Similarity Quantification Analysis to Detect Suffering Fetus

Iulian Voicu, Jean-Marc Girault

SIMILARITY QUANTIFICATION ANALYSIS TO DETECT SUFFERING FETUS

Iulian Voicu, Jean-Marc Girault

Université François Rabelais de Tours, UMR - S930, Tours, France
Inserm U 930, Tours, France

ABSTRACT

To characterize a dynamical system, a number of descriptors or invariants can be used. Here we proposed to review and generalize the analysis of recurrence by introducing the concept of similarity. In this study, we showed that it was possible to distinguish the fetal heart rate of healthy from suffering fetus using the similarities of a time series, while the analysis of recurrence does not allowed it.

Index Terms— dynamical system, recurrence plot, similarity plot, RQA parameters, fetal heart rate

1. INTRODUCTION

Most of the nonlinear dynamical systems are characterized by differential equations. These equations usually are not known and the only available information is often the output signal of the system. The lack of knowledge of differential equations leads to two strategies to characterize the system.

The first strategy is to calculate parameters using directly the output signal of the system. In this case, several parameters such as Lyapunov exponents, dimensions and entropies [1, 2, 3, 4] have been calculated to describe the different operating modes (periodic, random or chaotic) of the nonlinear system. Among them, the entropies count identical patterns of length m in the signal. The identical patterns are the patterns identical within a tolerance r.

The second strategy is to use the output signal to obtain the phase space of the system and then to calculate parameters. Although the real phase space can not be reconstructed without the knowledge of the differential equations, it is possible to construct an equivalent phase space (m-dimensional) [5]. In this equivalent phase space the topological properties of the attractor are preserved.

To characterize the phase space Eckman [6] introduced the recurrence plot (RP) method. The RP has enabled a two dimensional (matrix) representation of the recurrent information in the D-dimensional equivalent phase space. An element of the recurrence matrix is 1 if two points belonging to the trajectories in this space are recurrent (are confined within a sphere of radius r).

Different structures such as isolated points or lines, characterize the recurrence matrix [7]. To quantify these structures Zbilut and Webber [8, 9] proposed different parameters known as recurrence quantification analysis (RQA) parameters.

The analysis of the RQA parameters is important because the parameters change according to the periodic, random or chaotic regime of the system.

Because a pattern of m consecutive non-constant samples reflects a local variation of the signal, looking for recurrences involves to search in the signal only those variations with almost the same amplitudes. The disadvantage of this definition is that it removes identical variations of the signal located at different levels.

To quantify all similar variations regardless the signal level, we proposed a new approach based on similarity patterns. This approach that generalized the notion of recurrence to that of similarity was novel and simple to implement, since it applied directly to the output of a dynamical system.

The paper is structured as follows: section 2 presents the recurrence and modified similarity matrices together with the RQA parameter. The results obtained applying the RQA parameter in the two cases were presented in section 3, while in section 4 we discuss them. Finally, we conclude our study in section 5.

2. METHODS

In this section we presented the steps to obtain the recurrence and the similarity matrices. The steps to find the recurrence matrix, noted in this paper $\mathbf{RP}_m^D$, are described in section 2.1, while the steps corresponding to similarity matrix, noted $\mathbf{SP}_m^D$, are presented in section 2.2. At the end of this section we present the parameters used to distinguish between healthy and suffering fetus.

2.1. Recurrence plot matrix

The recurrence matrix was constructed by using patterns (noted $\mathbf{X_i}$) composed from m consecutive samples of the time series of length N. Thus we considered a phase space of dimension D = 1. Mathematically, the recurrence matrix computed in case of patterns of size m can be expressed as:

$$\mathbf{RP}_m^1(r,i,j) = \Theta(r - \|\mathbf{X}_i - \mathbf{X}_j\|), \quad \mathbf{X}_i, \mathbf{X}_j \in \mathcal{R}^m, \quad (1)$$

where $\Theta(\cdot)$ is the Heaviside function, $\|\cdot\|$ is the L_∞ norm, r is a fixed tolerance and $i, j = 1, \ldots, N-m+1$. This matrix is a binary matrix. To fill this matrix we made:

1. for each $i = 1, \ldots, N-m+1$ we formed the vectors;

$$\mathbf{X}_i = [x(i), x(i+1), \ldots, x(i+m-1)], \quad (2)$$

where $x(i)$ was the i^{th} sample of the signal, and $\mathbf{X}_i$ was the pattern composed of m samples.

