algorithm in hardware is the main point is to develop a classification system of radioactive elements. For referencing, this hardware it will call HSC, hardware to subtractive clustering. This hardware processes all the arithmetic computation, described in the section above, to calculate the potential of each point in the subtractive clustering algorithm.

The other component of this macro-architecture will be called SLC, component to storage, loading and control, which provides to the HSC the set of samples for the selection of cluster centers and stores the results of the calculated potential of each sample. This component also has the controller of the HSC. Figure 7.3 shows the components of the described macro-architecture.

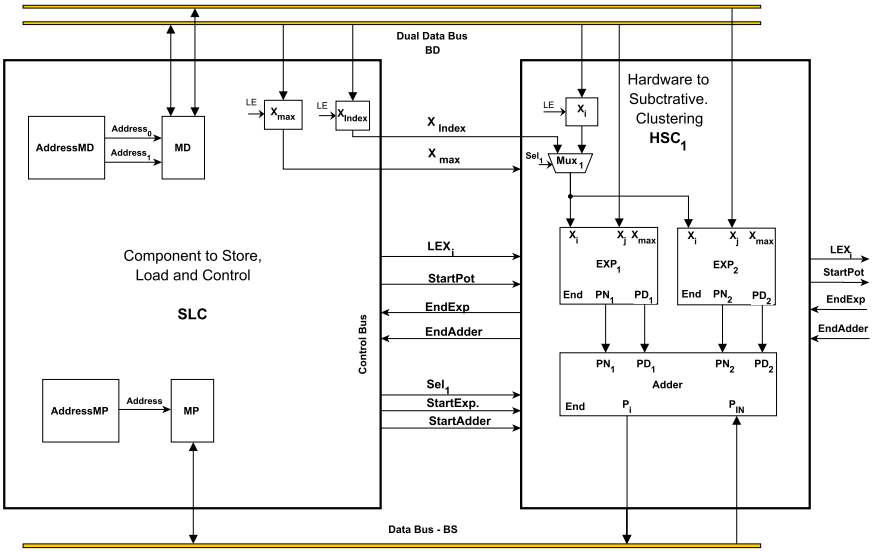

Fig. 7.3 Macro-architecture components - SLC e HSC

The SLC is a controller based on state machine. It includes a dual port memory MD that provides the data that has to be clustered and memory MP that allows for the bookkeeping of the potential associated with each clustered data. The registers X_{max}, X_i and X_{Index} maintain the required data until component EXP_1 and EXP_2 have completed the related computation. We assume the X_{max} value is available in memory MD at address 0. The X_{max} is the biggest value found within the data stored in MD. This register is used to the data normalization.

The two EXP components, inside HSC, receive, at the same time, different x_j values from the dual port memory MD. So the two modules start at the same time

and thus, run in parallel. After the computation of $e^{-\alpha\|x_i-x_j\|^2}$ by EXP_1 and EXP_2, component ADDER sums and accumulates the values provided at its input ports. This process is repeated until all data x_j, $1 \leq j \leq N$, are handled. So, this computation yileds the first P_i value to be stored in memory MP. After that, the process is repeated to compute the potential values of all data points in memory MD. At this point, the first cluster center has been found.

The SLC component works as a main controller of the process. Thus, the trigger for initiating the processing components EXP_1 and EXP_2 occurs from the signal *StartExp* sent by SLC. The component SLC has a dual-port memory MD which stores the samples / points to be processed. Memory MD allows the two components (EXP_1 and EXP_2) receiving a sample to calculate the exponential value and thus can operate in parallel . This sample for each component EXP are two distinct values x_j from two subsequent memory addresses.

The proposed architecture allows the hardware to subtractive clustering HSC can be scaled by adding more of these components in parallel to the computation of the factors $e^{-\alpha||x_j-x_i||^2}$. This provides greater flexibility to implement the hardware. Figure 7.4 shows how new components HSC are assembled in parallel.

