

Features Extraction and Classification of Raynaud's Phenomenon by IR Images

S. Di Zio - L. Ippoliti - A. Merla

Abstract This paper proposes a supervised classification approach for the differential diagnosis of Raynaud's Phenomenon on the basis of functional infrared imaging data. The classification of data from healthy subjects and from patients suffering for primary and secondary Raynaud's Phenomenon was first attempted by means of linear discriminant analysis (LDA). The direct application of LDA on raw data did not lead to satisfactory results. A substantial improvement of the classification was then achieved by addressing the problem of features extraction by means of functional variables and shape measures. The results of the proposed methodology are shown and discussed for a data set of 29 subjects.

Key words: Functional Infrared Imaging, Linear discriminant analysis, feature extraction, shape measures, Classification and pattern recognition, Functional data analysis.

1 Introduction

Raynaud's Phenomenon (RP) is a paroxysmal vasospastic disorder of small arteries, pre-capillary arteries and cutaneous arteriovenous shunts of extremities, typically induced by cold exposure and emotional stress [1]. RP usually involves the fingers of the upper and lower extremities, even though tongue, nose, ears, and nipples may result affected as well. RP can be classified as primary (PRP), with no identifiable underlying pathological disorder, and secondary (SSc), usually associated with a connective tissue disease, the use of certain drugs, or the exposition to toxic agents

S. Di Zio - L. Ippoliti
DMQTE, Viale Pindaro 42, Pescara, e-mail: s.dizio@unich.it, ippoliti@unich.it

A. Merla
Clinical Sciences and Bioimaging Department; Institute of Advanced Biomedical Technologies,
Foundation University G. d'Annunzio, e-mail: a.merla@itab.unich.it

[2]. Secondary RP is frequently associated with systemic sclerosis. In this case, RP typically may precede the onset of other symptoms and signs of disease by several years [1]. In this paper we are thus interested in classifying different forms of RP for an early and proper differential diagnosis.

Thermal infrared (IR) imaging is a non-invasive technique which has been widely used in medicine to evaluate cutaneous temperature. Considering a set of healthy controls (HCS), primary (PRP) and secondary (SSc) sclerosis RP patients, IR images will be used here for classifying the functional response of the fingers to a standard cold challenge [7]. To this end, several techniques for feature detection are discussed and the results of a supervised classification on a set of experimental data are shown.

The paper is organized as follows. In section 2 we present the data and how they have been obtained. In section 3 we address the problem of feature extraction proposing two different methods, based on functional data analysis and shape measures, respectively. A polar representation of the subjects is also proposed to visualize their differences. The results of the proposed methodology are shown in section 4.

2 The Data

Data for this study were provided by the Functional Infrared Imaging Lab - ITAB, Institute for Advanced Biomedical Technology, at the School of Medicine of the G. d'Annunzio University, Chieti, Italy. Data consist of finger temperature vectors documenting the thermal recovery from a standardised cold stress of both hands [7]. Specifically, the data consist of $n = 29$ subjects: 8 HCS, 6 PRP, and 15 SSc. The patients were classified according to the American College of Rheumatology criteria and standard exclusion criteria were observed [8]. Patients underwent to thermal IR imaging after having observed standard preparatory rules to the test [7]. Thermal high-resolution IR images were acquired every 30 seconds to monitor the response to a cold stress. Image acquisition started 2.5 minutes before (basal condition) and ended 20 minutes after the cold stress. The cold stress consisted in a two minute immersion of the hand in cold water (10°C), while wearing thin plastic gloves.

Thermal IR data were processed by means of in-home made Matlab software. For each subject, a set of 45 temperature images was collected. Each image series was corrected for motion artifacts by means of a contour alignment algorithm. Nail bed was considered as region of interest from which to extract the temperature curve vs. time (i.e. re-warming curves), and at each time point, data are obtained by averaging the temperature values of the pixels within the nail bed region. Figure 1 shows the registered re-warming curves for the subjects belonging to the three classes.

