Automatic Drawing Analysis of Figures Included in Neuropsychological Tests for the Assessment and Diagnosis of Mild Cognitive Impairment

M. Rincón[1(✉)], S. García-Herranz[2], M.C. Díaz-Mardomingo[2], R. Martínez-Tomás[1], and H. Peraita[2]

[1] Department of Inteligencia Artificial. E.T.S.I. Informática, UNED, Madrid, Spain
[2] Department Psicología Básica I. Facultad de Psicología, UNED, Madrid, Spain
mrincon@dia.uned.es

Abstract This proposal is framed within the group's general working line of applying artificial intelligence techniques to advance in early mild cognitive impairment diagnosis. If impairment in semantic production was studied in previous works, now we rely on the reduced ability to reproduce or copy simple figures, part of standardized neuropsychological tests designed to assess mild cognitive impairment. Although the long-term goal of this project is to work with all figures from these tests, in this paper we will focus on the automatic analysis of the alternating graphs figure. We develop a quantitative descrition of different features that appear to be very abstract in the test norms and define new features that are not considered so far. Results with just one figure are quite promising (77.7% precision and 77.1 recall).

Keywords: Drawing analysis · Mild Cognitive Impairment · Alzheimer disease · Drawings in neuropsychological tests · Test Barcelona

1 Introduction

It is known that mild cognitive impairment (MCI) is detected in the very early stages of Alzheimer disease (AD) and other neurodegenerative dementias, which is essential for achieving maximum effectiveness in pharmacological treatments and cognitive therapies. One part of standardized neuropsychological tests designed to assess MCI include reproducing or copying figures. The main idea of this work is to analyse, using artificial intelligence techniques, from a typical pattern or standard of each one of the figures with which we have worked in an investigation on early detection of mild cognitive impairment [1,2,3], the extent to which certain distortions from the standard may indicate of different profiles and degrees of MCI. These figures come from the Mini Examen Cognoscitivo (MEC) [4], the Test Barcelona [5], and the Rey complex-figure Test [6].

The processes and cognitive functions —cognitive domains— that are supposedly assessed by the execution of these figures are: the executive function, visual and/or visuospatial perception, motor skills, and spatial memory. In each

J.M. Ferrández Vicente et al. (Eds.): IWINAC 2015, Part I, LNCS 9107, pp. 508–515, 2015.
DOI: 10.1007/978-3-319-18914-7_53

of these functions or processes, the traits or components (control, inhibition, planning, etc.) involved in the patterns to be analyzed can be defined, for example, following the indications of Dr. Peña Casanova with regard to the drawings of alternating graphs and loops [5].

An important complementary aspect is to be able to obtain not only the distortion with regard to the standard figure at any given time, but also to be able to determine how this evolves over the years, an aspect that can be analyzed in the future, as it constitutes the data of our longitudinal research.

Although the long-term goal of this project is to work with all figures from the above-mentioned tests, we present herein only the results of the figure alternating graphs, which is one of the subtests of the Barcelona Test. The diagnostic value of this test, selected for this automatic exploratory analysis, lies in the fact that in its execution are involved some executive function components, such as seriation, planning, flexibility, inhibition, as well as praxic capacity.

The task consisted of copying one figure, one in which peaks and plateaus should alternate. Following the test norms, scoring is done according to the quality of the copy, assigning 0 to 2 points for each one of them. The distortions that can emerge in the execution of these figures can be of different types: variations in the size of the execution of the figure, alteration of features by addition or deletion, scribbling, perseverations, rotations, etc. These errors or alterations in the reproduction of the drawing may be markers of more severe dysfunctions, which can be of an apraxic type, in which the executive functions are also involved, as in Alzheimer's type dementia [7].

Whereas in other types of standardized tests that assess episodic memory, verbal fluency, etc. it is much easier to obtain normative data that allow scoring free from subjectivity, in these types that involve reproducing and copying figures, it is much more difficult, imposing the subjectivity and some discretionality by the evaluator. Although scoring criteria exist, there is a large component of subjectivity, that can undermine the reliability if discrepancies among the evaluators are not corrected [8]. As a result of this lack of agreement in the correction, there may be an important margin of error in the detection of certain problems of motor skills, visual and visuospatial perception, etc. within the framework of a general plan for early detection of MCI, either prior or not, to AD. The method proposed herein is a method of Artificial Intelligence (A.I), inexpensive, easy to apply and implement and that provides a convergence of criteria of both methods: the manual one and the automatic one. This could lead to discovering certain aspects of figure execution that may be significant for the early detection of alterations within the framework of specific cognitive domains.