2. we computed the maximum value of the m-dimensional absolute difference vector;

$$d(\mathbf{X}_i, \mathbf{X}_j) = max(|\mathbf{X}_i - \mathbf{X}_j|), \quad (3)$$

3. using a tolerance r, we assigned for each pair (i, j), $i, j = 1, \ldots, N$;

$$\mathbf{RP}^1_m(r, i, j) = \begin{cases} 1, & d(\mathbf{X}_i, \mathbf{X}_j) \leq r, \\ 0, & d(\mathbf{X}_i, \mathbf{X}_j) > r, \end{cases} \quad (4)$$

2.2. Similarity plot matrix

The key idea to obtain the similarity matrix was to remove the distance between the means of the compared patterns.

1. for each $i = 1, \ldots, N-m+1$ we formed the vectors;

$$\mathbf{X}_i = [x(i), x(i+1), \ldots, x(i+m-1)], \quad (5)$$

where $x(i)$ was the i^{th} sample of the signal, and $\mathbf{X}_i$ was the pattern composed of m samples.

2. we computed

$$\mathbf{X}_i^{'} = \mathbf{X}_i - \overline{M}, \quad (6)$$

where $\overline{M}$ was the mean of the vector $\mathbf{X}_i$

3. we computed the maximum value of the m-dimensional absolute difference vector;

$$d(\mathbf{X}_i^{'}, \mathbf{X}_j^{'}) = max(|\mathbf{X}_i^{'} - \mathbf{X}_j^{'}|), \quad (7)$$

4. using a tolerance r, we assigned for each pair (i, j), $i, j = 1, \ldots, N$;

$$\mathbf{SP}^1_m(r, i, j) = \begin{cases} 1, & d(\mathbf{X}_i^{'}, \mathbf{X}_j^{'}) \leq r, \\ 0, & d(\mathbf{X}_i^{'}, \mathbf{X}_j^{'}) > r, \end{cases} \quad (8)$$

To quantify the similarities present in the similarity matrix we introduced the notion of similarity quantification analysis (SQA).

2.3. SQA and RQA parameters

One of the parameter proposed by Zbilut [8] was the *recurrence rate*. This parameter was nothing else that the number of recurrent/similar points in the recurrence/similarity matrix. We computed this parameter according with the relation

$$RR = \frac{1}{N^2} \sum_{i,j=1}^{N} RP^1_m(r, i, j), \quad (9)$$

for the $\mathbf{RP}^1_m$ matrix, and with the relation

$$SR = \frac{1}{N^2} \sum_{i,j=1}^{N} SP^1_m(r, i, j) \quad (10)$$

for the $\mathbf{SP}^1_m$ matrix. We called this parameter the *similarity rate*.

3. RESULTS

In this section we exemplified in section 3.1 the difference between the recurrence and similarity matrix using a sine wave, while in section 3.2 we present the results obtained on fetal heart rate signals.

3.1. Recurrence and similarity matrices applied to the sinusoidal signal

Firstly we illustrated the difference between the two matrices $\mathbf{RP}^1_m$ et $\mathbf{SP}^1_m$, calculated in the case of a sine wave. Towards a better understanding of our approach we used for the sine wave, shown in figure 1, a frequency of 10 Hz sampled at 250 Hz.

We remember that finding the recurrence patterns means to find all patterns located at the same level of the signal. For example, the recurrence patterns of the plus sign-pattern (m = 2) in figure 1 appears at 26^{th} (asterisk pattern), 51^{th} and 76^{th} sample. On the other hand, we can see in the same figure 1 that the pattern located at the origin of the signal (sign

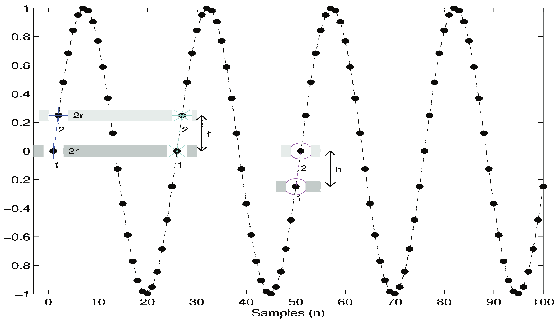
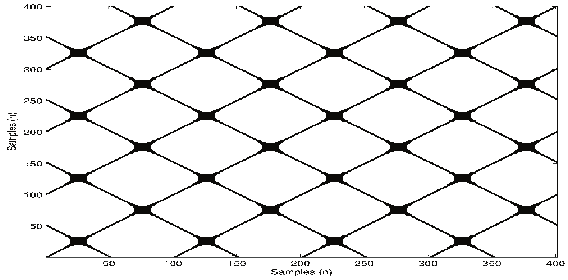
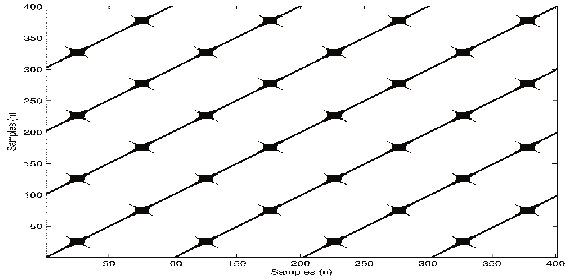