Each component HSC calculates in parallel the potential of a point i, the value P_i of the function 7.3. For this reason each module (HSC) must to receive and record a value of x_i to work during the calculation of the potential of a point. Since these values are in different adrress of the memory, this registry value x_i has to be done at different time because the memory can not have your number of ports increased as the number of components HSC is increased. To be not necessary to increase the number of control signals provided by the component SLC when new components HSC are added, the component HSC itself has to send some control signals for the thereafter.

These signs are to load the value x_i (LEX_i) and start the reduction potential of each point (*StartPot*), as showed in 7.3. Moreover, each component HSC should receive the signal *EndAdd* which indicates the end of the operation on the component

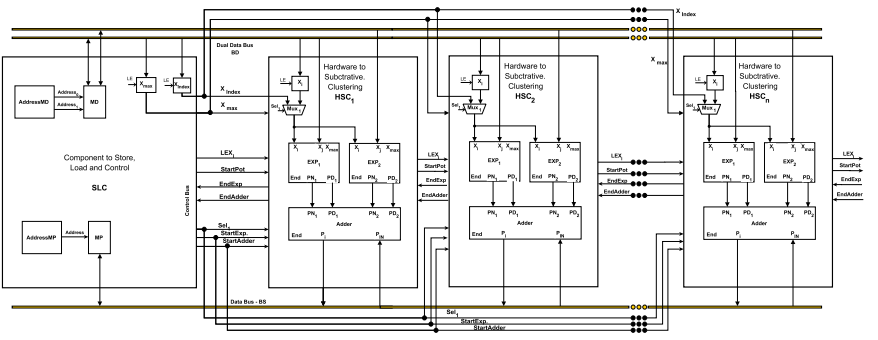

Fig. 7.4 Macro-architecture with HSC components in parallel

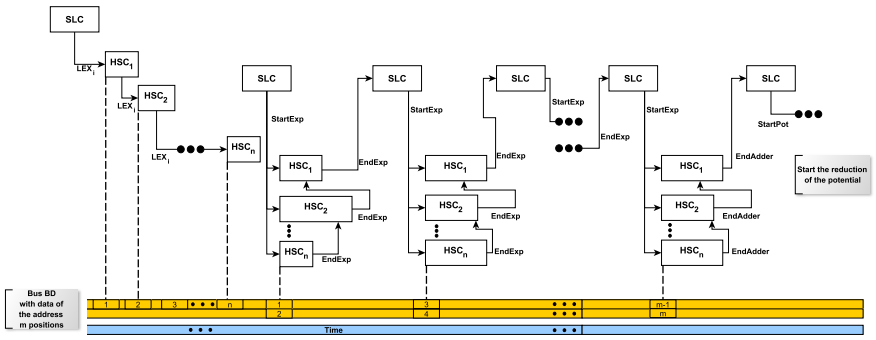

Fig. 7.5 Control signals with scaled architecture

ADDER of the thereafter component HSC. This ensures that the main control (SLC) only receive these signals after all the components of the HSC in parallel complete their transactions at each stage, allowing the hardware can be reconfigured without change in the main control. Figure 7.5 shows the effect of this scaling, simulating different processing times between the HSC.

The n components HSC, implemented in parallel, compute the potential of n points of the set of samples. As explained earlier, the record value of x_i, to be used in the calculation of the potential it has to be done in time different. It is shown in figure 7.5 that the first component HSC receives the signal LEX_i from SLC control and after registering it x_i, it sends the signal LEX_i for HSC thereafter. Only after all of the HSC to have recorded its value x_i, the signal to start the components EXP ($StartExp$) is sent with the first pair of values x_j in the dual bus BD.