In previous publications, our group has worked to define an economic procedure for MCI diagnosis by analysing cognitive alterations affecting declarative semantic memory [9,10]. As in this work, our goal there was to objectify and automate the analysis of a test that, because of its low cost, it could be used for routine clinical evaluations or screenings that could lead to more expensive and selective tests that confirm or rule out the disease accurately. We confirmed that, in this context, Bayesian networks are the most appropriate tool for this purpose because they allow us to combine previous knowledge with case data (the network structure, the qualitative part of the model, is obtained from psychology experts and epidemiological studies, and the network parameters, the quantitative part of the model, are learnt automatically from epidemiological studies and a linguistic corpus of oral definitions [11].

Other non-conventional diagnostic methods proposed to evaluate the cognitive state of patients or even to detect motor deficiencies caused by a brain haemorrhage from monitoring daily activity. In particular, Matic et al. analyse the act of getting dressed [12] and Kearns et al. analyse the tortuosity in movement paths—irregular movements—of elderly people with cognitive impairment [13].

The remainder of the paper is organised as follows: Section 2 describes the materials and methods, i.e., the sociodemographic and clinical data of the participants in the study and purpose specific automatic method for alternating graphs drawing analysis. In section 3 we present the experimental results. Finally, section 4 presents our concluding remarks as well as the future lines of research.

2 Materials and Methods

2.1 Participants

The sample (N = 40 participants) was recruited from a larger sample of participants in an ongoing longitudinal study (ref. SEJ 2004-04233 and SEJ 2007-63325) focused on determining the prevalence the different MCI subtypes [1,3]. The participants were recruited in the Autonomous Community of Madrid, Spain. They were assessed longitudinally with a neuropsychological battery during an average period of 3 years. MCI was defined as having a score of 1.5 SD below the mean in at least 2 of the tests applied. Depending on the data obtained through the different neuropsychological assessments, the participants were classified in one of the following cognitive profiles: healthy individuals (n=16)—expected performance according to references scales—or MCIs (n=24). Of the MCI group, after a 3-year follow-up, 10 of them had a diagnosis compatible with an initial phase of probable AD (see Table 1).

2.2 Methods

We have defined a purpose specific automatic method for alternating graphs drawing analysis. The general approach considered in this paper consists of the following steps: 1) drawing digitalisation, 2) drawing segmentation, 3) line extraction, 4) pattern matching and characterization and 5) MCI diagnosis. Figure 1 shows an example with the intermediate results of each one of these steps.

Drawing Segmentation. After digitalisation with a standard scanner, a gray-scale image is obtained (Fig. 1.a). The image contains two figures, the pattern

Table 1. Baseline sociodemographic and clinical descriptive data of healthy, MCI and MCI-converters

	Healthy n = 16	**MCI n = 14**	**MCI-Converters n = 10**
	Mean (SD)	Mean (SD)	Mean (SD)
Gender (female)	13 (81.25%)	11.00 (4.24)	4 (40%)
Age (years)	69.75 (4.85)	72.57 (5.31)	71.60 (5.08)
Formal education (years)	11.00 (4.24)	5.50 (6.51)	8.30 (5.53)
Geriatric Depression Scale (GDS) (Yesavage scale)[14]	3.44 (2.78)	3.93 (2.99)	4.20 (3.39)
Functional level (Blessed scale) [15]	0.59 (0.45)	0.67 (0.57)	1.00 (1.00)
Cognitive status (MEC 0-35)[4]	33.50 (2.47)	29.14 (3.78)	30.30 (2.00)

and the manual drawing. Histogram analysis is performed in order to segment the objects from the background. It is assumed that background is white and that the drawing is a continuous black line that comprises a small amount of the image pixels. Due to the fact that posterior pattern analysis is simplified if segmentation obtains a thin object, an iterative threshold selection is used that evaluates the results and stops when more than two big objects are obtained (Fig. 1.b). After that, the region of interest is rescaled in x and y to a standard size. The segmented line quality depends tremendously on the conditions under which the drawing is performed, the pen type, and the scanning process.

