Visual analysis and quantitative assessment of human movement
Soancatl Aguilar, Venustiano

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2018

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Distinguishing Patients with a Coordination Disorder from Healthy Controls Using Local Features of Movement Trajectories during the Finger-to-Nose Test

Abstract

Quantitative assessment of movement disorders is valuable for monitoring progression of patients, distinguishing healthy and pathological conditions, and ultimately aiding in clinical decision making, thereby offering the possibility to improve medical care or rehabilitation. A common method to assess movement disorders is by using clinical rating scales. However, such scales depend on the evaluation and interpretation of an observer and thus contain a subjective component. Objective and more accurate methods are under continuous development but gold standards are still scarce. Here, we show how a method we previously developed, originally aimed at assessing dynamic balance by a probabilistic generalized linear model, can be used to assess a broader range of functional movements. We here apply this method to distinguish patients with coordination disorders from healthy controls. We focused on movements recorded during the finger-to-nose task (FNT), which is commonly used to assess coordination disorders. We also compared clinical FNT scores and model scores. Our method achieved 84% classification accuracy in distinguishing patients and healthy participants, using only two features. Future work could entail testing the reliability of the method for distinguishing groups of patients and/or controls using other clinical tests such as finger chasing or quiet standing and using other tracking devices such as depth cameras or force plates.

5.1 Introduction

Quantitative analysis of human movement can be valuable for diagnosis and monitoring of motor disorders; it can aid in distinguishing healthy and pathological conditions and in following the progression of patients over time as well as the efficacy of interventions [156]. A common method to assess human movement in a clinical setting is provided by clinical rating scales which are easy to administer and have been validated and standardized [108]. However, one of the main drawbacks of rating scales is that they depend on the evaluation and interpretation of an observer and thus contain a subjective component. Moreover, clinical scales are not enough to assess different motor control strategies during
the execution of movements [108, 114, 142]. Thus, objective measures and methods that quantify aspects of movement could be valuable in neurology, rehabilitation and other fields of medicine. As added benefit, they could be used in combination with clinical rating scales [25].

One of the challenges when developing quantitative and objective methods to assess human movements is the lack of methods to establish the validity of the measurements, as the more subjective rating scales usually provide the only reference. In a previous study (see [129, 130]) we used the movement performance of younger and older participants as a proxy of better and worse movement, knowing that movement in older participants is generally worse than in younger participants. We then successfully used generalized linear models (GLMs) [159] to predict movement category (young or old) based on features derived from the movement trajectory. One characteristic that makes GLMs appropriate to classify human movement as better or worse is that their outcomes can be probability values that reflect movement performance. In case of diagnostic applications, if we assume that probability 0 represents “healthy” performance and probability 1 represents “pathological” performance, we propose that a probabilistic GLM could be used in a similar way. Intermediate probability values as estimated by the GLM would then indicate how similar the movements are to those movements that reflect pathological performance.

Here, we evaluate how the method that we used in our previous study to distinguish young and old participants [129] performs when applied to the problem of distinguishing patients with a coordination disorder from healthy controls during a movement task that is recorded using inertial measurement units (IMUs). We focused on movements recorded from the finger to nose task (FNT). The FNT is a neurological examination that measures smooth and coordinated upper-limb movements between the tip of the nose of the participant and the tip of an examiner’s index finger[70]. This test can be used to assess coordination in diseases such as Early-Onset-Ataxia (EOA) or Developmental Coordination Disorder (DCD) [81]. Although there are valid clinical scales measuring the severity of ataxia [155] such as the International Cooperative Ataxia Rating Scale (ICARS) and the Scale for the Assessment and Rating of Ataxia (SARA), assessing the presence of dysmetria (difficulty in controlling the range of movement resulting in undershooting or overshooting a target [70]) and tremor in FNT trials is challenging. Even experienced clinicians have exhibited low reliability in the assessment [135]. Our aim is to contribute to the development of objective measures and methods to quantify movement in patients with a coordination problem thereby eventually aiding in diagnosis and monitoring of such patients. A very first step in achieving this is to be able to distinguish between patients and controls [81].