The application of a standardised cold stress determines an immediate temperature drop from the stationary steady state. The recovery is monitored starting at the sixth image of the series, that is right after the cessation of the cold stress. The dynamics of the re-warming curves appear quite different in the three study populations. PRP group exhibits an extremely low interindividual variability, whereas the SSc

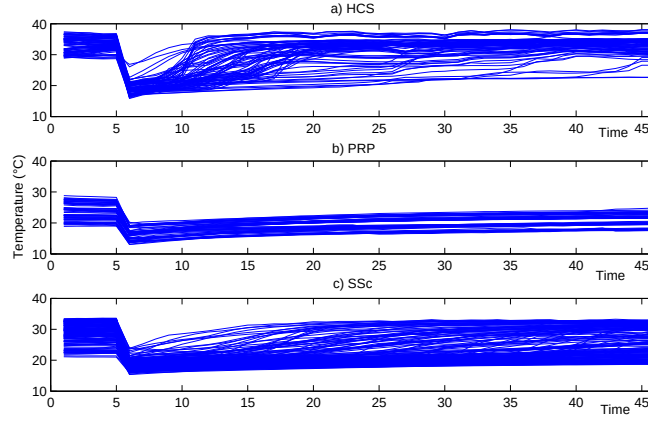


Fig. 1 Time series of re-warming curves for the three groups of HCS, PRP and SSc

group shows larger finger to finger and interindividual variability, according to the advancement and the presentation of the pathology ([3], [2], [7]). HCS group recovers to the baseline temperature after a few minutes, while PRP patients do not appear to activate any effective re-warming capability. SSc patients appear to show an intermediate recovery behaviour.

3 Discriminating Raynaud's Phenomenon

There exists a wealth of different classifiers and classification techniques and there is equally a large amount of literature on the subject [5]. In the following we consider the technique of linear discriminant analysis (LDA - [6]) which provides a way of classifying observations into groups based on specific predictors.

The LDA classifier is first implemented on the matrix of raw data. Each subject is represented by an observation vector, $\mathbf{x}_i$, $i = 1, 2, \dots, 29$, obtained by averaging the 10 re-warming curves corresponding to each finger. This procedure leads to the data matrix $\mathbf{X}$ of dimension (29×46) , where 46 are the time recording points. The performance of LDA can be assessed considering various approaches and one of the simplest is based on the calculation of the error rate of classification (i.e. the ratio of correctly classified subjects to the total number of subjects) within a leave-one-out cross-validation (LOCV) procedure. Applying the LOCV procedure on the raw data, the misclassification rate results equal to 20.70%, which is only partially satisfactory. Exactly the same results are achieved by first transforming the full feature space into a lower dimensional space by using the first three principal components representing the 96% of the variability. The estimated error rate corresponds to a misclassification of 6 subjects and specifically, 2 HCS and 2 PRP assigned to SSc

group, and 2 SSc assigned one to the group of HCS and one to the PRP group. One way of reducing the misclassification rate is to extract useful "latent" features from the re-warming curves and in the following we thus suggest two different approaches: the first one exploits the functional representation of the data while the second considers the time needed by the fingers to recover specific thresholds of temperature. While the former approach aims at extracting features from the pattern of the re-warming curves, the latter approach specifies both simple summary statistics (eg. mean, standard deviation and percentages) and shape measures.

3.1 Curve Interpolation

Although measured at discrete intervals, re-warming curves have an underlying continuous functional form and can thus be considered as samples of functional observations. The replication of curves, one for each finger, invites an exploration of the ways in which they vary and, for example, classical summary statistics would undoubtedly be helpful to explore their features.

In general, assuming that raw discrete data for the i -th statistical unit is $x_{ij}, i = 1, \dots, n; j = 1, \dots, p$, the first step is to convert these values into a function y_i . We assume that the values are errorless so that we are interested in the *interpolation* process. Notwithstanding the apparent "simplicity" of the pattern of the re-warming curves, polynomial functions do not provide satisfactory results in fitting the data thus, natural cubic splines [9] are used here to represent their temporal dynamic.