Fig. 7.7 shows the architecture of the module EXP_1 and EXP_2 that permits the calculation of the exponential value $e^{-\alpha\|x_i - x_j\|^2}$. The exponential value was approximated by a second-order polynomial using the least-squares method [8] . Moreover, this architecture computes these polynomials and all values were represented using fractions, as in (7.4).

$$e^{-\alpha\|x\|} = \frac{N_a}{D_a}\left(\frac{N_v}{D_v}\right)^2 + \frac{N_b}{D_b}\left(\frac{N_v}{D_v}\right) + \frac{N_c}{D_c} \tag{7.4}$$

wherein, factors $\frac{N_a}{D_a}$, $\frac{N_b}{D_b}$ and $\frac{N_c}{D_c}$ are some pre-determined coefficients. $\frac{N_v}{D_v}$ is equivalent to variable (αx) in the FPP representation. For high precision, the coefficients were calculated within the range [0, 1[, [1, 2[, [2, 4[and [4, 8]. These coefficients are shown respectively in the quadratic polynomials of (7.5).

$$e^{-(\alpha x)} \cong \begin{cases} P_{[0,1[}\left(\frac{N_v}{D_v}\right) = \frac{773}{2500}\left(\frac{N_v}{D_v}\right)^2 - \frac{372}{400}\left(\frac{N_v}{D_v}\right) + \frac{9953}{10000} \\ P_{[1,2[}\left(\frac{N_v}{D_v}\right) = \frac{569}{5000}\left(\frac{N_v}{D_v}\right)^2 - \frac{2853}{5000}\left(\frac{N_v}{D_v}\right) + \frac{823}{1000} \\ P_{[2,4[}\left(\frac{N_v}{D_v}\right) = \frac{67}{2500}\left(\frac{N_v}{D_v}\right)^2 - \frac{2161}{10000}\left(\frac{N_v}{D_v}\right) + \frac{4565}{10000} \\ P_{[4,8[}\left(\frac{N_v}{D_v}\right) = \frac{16}{10000}\left(\frac{N_v}{D_v}\right)^2 - \frac{234}{10000}\left(\frac{N_v}{D_v}\right) + \frac{835}{10000} \\ P_{[8,\infty[}\left(\frac{N_v}{D_v}\right) = 0 \end{cases} \tag{7.5}$$

The accuracy of these calculated values, i.e. the introduced error is no more 0.005, is adequate to properly obtain the potential values among the data provided during the process of subtractive clustering. The absolute error introduced is shown in Fig. 7.6. Depending on the data, this requires that the number of bits to represent the numerator and denominator have to be at least twice the maximum found in the data points provided.

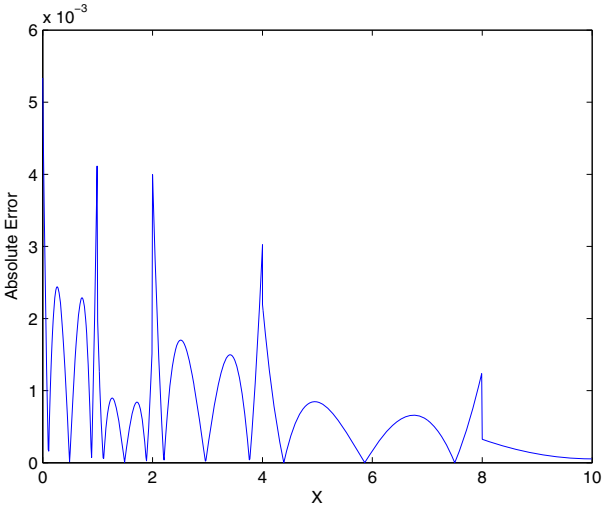

Fig. 7.6 Absolute error introduced by the approximation computation

The architecture of the Fig. 7.7 presents the micro-architecture of components EXP_1 and EXP_1. It uses four multipliers, one adder/subtracter and some registers. These registers are all right-shifters. The controller makes the adjustment of the binary numbers with shifts to the right in these registers in order to maintain the frame