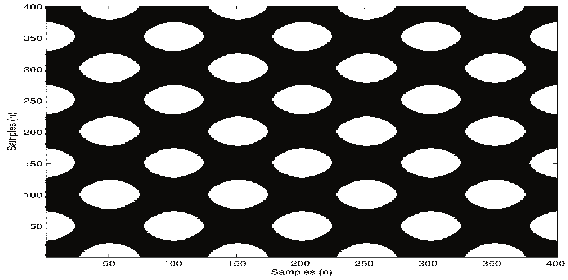
Fig. 1: The sinusoidal signal. Example of a pattern m=2 placed at different levels in the signal.



(a)



(b)



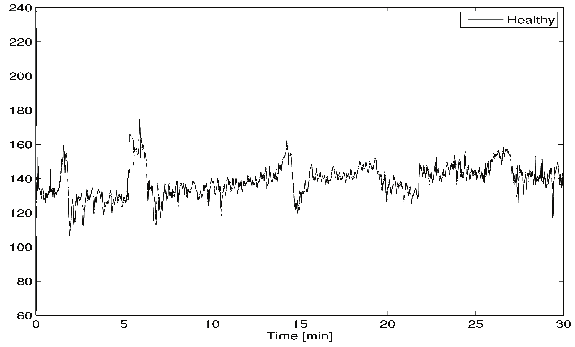
(c)

Fig. 2: (a) The recurrence plot matrix using m = 1 (RP_1^1); (b) The recurrence plot matrix using m = 3 (RP_3^1); ; (c) the similarity plot matrix using m = 3 (SP_3^1). The matrices were constructed for a value of r equal to 0.1 from the standard deviation of the sinusoidal signal.

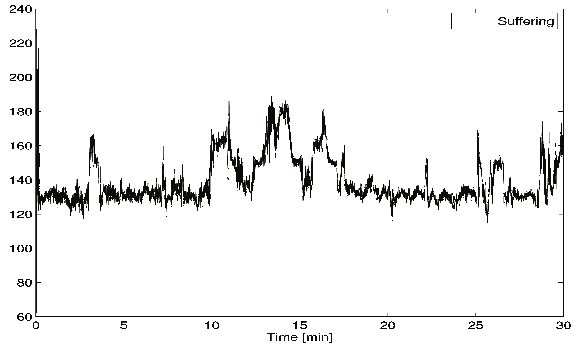
plus-pattern) had a similar pattern (circle-pattern), which was located at a different level in the signal.

We applied the equations 2-4 for two cases: m = 1, m = 3. The tolerance r was chosen equal to 0.1 from the value of the standard deviation of the sinusoidal signal. To obtain a better resolution of the figures that illustrated the recurrence and similarity matrices we increased the sample frequency at 1000 Hz for the sinusoidal signal of 10 Hz.

As we can see in figure 2a the recurrence matrix for m = 1 is formed by the diagonal and anti-diagonal lines structures. The main diagonal of the matrix indicates that a pattern was identical (recurrent) with itself. The distance between the lines that are parallel with the main diagonal line indicates the frequency of the signal, which in our case was 100 samples. Each column i of the matrix identified recurrences of the pattern that started at i^{th} sample.



(a)



(b)

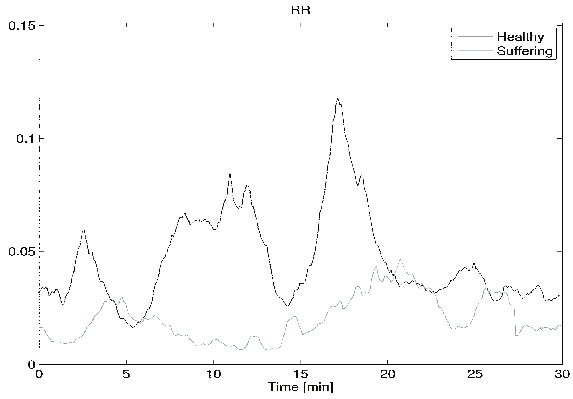
Fig. 3: The fetal heart rate: (a) of a healthy fetus; (b) of a suffering fetus.