Once the data have been converted in functional form, then several classification variables are readily available. For example, *velocity* curves (i.e. first derivatives) can be easily obtained. Further useful variables are also the *integral* and the *arc-length* of the curves. For the latter, we recall that given a function, $y = f(t)$, with continuous derivatives, the distance $d(v, s)$ between two points, v and s , on the curve can be computed as [4]

$$d(v, s) = \int_v^s \sqrt{1 + \left(\frac{dy}{dt}\right)^2} dt. \quad (1)$$

3.2 Shape Measures

Figure 1 shows that several re-warming curves do not reach specific thresholds of temperature, such as the room temperature fixed around 23° . At the same time, due to the different thermal and blood flow properties, it can also be observed that there are several fingers that are able to reach the same threshold temperature after different time points. This suggests to create classification variables which highlight these dynamic aspects of the re-warming behaviour. Given the different pattern shown by healthy and unhealthy re-warming curves of the same subject, it appears extremely

important to preserve the information of each single finger. Here we propose to represent the 10 fingers of a subject on a polar diagram with each finger represented by a radius placed at a fixed distance of 36° from each other. The radius then represents the time needed by the finger to reach a fixed threshold temperature, τ ; for a given threshold, we thus have a polar representation of a subject. Then, by joining the extremities of the radii, we obtain irregular planar figures and, for example, for each of them we may compute simple shape measures such as the area and the perimeter. As an example, for $\tau = 21^\circ$, Figure 2 below shows a comparison of polar representations for two subjects belonging to HCS and SSc groups, respectively. As can be seen, while the HCS shows almost a regular pattern, the SSc one is much more irregular with some of the fingers not reaching at all the threshold temperature of 21 degrees. Notice that for illustrative purposes, the fingers that do not reach the threshold (within the end of the experiment) are conventionally represented as a point at the origin - Figure 2b. For data processing purposes, instead, the recovering time at that threshold is predicted through trend analysis.

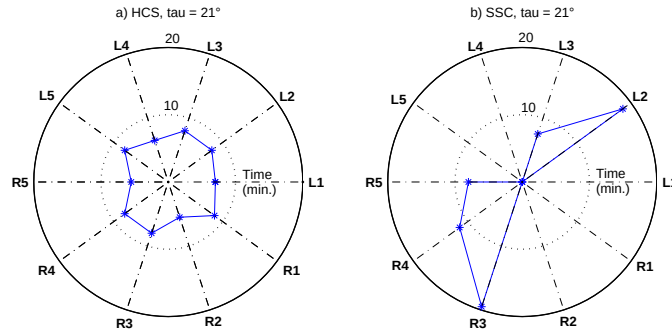


Fig. 2 Polar representation for two subjects belonging to HCS and SSc groups. Each radius represents the time needed by each finger to reach the fixed threshold temperature of $\tau = 21^\circ$.

Furthermore, in the representation we also impose a condition of circularity in which the radii represent the fingers according to their natural order distinguishing by "L" and "R" the left and right hands, respectively (see Figure 2). Of course, whereas we do recognize that a different order of the fingers can lead to different perimeters of the figure, we also highlight the following points: 1) keeping fixed the type of ordering, the procedure is consistent for all the subjects, 2) the area is invariant with respect to the ordering of the fingers and finally, 3) the polar representation is above all useful for illustrative and communicative purposes and without calculating shape measures, simple summary statistics can be used. Some examples are: a) the average time needed by each subject to reach the temperature τ , b) the standard deviation of

the time needed by each subject to reach the temperature τ and c) the percentage of the fingers which reach the threshold τ .

4 Discrimination Results

In this section we discuss discrimination results on the data described in section 2. The analysis is performed step by step considering different cases of interest.

Case 1: modelling the mean functions. In this section we first consider discrimination results using classification variables derived from the functional approach. Specifically, since for each subject 10 functions, y_k , are available (i.e. one for each finger), classical summary statistics such as, the *mean* and the *standard deviation* functions, can be obtained by considering the functions pointwise across replications.

