Evaluation of the fine motor skills of children with DCD using the digitalised visual-motor tracking system

Ruimin Li[1,2,3], Bin Li[1,2,3], Shixiong Zhang[3], Hong Fu[3], Wai-Lun Lo[3], Jie Yu[4], Cindy H.P. Sit[4], Desheng Wen[1]

[1]*Xi'an Institute of Optics and Precision Mechanics of CAS, Xi'an, People's Republic of China*
[2]*University of Chinese Academy of Sciences, Beijing, People's Republic of China*
[3]*Department of Computer Science, Chu Hai College of Higher Education, Hong Kong*
[4]*Department of Sports Science and Physical Education, the Chinese University of Hong Kong, Hong Kong*
E-mail: hfu@chuhai.edu.hk

Published in *The Journal of Engineering*; Received on 30th September 2017; Revised on 7th December 2017;
Accepted on 20th December 2017

Abstract: The study on the coordination between vision and motion of children with developmental coordination disorder (DCD) can help understand the mechanism of DCD for timely and appropriate intervention. Whereas the existing visual-motor integrated systems rely on markers attached to the subject to track the eye gaze and body movements, which is too expensive and not suitable for DCD assessment. In this study, a markerless visual-motor tracking system which consists of an eye-tracker used to track the eye gaze and a Kinect used to capture the body movements is designed to monitor the behaviour of children in the fine motor tasks. Then the eye gaze position of the subject is matched into the motion image captured by Kinect. The current data of children placing pegs captured by the proposed system are analyzed quantitatively. We find that the visual movement speed of the children with DCD or at risk of DCD is slower than that of the typical developing children to focus on the target while their hand movement speed is almost the same. In addition to DCD analysis, the proposed system is meaningful to the monitoring of other diseases related to visual-motor coordination.

1 Introduction

Developmental coordination disorder (DCD), also known as developmental dyspraxia, is a disorder of visual-motor integration beginning in childhood that mainly affects motion control, planning of movements and coordination. About five to six per cent of school-aged children are hassled by DCD. Children with DCD seem clumsy, awkward and poorly coordinated when they perform some fine or gross tasks, but have normal or above average intellectual abilities [1]. However, their coordination disorder in motor skills may withdraw themselves from participation in physical or motor-based activities [2] and further impact their social integration, academic process and emotional development. Besides, this disorder would persist into adulthood without timely intervention, therefore making it a perpetual defect. Currently, the leading assessment tools of DCD are movement assessment battery for children-2 (Movement ABC-2) [3], DCD Questionnaire-Revised [4] and Bruininks–Oseretsky Test of Motor Proficiency-2 [5], which grades the results of participants in certain specific tasks, whereas DCD is also a disorder of motion and vision. These methods mentioned above focus on the final results of the tasks while ignoring the detailed information during the tasks. So, it is necessary to study the coordination of vision and motion [6], which helps to understand the mechanism of DCD for timely and appropriate intervention. In the course of assessment, examiners are required to observe the examinee's motion posture and gaze direction of eyes as well as to record the results and monitor the operation mistakes. Actually, it is almost impossible to quantitatively record the data of gaze direction and motion posture [7] by human observation without missing any details. Hence, the digitalised recording and analysing system is imperative in DCD assessment [8].

At present, there is rare research on the visual-motor coordination with digitalised equipment [9–12]. For example, David *et al*. [9] investigated whether obesity affected the visual motor coordination via a special task which required participants to swing the pendulum synchronically with a moving visual signal displayed on a screen. There was no digitalised system to monitor the participants' performance and the conclusion was depended on the final results of the task, lacking the analysis of the details. Actually, some companies are committed to one of these two technologies (eye-tracker and motion detection) and develop several mature products, such as SMI eye tracking glasses, Ergoneers Dikablis eye tracker, VICON motion capture systems and Qualisys track manager. With a combination of these two technologies, the coordination of visual-motor can be tracked and analysed. Essig *et al*. [13] designed the VICON-Eye Tracking Visualiser which consists of SMI mobile eye-tracking and VICON motion-capturing to calculate the 3D gaze vector and hand motion. Miles *et al*. [14] demonstrated Quiet Eye Training enhanced visuomotor coordination in children with DCD by combining Applied Science Laboratories (ASL) eye-tracker with a digitalised single-lens reflex (SLR) camera capturing participants' movements. Whereas plenty of reflective markers used for movements detection were needed to attach to subjects in the above two systems, which limited their applied range in some tasks, such as DCD assessment. Besides, both the eye-tracker system and the movements' detection system are really pricey. So, there is an urgent need to develop an inexpensive and markerless digitalised visual-motor tracking system.