We increased the size of the pattern and we used now m = 3 in the equations 2-4. We observed that when we computed the recurrence matrix for the pattern with m = 3, the anti-diagonal vanished.

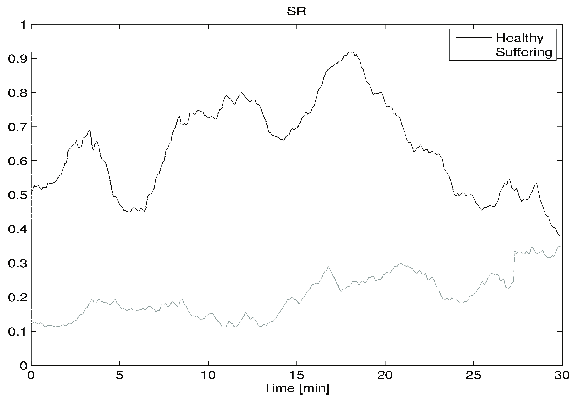
Next step was to compute the similarity matrix using equations 5-8 for m = 3 and the same values r. We illustrated the similarity matrix in figure 2c. The similarity matrix has more points coded in black, which means that the number of similar patterns was larger (we found vertical lines for each pattern) than the recurrent patterns from a sinusoidal signal.

3.2. Recurrence and similarity matrices applied to fetal heart rate

We monitorized the cardiac rhythm of the fetuses over 30 minutes using the ultrasound Actifoetus device [10]. The Doppler signals were acquired at CHRU "Bretonneau" Tours, France. The consent of each patient was obtained and the study was approved by the ethics committee of the Clinical Investigation Centre for Innovative Technology of Tours (CIC-IT 806 CHRU of Tours). Patients were older than eighteen years and pregnancy was singular. The recordings were made during the twenty-fifth and fortieth gestational weeks. The fetuses with congenitally malformations were not considered by



(a)



(b)

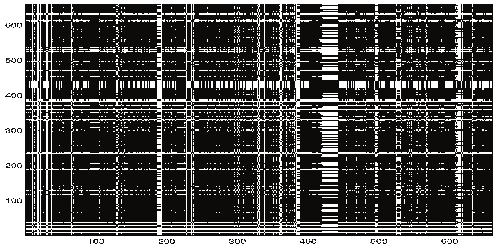
Fig. 4: (a) The recurrence rate RR and (b) the similarity rate SR parameters computed for a pattern m = 3 in the case of a healthy and suffering fetus.

the physicians in the group of suffering fetuses. In the group of suffering fetuses were included all fetuses for which a reversed end diastolic flow velocity in the fetal umbilical artery was present.

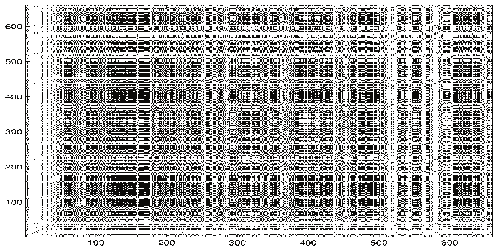
The figure 3 shows the heart rates of a healthy and suffering fetus, respectively. The rhythm was evaluated every 250 ms, obtaining 7200 values for a period of 30 minutes, let 240 values per minute.

We computed the recurrence and the similarity matrices for the fetal heart rates series. However, before computing the matrices a normalization of the heart rates was made, thus both of them were converted into zero mean variables with unity variances. Because the signal is non-stationary we applied a moving window technique by fixing a three minutes value for the window size. The $\mathbf{RP}_3^1$ and $\mathbf{SR}_3^1$ were evaluated for each window (720 values per window) together with the recurrence rate (RR) and similarity rate (SR) parameters (the equations 9 et 10).

The evolution in time of the RR parameter for both healthy and suffering fetus was illustrated in figure 4a, while



(a)



(b)

Fig. 5: (a) The RR and (b) the SR parameters computed for a pattern m = 3 in case of a healthy and suffering fetus.

the evolution of SR parameter in figure 4b. We used m = 3 for the pattern dimension and r = 0.1.

We can observe in figure 4a that the parameter RR ranged in the interval 0.006-0.118 when it was evaluated for the healthy and suffering fetus. The two curves could not be separated in this case.