Line Extraction. Line thickness is various pixels wide and we therefore need to thin it down. We have used mathematical morphology for extracting the skeleton, which is very noisy and contains many small branches that must be eliminated (Fig. 1.c) Assuming that both line ends are in the x-coordinate extremes, the rest of braches are eliminated by deleting iteratively the endpoints. If small loops are detected, they are broken and the intermediate branches are eliminated again. The final result is a continuous line with no loops. Then, we use the recursive Douglas-Peucker line simplification algorithm to approximate the curve with line segments to a specified tolerance (Fig. 1.d).

Pattern Matching and Characterization. In this point, we have to make clear that our interest is not limited to recognise the alternating graph pattern or to assess if the pattern is copied correctly or not, instead, we are interested in wider assessment metrics that serve us to discriminate between groups of people. Because the alternating graph consists of repeating five times the peak-plateau

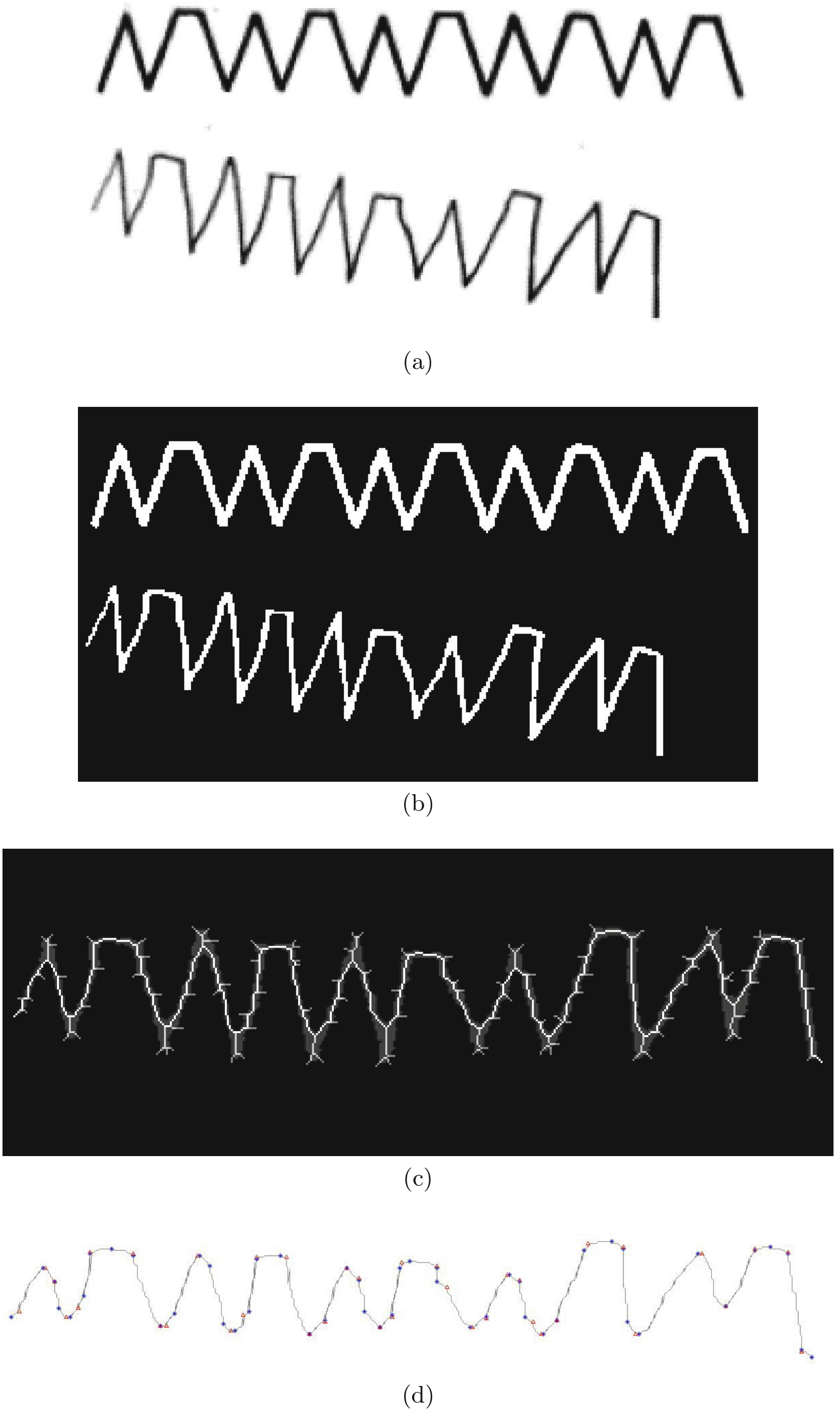

Figure 1. Peak-Plateau drawing segmentation and segment decomposition. a) scanned image; b) segmented image; c) skeleton extraction and d) segment approximation