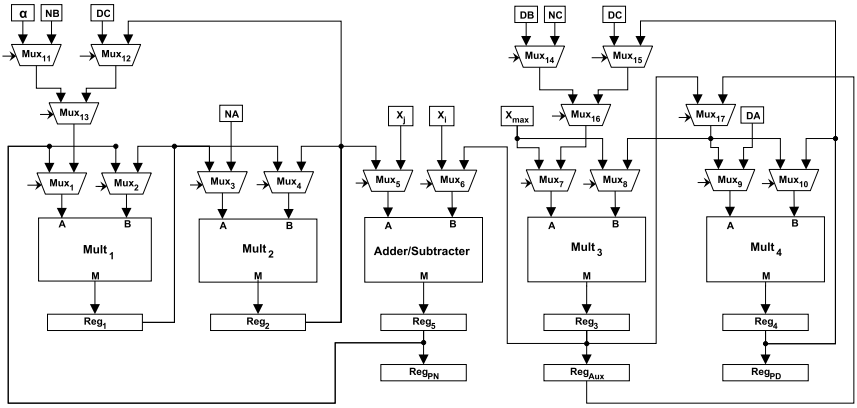

Fig. 7.7 Architecture of EXP Modules to compute the exponential $e^{-\alpha\|x_i-x_j\|^2}$

of binary numbers after each operation. This is necessary to keep the results of multiplication with the frame of bits used without much loss of precision. The closest fraction is used instead of a simple truncation of the higher bits of the product.

In this architecture, multipliers MULT_1, MULT_2, MULT_3 and MULT_4 operate in parallel to accelerate the computation. The state machine in the controller triggers these operations and controls the various multiplexers of the architecture. The computation defined in (7.4) is performed as described hereafter.

- Step 1: Compute $NV \times NV$, $NB \times NV$, $DV \times DV$ and $DB \times DV$;
- Step 2: Right-shift registers to render the frame of bits to the original size and in parallel with that, compute $A = NA \times NV \times NV$, $C = NB \times NV \times DC$, $D = DB \times DV \times NC$ and $E = DB \times DV \times DC$;
- Step 3: Add of $C + D$ and, in parallel with that, compute $B = DA \times DV \times DV$;
- Step 4: Add $\frac{A}{B} + \frac{C+D}{E}$.

7.5 Performance Results

The data shown in figure 7.2 were obtained using a simulation program called *Real Gamma-Spectrum Emulator*. These data are in spreadsheet format of two columns, where the first column corresponds to the channel and the second is the number of counts accumulated in each channel. To validate the method chosen (subtractive clustering), the algorithm was implemented with Matlab, using the simulated data. As seen in the introduction, these data simulate a radioactive source consists of Cs137 and Co60. To apply the subtractive clustering algorithm in Matlab data provided by the simulation program needed to be converted into one-dimensional data in one column. For example, if channel 1324 to accumulate 100 counts, means that the value 1324 should appear 100 times as input. only in this way the clustering algorithm is able to split the data into subgroups by frequency of appearance. In a

real application this data would be equivalent to the output of AD converter of a gamma spectrometry system, as shown in the introduction.

In the spectrum of Fig. 7.2, one can see three peaks. The first one in the channel 1324 is characteristic of Cs137 (0.662 MeV). The second and third peaks correspond the energy of Co60. The circular marks near the first and second peaks show the result of applying the subtractive clustering algorithm on the available data with Matlab software. These circular marks are center of the found clusters. These found clusters are very near (one channel to the left) of the signal peaks, the expected result. With the configuration to the algorithm in Matlab, the third peak was not found. This result can change with an adjust of the radius r_a in 7.2. This is enough to conclude that the data provided belongs to a radioactive source with Cs137 and Co60 and the subtractive cluster method can be used to identify these radionuclides.

As the proposed architecture is based on the same algorithm, is expected to find the same result. The initial results show that the expected cluster center can be identified as in Matlab specification. The hardware takes about 12660 clock cycles to yield one sum of exponential values ($\sum_{j=1}^{n} e^{-\alpha \|x_i - x_j\|^2}$). Considering the one hundred points in the avaiable data set of the case study, the identification of the first cluster center would take ten times that amount, i.e. about 126600 clock cycles. However, finding the center of the second cluster is faster. It should take about 13000 clock cycles. This result can change with the data and depends of the amount of adjustment required to the right in the shift registers during the process. The simulation results of an instance of this process is shown in Fig. 7.8.