Thus, for example, we have 29 mean functions, each representing a subject. For each of them we have calculated the following 15 variables: 1) the slope of the curve between the time points t_5 and t_6 (V1), 2) the first derivatives for the interval $[t_6, t_{15}]$ (V2-V11), 3) the arc-length (V12) and the integral (V13) in the interval $[t_6, t_{46}]$, 4) the integral ratio (IR) - (V14), i.e. the ratio between the integral of the curve within the interval $[t_6, t_{46}]$ and the area of the rectangle having the same interval as base and the maximum temperature as height, 5) the fraction of not recovering the basal temperature (V15), that is the relative quantity of temperature that is not recovered at the end of the experiment. Stepwise discriminant analysis with forward selection was used to choose the best subset of variables consisting in: V1, V11, V12, V13 and V14. The confusion matrix gives an estimate of the apparent error rate equal to 10.34%, but performing the LOCV procedure, the estimated error rate increases up to 13.8%. However, with respect to the results obtained using the raw data, we have an improvement of about 7%. All the HCS subjects are now correctly classified but still remain 2 PRP classified as SSc and 2 SSc classified as PRP.

Case 2: modelling the coefficient of variation functions. We know that healthy and unhealthy fingers (of the same subject) show a different pattern of the re-warming curves. Accordingly, for discriminant purposes, we thus expect to obtain more insights by studying the functions of the coefficient of variation - CV -, obtained as the (pointwise) ratio between the standard deviation and mean functions. The CV functions provide information on the inter-individual relative (to the mean) variability of the subjects and we therefore expect to extract different information by deriving on these functions the same functional features described in Case 1.

Specifically, we observe a completely homogeneous pattern for the PRP group which also demonstrates an intensity of the variability largely different from the other two groups. On the other hand, the CV functions for HCS and SSc subjects show much more variable patterns. Furthermore, HCS functions are more variables within the first 20 time points and this suggests that we may improve the discrimination results by including also classification variables identified into two different

intervals of the series trying to capture the different variability of the HCS and SSc groups. The truncation point, t_k , used to split the two intervals is chosen by minimizing the LOCV error rate with k varying in the interval $[t_{10}, t_{35}]$. Following this procedure we have found that the minimum LOCV error rate is achieved for $k = 21$ (i.e. 8 mins after the cold stress). Notice that this result is physiologically relevant since HCS subjects normally recover the room temperature after 7-8 minutes. This procedure has led to the identification of 41 functional variables reproducing the features described in Case 1, repeated for the whole length period of the series and for the series split in the two intervals $[t_1, t_{21}]$ and $[t_{22}, t_{46}]$.

Stepwise discriminant analysis with forward selection was again used to choose the best subset of variables consisting in the following variables: 1) arc-length for the interval $[t_6, t_{46}]$, 2) the integrals for the intervals $[t_1, t_{46}]$ and $[t_1, t_{21}]$, 3) the integral ratio - IR -, and 4) the first derivative at time point t_{17} , for a total of 5 classification variables. The confusion matrix resulting from LDA gives an estimate of the apparent error rate equal to 3.4%, but performing the LOCV procedure, the estimated error rate significantly increases up to 20.7%. In particular 2 HCS, 1 PRP and 3 SSc are wrongly assigned. This may be explained by the great inter-individual variability existing among the subjects.

Case 3: modelling shape measures. We discuss here discrimination results from LDA applied on classification variables created following the approach described in section 3.2. Considering the shape measures of area (A), perimeter (P), and the shape factor $SF = P/\sqrt{A}$, three classification variables are computed for each value of τ . Specifically, for $\tau \in [20, 28]$, 27 classification variables are included in the model. The stepwise procedure has selected only two features corresponding to the perimeter at $\tau = 25$ and SF at $\tau = 28$. The LOCV error rate and the apparent error rate coincide, and are equal to 13.8%, the same achieved with functional approach using the mean functions. However, differently from the mean function approach, we observe here that the estimated model misclassifies 1 HCS and 3 SSc.