pattern, which in turn consists of a sequence of various segments with different orientations (ascending, descending and horizontal segments), we analyse whether the segments found in the drawing match the pattern or not. The result is a measure of the number of valid and invalid segments found in the drawing. Table 2 summarises the list of features used for characterising the alternating graph that allows us to compare it with the model.

Table 2. List of features used for characterising the peak-plateau drawing

Feature	Description
height_pattern_diff	height difference between pattern and drawing
width_pattern_diff	width difference between pattern and drawing
drawing_tilt	tilt of the drawing with respect to the sample.
vertical_dist_to_model	vertical distance between drawing and model
X_scale_factor	scale factor in X coordinate used for drawing normalization
Y_scale_factor	scale factor in Y coordinate used for drawing normalization
alternance	variable associated with a perfect match between drawing and model in the five peak-plateau pattern
#valid_patterns	number of valid peak-plateau patterns
phase_diff	average x coordinate difference between pattern and drawing
plateau_width	average width of the horizontal segment in plateau subpatterns
#segments	total number of detected segments
#valid_segments	number of drawing segments that match the pattern segments
#invalid_segments	number of drawing segments that does not match the pattern
average_segment_tilt	average tilt of the drawing segments
last_segm_match	last consecutive valid segment

DCL Diagnosis. We have built a supervised machine learning classifier using the dataset described in subsection 2.1. In this stage of the study, we are interested in the analysis of the discrimination power of the different features. Therefore, we have used a J48 decision tree to implement the classifier because of ease of understanding. Due to the small sample size, we have used leave-one-out cross-validation for evaluating the classifier, which allows to use the largest possible training sample while keeping a reliable performance estimation.

3 Experimental Results

The final features included in the decision tree are `#invalid _segments`, `X_scale_factor` and `vertical_dist_to_model`. It is worth noting that `#valid _segments` is not very discriminative and it is highly correlated with `#invalid_segments`. Besides, the most discriminative measures, `vertical_dist_to_model` and `X_scale_factor`, are not considered in the neuropsicological tests. The results are quite promising (0.777 precision / 0.771 recall) having in mind the limitations imposed by the small sample size and the fact that the analysis is based on just one item of one test. As mentioned in the introduction, the expert's score in this test is quite generic and only distinguishes three possible values: good performance (2), good performance with some defects (1) and mistake (0).

4 Conclusions y Further Research

In this paper we have demonstrated that AI techniques can offer solutions to support the automatic analysis of drawings included in neuropsychological tests for the assessment and diagnosis of MCI. The main advantage of the automatic analysis is that it includes a larger amount of metrics for characterising the drawings, which makes it more quantitative, robust and user independent.

We have implemented an automatic system alternating graphs analysis and we have found discriminative metrics for diagnosing MCI. The fundamental problem in machine learning AD diagnosis, as in most neurological studies, is the absence of a training dataset large enough to build a reliable AD diagnosis system using supervised learning. We have to recognise that with this small sample size, we can only conclude that these proposed features are good candidates for MCI diagnosis, but they alone can not distinguish between different MCI types. In future works we will broaden the sample and will combine features from different figures to improve the classification performance. Therefore, this work intends to be a pilot study of a much broader work in which, from the longitudinal data already available, we can make different types of analysis:

- Longitudinal analysis of healthy controls: comparing, over a series of years, the execution patterns of healthy control subjects, to verify their stability in the execution of the figures.
- Longitudinal analysis of the stable MCIs: the same type of comparative approach but in subjects with stable MCI, to attempt to answer the following questions: Is the execution of the figures stable? Of all of them? For how long?
- Longitudinal analysis of the MCIs that evolve to AD: the same in subjects with an evolutionary MCI to Alzheimer's disease or another dementia.
- Transversal analysis of the different types of MCI: the same in subjects with various types of MCI: amnestic, multidomain and nonamnestic.
- Relationship between the figure drawing and ideomotor praxia: to study the relationship between the execution of the figures and the ideomotor praxia obtained by the same subjects in other tests.
- Socio-demographic analysis: to analyze the relationships between socio-demographic variables, such as age, gender and level of education, and the execution patterns of the above-mentioned figures.