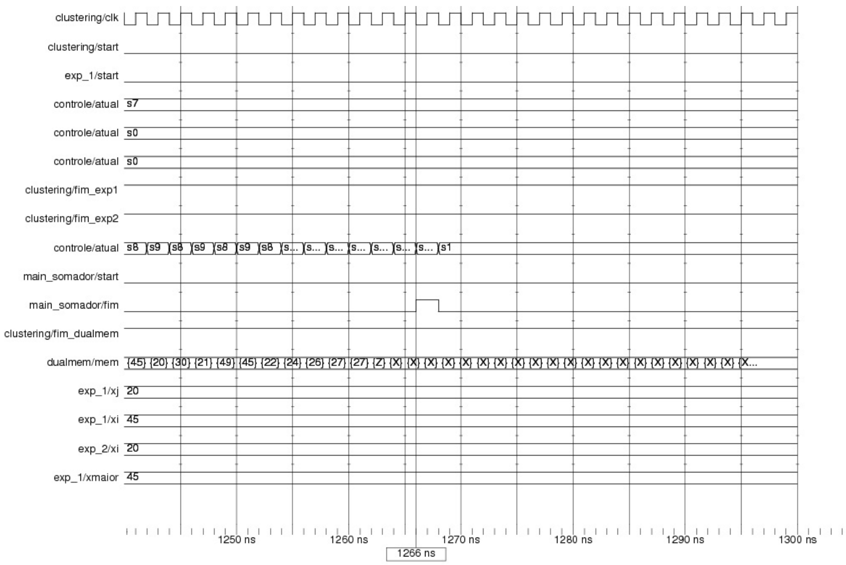

Fig. 7.8 Simulao de forma de onda de deslocamentos a direita para ajuste no nmero de bits do resultado

7.6 Summary

This chapter describes the implementation of subtractive clustering algorithm in hardware. The results shows the expected cluster center can be identified with a good efficiency. In data from the simulation of signals of radioactive sources, after conformation of the signal and its conversion into digital , the cluster center represents the points that characterize the energy provided by a simulated radionuclides. The identification of these points can sort the radioactive elements present in a sample. With this hardware it was possible to identify more than one cluster center, which would recognize more than one radionuclide in radioactive sources.

These results reveal that the proposed hardware can be used to develop a portable system for radionuclides identification. This system can be developed and enhanced integrating the proposed hardware with a software to be executed by a processor inside the FPGA, bringing reliability and faster identification, an important characteristics for these systems. Following this work, we intend to develop a software-only implementation using an embedded processor or a micro-controller to compare it with the hardware-only solution.

References

1. Knoll, G.F.: Radiation Detection and Measurement. John Wiley and Sons, New York (1989)
2. Performance Criteria for Hand-held Instruments for the Detection and Identification of Radionuclides. ANSI Standard N42.34 (2003)
3. Gilmore, G., Hemingway, J.: Practical Gamma Ray Spectrometry. John Wiley and Sons (1995)
4. Dunn, J.C.: A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. Journal of Cybernetics 3, 32–57 (1973)
5. Hathaway, R., Bezdek, J., Hu, Y.: Generalized fuzzy C-means clustering strategies using Lp norm distances. IEEE Transactions on Fuzzy Systems, Proc. of SPIE Conf. on Application of Fuzzy Logic Technology, pp. 246–254 (1993)
6. Yager, R.R., Filev, D.: Learning of Fuzzy Rules by Mountain-Clustering. In: Proc. IEEE Internat. Conf. on Fuzzy Systems, pp. 1240–1245 (1994)
7. Chiu, S.L.: A Cluster Estimation Method with Extension to Fuzzy Model Identification. In: Proc. IEEE Internat. Conf. on Fuzzy Systems, pp. 1240–1245 (1994)
8. Rao, C., Toutenburg, H., Fieger, A., Heumann, C., Nittner, T., Scheid, S.: Linear Models: Least Squares and Alternatives, New York. Springer Series in Statistics (1999)