Figure 5.1 illustrates the general steps of our method. In step 1, the three-dimensional (3D) trajectories are collected (during a functional
5.2 METHODS

This study was performed using data acquired in the context of the project Quantification of symptoms of movement disorders employing motion sensors [83]. Part of the data were previously used to investigate whether a random forest classifier employing 14 features derived from 3D movement trajectories during the FNT could classify children with coordination problems and age-matched healthy control children. In the present study we only use two features, which are different from those used in [83] and include additional participants.
5.2.1 Participants

In the present study we included two sets of participants. The first set concerns the data used in the study mentioned above and consisted of 34 children: nine children with EOA (mean age 13.3 years, SD 4.0 years), seven children with DCD (mean age 9.4 years, SD 2.2 years), and 18 healthy age-matched control children (mean age 11.8 years, SD 3.4 years). For the second set, the data were collected at a different time and involved 36 participants: 12 children with EOA (mean age 13.5 years, SD 2.8 years), 22 adults with ataxia (mean age 54.9 years, SD 14.7 years), and two healthy participants (age unknown). By including patients and controls over a wider age range, we increased the complexity of the data compared to the previous study (see [83]). All parents of the children and all adult participants provided written informed consent. Children who were 12 years or older provided informed assent. Inclusion criteria for ataxia patients were a clinical diagnosis of pediatric ataxia or recognition of ataxia as a primary movement disorder as assessed by three experts in movement disorders. The DCD inclusion criterion was an official diagnosis as determined by a rehabilitation center. Exclusion criteria for healthy participants were a neurological and/or orthopedic disorder and/or any medication with a negative effect on coordination. Furthermore, healthy children were declared to be healthy by their parents.

5.2.2 Data collection and preprocessing (steps 1 and 2)

Each participant performed the FNT during 21.8 sec on average (SD 7.8 sec) with both hands, left and right. The trials were video recorded. Three pediatric neurologists additionally assessed the FNT executed by EOA and DCD participants, according to the official SARA guidelines [119], for the first set of participants only. SARA assessment was not performed for the second set of participants. During task execution, participants wore three inertial measurement units (IMUs Shimmer3, Shimmer, Dublin, Ireland-based Realtime Technologies) on the upper arm, fore arm, and index finger. The data collected by the IMUs at 51.2 Hz were used to estimate 3D trajectories of the participants’ index finger using an upper limb model [83] implemented in Labview (Austin, Texas, United States of America). We subsequently applied a moving average filter of the 3D trajectory data using a window of 15 samples to smooth the signals.

5.2.3 Estimating local features (step 3)

Local features are those that can be estimated for short segments taken from the 3D trajectories such as curvature, torsion, instantaneous speed.
5.2 Methods

and their time-derivatives [128, 153]. Local features have the added value compared to global features that they offer the possibility to assess performance in "real-time" and provide immediate feedback. As local features we selected local curvature and instantaneous speed because they allowed high classification accuracy in our previous study involving movement of younger and older participants [129] and because they are expected to provide relevant information about the ability of the current participants to perform the FNTs. Local curvature measures how smooth a 3D trajectory is for each three consecutive points, while instantaneous speed is determined between each two consecutive points. By visualizing curvature and speed signals and identifying the repetitive FNT movement, samples that preceded or followed FNT execution were excluded from further analysis. After sample exclusion, trials lasted 17.5 sec on average (SD 6.4 sec). Then, local curvature ($\kappa$) and instantaneous speed ($s$) were estimated from the FNT trajectories according to the method of Soancat-Aguilar et al. [128] and subsequently log-transformed. Finally, mean speed ($\bar{s}$) and mean curvature ($\bar{\kappa}$) were estimated for each participant $k$ and used as predictors in the GLM.