Acknowledgments. This research has been supported by project No. 018-ABEL-CM-2013 of NILS Science and Sustainability program coordinated by Universidad Complutense (Spain), and UNED projects SEJ-2004-04233 and SEJ-2007-63325.
We thank all voluntary participants of las Rozas, Pozuelo and Madrid for taking part in the longitudinal study whose data have led to this work for their selfless collaboration, as well as the municipalities and institutions that facilitated contact details and provided an adequate logistics to carry out evaluations.

References

1. Díaz, C., Peraita, H.: Detección precoz del deterioro cognitivo ligero de la tercera edad. Psicothema 20(3), 438–444 (2008)
2. Díaz-Mardomingo, M.C., García-Herranz, S., Peraita, H.: Detección del DCL y conversión a la EA: un estudio longitudinal de casos. Psicogeriatría 2(2), 105–111 (2010)
3. Peraita, H., García-Herranz, S., Díaz-Mardomingo, M.C.: Evolution of Specific Cognitive Subprofiles of Mild Cognitive Impairment in a Three-Year Longitudinal Study. Current Aging Science 4, 171–182 (2011)
4. Lobo, A., Ezquerra, J., Gómez, F., Sala, J.M., Seva, A.: El mini-examen cognoscitivo. Un test sencillo, práctico, para detectar alteraciones intelectivas en pacientes médicos. Actas Luso Españolas de Neurolología, Psiquiatría y Ciencias Afines 3, 189–202 (1979)
5. Peña-Casanova, J.: Programa integrado de exploración neuropsicológica "test Barcelona". Normalidad, semiología y patología neuropsicológica. Masso, Barcelona (1991)
6. Rey, A.: Rey. Test de copia y de reproducción de memoria de figuras geométricas complejas. TEA, Madrid (2003)
7. Freeman, R.Q., Giovannetti, T., Lamar, M., Cloud, B.S., Stern, R.A., Kaplan, E., Libon, D.J.: Visuoconstructional problems in dementia: contribution of executive systems functions. Neuropsychology 14(3), 414–426 (2000)
8. Urbina, S.: Claves para la evaluación con tests psicológicos (Kaufman, A.S., Kaufman, N.L (trad.)). TEA Ediciones (publicado originalmente en 2004), Madrid (2007)
9. Guerrero Triviño, J.M., Martínez-Tomás, R., Peraita Adrados, H.: Bayesian Network-Based Model for the Diagnosis of Deterioration of Semantic Content Compatible with Alzheimer's Disease. In: Ferrández, J.M., Álvarez Sánchez, J.R., de la Paz, F., Toledo, F.J. (eds.) IWINAC 2011, Part I. LNCS, vol. 6686, pp. 419–430. Springer, Heidelberg (2011)
10. Guerrero, J.M., Martínez-Tomás, R., Rincón, M., Peraita, H.: Bayesian network model to support diagnosis of cognitive impairment compatible with an early diagnosis of Alzheimer's disease. Methods of Information in Medicine (in revision, 2015)
11. Peraita, H., Grasso, L.: Corpus lingüístico de definiciones de categorías semánticas de personas mayores sanas y con la enfermedad de Alzheimer. Una investigación transcultural hispano-argentina. Fundación BBVA (2010), http://www.fbbva.es/TLFU/dat/DT%203_2010_corus%20linguistico_peraita_web.pdf
12. Matic, A., Mehta, P., Rehg, J.M., Osmani, V., Mayora, O.: Monitoring Dressing Activity Failures through RFID and Video. Methods Inf. Med. 51(1), 45–54 (2012)
13. Kearns, W.D., Nams, V.O., Fozard, J.L.: Tortuosity in Movement Paths Is Related to Cognitive Impairment. Methods Inf. Med. 6(49), 592–598 (2010)
14. Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O.: Development and validation of a geriatric depression scale: A preliminary report. Journal of Psychiatry Research 17(1), 37–49 (1983)
15. Blessed, G., Tomlinson, B.E., Roth, M.: The association between quantitative measures of dementia and of senile changes in the cerebral grey matter of elderly subjects. British Journal of Psychiatry 114, 797–811 (1968)