Alternatively, by using the three summary statistics described in a), b) and c) at the end of section 3.2, the rate of misclassification slightly improves decreasing to 10.3%. Differently from the shape measure approach, we have that the estimated model here misclassifies 2 HCS (attributed to SSc) and 1 SSc (attributed to PRP).

Case 4: the integrated approach. In this section we discuss discrimination results from LDA performed by joining together the best-set of classification variables presented in Cases 1-3. Specifically, through the stepwise discriminant analysis performed on the resulting 18 variables, the following 5 classification variables have been selected: 1) arc-length of the mean and CV functions for time points t_6 - t_{46} (V1-V2), 2) the first derivative of the CV functions at time points t_{17} (V3), 3) the integral ratio, IR, for the CV functions (V4) and finally, 5) the average time needed by the subject to reach the temperature $\tau = 25^\circ$ (V5). With 3 groups, two discriminant functions can be estimated; the first function accounts for the 88.8% of the common variance while the second one accounts for the additional 11.2%. The confusion matrix resulting from LDA gives an estimate of the apparent error rate equal to 0%; apart from one unit for which we observe a probability of 0.478 and 0.522 to

belong to the PRP and SSc classes, respectively, the probabilities of the individuals to belong to their own class are very large (very close to 1). This justifies the results from the LOCV procedure, for which the estimated error rate is 3.4%, which corresponds exactly to the misclassification of 1 SSc wrongly assigned to the PRP class with a probability of 0.88.

5 Conclusions

As seen in Case 4, very good discrimination results can be achieved by taking into account all the relevant features of the re-warming process. The three different approaches (cases 1-3), considered separately, provided larger misclassification errors since each of them is only capable to capture specific aspects of the phenomenon. Finally, we highlight that to overcome the problem of the small sample size we considered different approaches for estimating the classification error rate. Specifically, both bootstrap and H-fold corss-validation techniques agree with the results obtained by LOCV. Unfortunately, the small sample size can hardly be changed in practical problems like this but, however, we aim at extending the analysis to a wider data set in further works. A comparison of different classifiers, such as PLS-DA and Neural Network, will also be considered in the next future.

References

1. Belch, J.: Raynaud's phenomenon. Its relevance to scleroderma. *Ann. Rheum. Dis.*, **50**, 839–845 (2005)
2. Block, J.A., W. Sequeira: Raynaud's phenomenon. *Lancet*, **357**, 2042–2048 (2001)
3. Di Carlo, A.: Thermography and the possibilities for its applications in clinical and experimental dermatology. *Clin. Dermatol.*, **13**, 329–336 (1995)
4. Gullberg, J.: *Mathematics from the Birth of Numbers*. W.W. Norton and Company, New York (1997)
5. Krzanowsky, W.J., Marriot, F. H. C.: *Multivariate Analysis: Classification, Covariance Structures and Repeated Measurements*, Vol 2. London: Arnold (Holder Headline Group) (1995)
6. Mardia, K. V., Kent, J. T., Bibby, J. M.: *Multivariate Analysis*, Academic Press, Duluth, London (1979)
7. Merla, A., Romani, G.L., Di Luzio, S., Di Donato L., Farina G., Proietti M., Pisarri S., Salsano, S. Raynaud's phenomenon: infrared functional imaging applied to diagnosis and drug effect. *Int. J. Immunopathol. Pharmacol.*, **15**(1), 41–52 (2002a)
8. Merla, A., Di Donato, L., Pisarri, S., Proietti, M., Salsano, F., Romani G.L.: Infrared Functional Imaging Applied to Raynaud's Phenomenon. *IEEE Eng. Med. Biol. Mag.*, **6**:73 (2002b)
9. Wahba, G.: *Spline Models for Observational Data*. Philadelphia: Society for Industrial and Applied Mathematics (1990)