In contrast, when we evaluated the parameter SR from the similarity matrix, the two curves shown in figure 4b increased, but the increase was more pronounced in the case of the healthy fetus and a clear separation between the two curves was possible. We found that for the healthy fetus the SR curve varied in the interval [0.37, 0.92], whereas in the case of suffering fetus the curve varied in the interval [0.11, 0.35]. In this case we were able to separate the two fetus.

4. DISCUSSIONS

The $\mathbf{RP}_3^1$ matrix quantified the percentage of "similar" patterns in the signal that had the same variation. In comparison, the $\mathbf{SP}_3^1$ matrix quantified the percentage of patterns with the same variation and modulated in amplitude according to Groth [11]. In contrast to symbolic patterns considered by Groth, the amplitude modulation of the similar patterns used to compute the similar matrix is more restrictive. If the approach of Groth took into account only the sign of the variation between consecutive elements of the pattern, we also considered the value of this variation (the value of the derivative).

The more pronounced increasing of the parameter SR in the case of the healthy fetus means that the same kind of variation (similar patterns) is more present comparing with

the case of the suffering fetus. The increased number of the similar patterns in the case of healthy fetus was determined by the increased number of different patterns and by the increased probabilities of the similar patterns. A higher diversity and probability of similar patterns is the characteristics of a chaotic system.

On the other hand we observed, see in figure 5a-5b, that the decreasing of the parameter SR in the case of suffering fetus was the consequence of a loss or fragmentation of the vertical lines in the matrix $\mathbf{SP}_3^1$. According to Marwan [7], the vertical lines of the matrix characterize chaos-chaos transitions. The disappearance of the vertical lines can signify the loss of the chaotic nature of the heart.

If the diversity of similar patterns decreases or there is a too much diversity of very unlikely patterns, then we can imagine that the operating mode of the fetal heart tends to a random one. On the other hand, if diversity decreases but the occurrence of the patterns is important, the operating mode of the heart tends to a periodic regime.

In the end of this section we mention that the similarity matrix could be applied only for patterns of length $m \geq 2$. For the pattern m=1 we could not applied this method because all the patterns became similar after the subtraction of the mean.

5. CONCLUSIONS

We modified the definition of the recurrence of the patterns. Using the new definition we calculated the similarity matrix. From this matrix we were able to discriminate the heart rate of a healthy from a suffering fetus while it was not possible using the recurrence matrix. A loss of the chaotic nature of the fetal heart, measurable by the similarity matrix, could explain the discrimination.

6. REFERENCES

[1] S. M. Pincus, "Approximate Entropy (ApEn) as a Measure of System Complexity," in *Proc of Nat Acad Sciences*, 1991, vol. 88, pp. 2297–2301.

[2] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman, "Sample Entropy Analysis of Neonatal Heart Rate Variability," *Am J Phys - Regul, Int & CompPhys*, vol. 283, no. 3, pp. R789–R797, 2002.

[3] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale Entropy to Distinguish Physiologic and Synthetic RR Time Series," *Comput Cardiol*, vol. 29, pp. 137–140, 2002.

[4] C. Bandt and B. Pompe, "Permutation Entropy: A Natural Complexity Measure for Time Series," *Phys Rev Lett*, vol. 88, no. 17, Apr. 2002.

[5] F. Takens, *Detecting Strange Attractors in Turbulence*, vol. 898, Springer, Berlin, Germany, 1981.

[6] J.-P. Eckman, S. Oliffson Kamphorst, and D. Ruelle, "Recurrence Plots of Dynamical Systems," *Europhys Lett*, vol. 4, no. 91, pp. 973–977, 1987.

[7] N. Marwan, *Encounters with Neighbours: Current Developments of Concepts Based on Recurrence Plots and Their Applications*, Ph.D. thesis, University of Potsdam, 2003.

[8] J.P. Zbilut and C.L. Webber, "Embeddings and Delays as Derived from Quantification of Recurrence Plots," *Phys Lett A*, vol. 171, no. 3–4, pp. 199–203, Dec. 1992.

[9] C.L. Webber and J.P. Zbilut, "Dynamical Assessment of Physiological Systems and States Using Recurrence Plot Strategies," *J Appl Physiol*, vol. 76, no. 2, pp. 965–973, Feb. 1994.

[10] I. Voicu, J. M. Girault, C. Roussel, A. Decock, and D. Kouame, "Robust Estimation of Fetal Heart Rate from US Doppler Signals," *Physics Procedia*, vol. 3, no. 1, pp. 691–699, 2010.

[11] A. Groth, "Visualization of Coupling in Time Series by Order Recurrence Plots," *Phys Rev E*, vol. 4, 2005.