5.2.4 GLM definition and GLM fitting (steps 4 and 5)

Following the steps described in [159] we specified a GLM as follows. First, we defined an outcome variable ($d$) as binary (0 - healthy class, 1 - coordination disorder class) and assumed that it follows a Bernoulli distribution. Second, a linear model was specified as a function of mean speed $\bar{s}$ and mean curvature $\bar{\kappa}$. Third, the logit function [24] was used to transform the probability distribution constrained between 0 and 1 into a function that can take any real value. Mathematically:

$$d_k \sim Bernoulli(\mathbb{P}_k), \quad k = 1 \ldots n$$

$$\logit(\mathbb{P}_k) = \alpha + \beta_1 \cdot \bar{\kappa}_k + \beta_2 \cdot \bar{s}_k,$$

(5.1)

$$\alpha \sim \mathcal{N}(0, 10), \quad \beta_1 \sim \mathcal{N}(0, 50), \quad \beta_2 \sim \mathcal{N}(0, 50)$$

where $n$ is the number of participants, $\alpha$ is the intercept, $\beta_1$ and $\beta_2$ are the slopes, and $k$ is a participant index. The logit function is defined as the logarithm of the odds (log-odds) [24], where the odds of $\mathbb{P}_k$ is $\mathbb{P}_k/(1 - \mathbb{P}_k)$. Thus,

$$\logit(\mathbb{P}_k) = \log\left(\frac{\mathbb{P}_k}{1 - \mathbb{P}_k}\right) = \alpha + \beta_1 \cdot \bar{\kappa}_k + \beta_2 \cdot \bar{s}_k,$$

(5.2)

and solving for $\mathbb{P}_k$

$$\mathbb{P}_k = \frac{e^{\alpha + \beta_1 \cdot \bar{\kappa}_k + \beta_2 \cdot \bar{s}_k}}{1 + e^{\alpha + \beta_1 \cdot \bar{\kappa}_k + \beta_2 \cdot \bar{s}_k}}$$

(5.3)

$\mathcal{N}$ represent a normal distribution with 0 mean and standard deviation 10 for the intercept and standard deviation 50 for the slopes ($\beta_1$ and $\beta_2$).
To fit the GLM (Eq. 5.1) we built a model in Stan, which is a probabilistic programming language [43], using the rethinking R package [87].

5.2.5 **GLM performance (step 6)**

We performed leave-one-out cross validation (LOOCV) [64] to test the performance of the model on new data. Suppose that the set $U$ contains the pairs $(\bar{\kappa}_k, \bar{s}_k)$ collected from the two sets of participants ($k = 1 \ldots 70$). Then, for each participant $k$ in $U$ we fitted a model on the set $\{U - (\bar{\kappa}_k, \bar{s}_k)\}$ and used the fitted model to predict the probability that participant $k$ belongs to the coordination disorder class. The predicted probabilities were used to estimate an optimal threshold to classify FNT trials as belonging to a healthy or coordination disorder participant. This threshold was estimated as the point with the best sum of sensitivity and specificity known as the Youden index [157], closest to the point $(0,1)$ of the receiver operating characteristics (ROC) curve [38]. Sensitivity is the proportion of correctly classified patients. Specificity is the proportion of correctly classified healthy participants. The threshold was estimated using the pROC R-package [111].

5.2.5.1 **SARA scores compared to GLM scores**

To gain further understanding of any misclassifications, SARA scores and model scores were compared for the first group of patients only. The mean SARA score across observers for each patient in the first data set was determined. Then, to investigate to what extent SARA scores coincide with model scores a scatter plot was used. For specific cases, we visualized 3D trajectories and the distribution of curvature and speed values as violin plots to gain further understanding.

5.3 **Results**

Figure 5.2 provides an example of local curvature and instantaneous speed (in log scale) as a function of time for a healthy participant and a participant with EOA. Both measures, speed and curvature, are clearly regular and repetitive for the healthy participant. For the patient, however, both measures behave more irregularly. Taking into account the range of the measures, the healthy participant displayed faster and smoother movements than the EOA participant, as indicated by higher speed values and lower curvature values. It can also be observed that high speed values coincide with low curvature values and, vice versa, low speed values coincide with high curvature values. This is known as the power law relation between curvature and speed in log scale [51].
5.3.1 GLM classification

After performing LOOCV and using the ROC curve, the probability threshold that best separates healthy participants from patients was found to be 0.587. Using this threshold, 84% of the healthy participants were correctly classified and 84% of the patients were correctly classified. Figure 5.3 shows the LOOCV predictions of model (5.1). This visualization shows that there generally is a good separation between healthy participants and participants with a coordination disorder. Most of the healthy participants are grouped in the top left corner of the graph; this group represents participants who scored probability values lower than the threshold, and were classified as healthy participants. Most of the patients are in the group dispersed between the center and the bottom right corner of the graph; this group represents participants who scored probability values higher than the threshold and were classified as patients. These findings again illustrate that in general healthy participants displayed faster and smoother FNT movements than patients. Some overlap between the two groups of participants, however, prevents a better separation. For example, some healthy participants (5, 25, and 23) score similar probabilities as patients, while one DCD patient and one EOA patient (61 and 32, respectively) score similar probabilities as healthy participants. Thus, according to model (5.1) participants 32 and 61 behave very much as healthy participants.
Figure 5.3: Visualization of the predictions of model 5.1. Black edges of the shapes represent participants classified as healthy, while green edges represent participants classified as patients.
5.3.2 SARA scores compared to GLM predictions

In Figure 5.4 GLM scores are plotted against SARA scores to gain further understanding of misclassified patients from group 1. From this figure we can observe that most of the misclassified patients had relatively low SARA scores, meaning that the observers noticed only small or no tremor at all (smooth FNT trajectories). However, some patients received a low SARA score (suggesting that the FNT trajectories looked smooth to the observers) whereas the model score was high (participants 4, 24 and 30, in the top left corner of Figure 4). From a classification point of view, this suggests that the model classifies patients not only on the presence of irregularities (tremor) which would result in high curvature.

![Diagram of model scores against SARA scores](image)

Figure 5.4: Model scores against SARA scores. SARA scores are averaged over left and right hands and observers. The vertical axis indicates the probability of having a coordination disorder. The numbers represent participants. The misclassified patients from group 1 are in the lower left corner of the plot (numbers 8, 26, 28 and 32).

To understand why the model classifies some patients with no visible or minor tremor as healthy and others (correctly) as patients, we present violin plots of the curvature and speed distributions in Figure 5.5 and 3D trajectories in Figure 5.6, for each hand of participants 4, 8, 14, 24, 30, and 31. We included participant 8 as an example where no visible tremor was observed and the model classified the patient as healthy, and participants 4 and 24 as examples where no visible tremor was observed but the model correctly classified them as patients. For comparison, we also included patients 30 and 14 where minor to moderate tremor was observed and the model correctly identified them as patients. Finally, participant 31 was included as an example of a healthy participant. To
start with this participant, the trajectory is regular and smooth (Figure 5.6) with relatively high speeds and low curvatures (Figure 5.5). In strong contrast, participant 14, who exhibited moderate tremor, had relatively low speeds and high curvatures during very irregular trajectories resulting in a high model score. A similar, although more subtle difference compared to the healthy participant (31) can be observed for participant 30, who exhibited minor tremor, but also had relatively low speeds and high curvatures during trajectories that were also irregular, although less than in participant 14. This explains the high model score for this patient, as well. Participants 4 and 24, who had no visible tremor, did have relatively low speeds and high curvatures, while their trajectories looked very similar to those of the healthy participant (31), explaining the high model score as well as why no tremor was observed. Finally, participant 8, who had no visible tremor either, but was scored as healthy by the model, indeed had curvatures and speeds that were very similar to those of the healthy participant (31).

In summary, the model seems to classify some patients with no visible or minor tremor as patients, because it picks up features from the movement trajectories that have been recorded by the IMUs and that are not visible to the naked eye. On the other hand, if the trajectory of a patient is similar to that of a healthy participant in terms of speeds and curvatures, as may be the case for some of the (mildly affected) DCD patients, it seems the patient will be classified as healthy.
Figure 5.6: 3D trajectories collected from a healthy participant and 5 patients with a diagnosed coordination disorder. The labels on the left represent participants and SARA scores.
5.4 Discussion

The goal of the present study was to apply a recently developed method to distinguish between patients with coordination disorders and controls who performed the FNT. We expected that FNT movement trajectories would be smoother and faster in healthy participants than in patients and that these movement characteristics should be reflected in lower local curvature and higher instantaneous speed values in healthy participants, which was indeed confirmed. Using local curvature and instantaneous speed as features the method achieved 84% accuracy distinguishing patients and controls.

First a (probabilistic) GLM was defined as a function of curvature and speed to estimate the probability that the FNT trajectories were collected from a patient. Then, to test the GLM on new data we performed LOOCV resulting in 84% accuracy. In addition, we expected that misclassified patients would exhibit FNT trajectories similar to those of healthy participants, exhibiting smooth trajectories as reflected in low model scores. For further understanding of misclassifications, we plotted SARA scores against model scores, as well as violin plots of the local curvature and instantaneous speed distributions of selected participants. This suggested that the model classifies some patients with no visible or minor tremor as patients, because it detects features from the movement trajectories recorded by the IMUs that are not visible to the naked eye. On the other hand, if the trajectory of a patient is similar to that of a healthy participant in terms of speed and curvature, it seems the model classifies the patient as healthy.

Our accuracy results are consistent with other studies [78, 83, 141] that tried to distinguish between healthy and pathologic FNT trials. It is remarkable that by using only two features we still achieve similar accuracy, whereas the authors in [78, 83, 141] used 33, 11, and 14 features respectively. A benefit of using only two features is the ability to visualize groups without the need of dimensionality reduction techniques such as principal component analysis [66] or Sammon mapping [117]. Additionally, only one optimal threshold (simple classifier) was used in our approach to classify healthy participants and patients.

One limitation of the present study is that the severity of ataxia in our group of participants, as assessed by neurologists, does not cover the whole range of the SARA FNT scale. The highest score provided by the neurologists is not higher than 2 (tremor smaller than 5 cm), whereas the maximum score is 4 (unable to perform the pointing movements). Including participants with more severe symptoms of a coordination disorder may change the results. However, the classification accuracy should be similar, as more severe symptoms should result in even higher curvature values and slower speed values. In other words, we could expect such participants to be in the lower right corner of Figure 5.3
(high curvature and slow speed values), where we expect the model to classify them as patients.

The need for objective and quantitative assessment of human movement to reinforce and support the use of clinical rating scales is evident [25, 32, 135]. A benefit of the presented method is that it can be applied to a broad range of human movements commonly used in clinical tests such as gait, static postural control, dynamic postural control, finger chasing, path drawing spirals, circles, squares, or figure-8 shapes, and fast alternating hand movements [119]. In addition, the methodology can be applied independently of the tracking device such as force plates or depth cameras. The quantification of smooth movements plays an important role in the assessment of coordination disorders [10]. In previous studies [128, 130] we proposed curvature as a measure of smoothness of body movements. Here, we provide additional evidence of the usefulness of this measure as a measure to differentiate pathologic from healthy movements.

In conclusion, we have shown that the (probabilistic) GLM we developed to assess dynamic balance can also be used to assess patients with coordination disorders. The method is useful to distinguish patients and healthy participants based on an instrumented version of the FNT. Future work could entail testing the reliability of the method for distinguishing groups of patients and/or controls using other clinical tests such as finger chasing or quiet standing, and using other tracking devices such as force plates or depth cameras that can be used to track body movements without the need for markers or wearable measurement devices.