In this study, a markerless digitalised system is designed to track both eye gaze and body movements in the fine motor tasks. Examinees are required to wear an eye-tracker used to track their eye gaze with a Kinect put ahead to detect the fine movements of the upper body especially the hand's motion. Then the fusion algorithm is developed to integrate eye and body movement data for an in-depth study on the mechanism of DCD. The proposed system is used to monitor the performance of children in the task of placing pegs. Also, the data captured by the system are exploited to analyse the behaviour of DCD and normal children in detail.

The rest of this paper is organised as follows. The procedure of the proposed system used in the fine motor task of placing pegs is presented in Section 2. Then in Section 3, the pre-processing of the integration of the eye gaze position with the hand movements is described in detail. In Section 4, the behaviours of two children (one with DCD and the other not) in placing pegs monitored by the proposed system are analysed to investigate the different patterns of them. Finally, Section 5 concludes the paper and makes a prospect for the future work.

2 Methods

The Movement ABC-2 is a standard test served as an assessment instrument of children with motor difficulties and it is also an authoritative tool for DCD assessment in clinical and scientific research. Participants are required to perform several tasks in a strictly specified way. In addition, the objective, quantitative information on the movement competence is provided. The tests in the Movement ABC-2 are divided into three age bands (Abs) and diverse tasks are designed for different age bands. Here, we study the behaviour of children aged 7–10 years (AB2) in the fine task of placing pegs using the proposed markerless digitalised visual-motor tracking system.

2.1 Participants

Participants were 23 children aged 7–10 years old who were randomly recruited from primary schools in Hong Kong. Children with known physical disabilities, psychiatric/emotional disorders, autistic tendencies, or neurological disease were excluded from this study. By using Movement ABC-2, children were scored by the professional instructor to determine whether suffering from DCD or not. The total score of a child is below or equal to the 5th percentile, which is considered as 'highly likely to have a motor disorder'. After evaluation, three of the 23 participants were regarded as suffering from DCD and five children were 'at risk' of having a motor disorder. Before attending the tasks, these children did not receive any special training of the projects in the tasks.

Full parental and participant's consent was acquired prior to conducting the experiments. Besides, ethical approval was obtained from the local ethics committee and the use of data in this research was permitted by the local Institutional Review Board in order to protect the human subjects. All children received a gift for participating.

2.2 Tasks

There are eight tasks in total used to measure the movement skills of children including manual dexterity (placing pegs, treading lace and drawing trial), aiming & catching (catching with two hands, throwing beanbag onto mat) and balance (one-board balance, walking heel-to-toe forwards and hopping on mats). Here, placing pegs in the manual dexterity is belonging to fine motor tasks in which the subtle movements of the upper limbs, especially the hands, are the focus of observation. The materials to do the task of placing pegs include a blue pegboard, 12 yellow mushroom pegs, a blue bank box base, a table-top mat and a stopwatch as shown in Fig. 1.

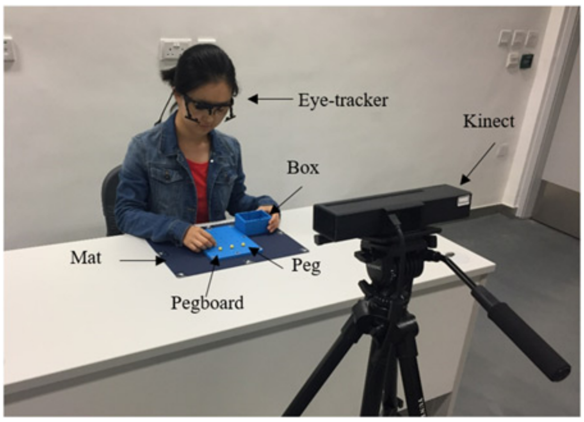


Fig. 1 *Schematic diagram of the proposed system used in placing pegs task*

In this task, the examinee is required to hold the box containing 12 pegs steady with one hand and put the other hand on the mat. At a signal given by the examiner, the examinee should pick up one peg and insert it into the pegboard as quickly as possible until the entire 12 pegs are inserted into the board. In Fig. 1, the subject is conducting the task using the right hand and aims to insert a peg into the third hole in the first line of the pegboard. The time interval from the free hand leaving the mat to the last peg inserted into the pegboard is recorded in the standard form. Both hands are tested and the preferred hand is tested firstly, then the other. The eye gaze position and the hand movements are important to study the visual-motor coordination of the subject.

2.3 System

In this section, the digitalised markerless visual-motor tracking system for fine motor skills assessment is introduced in detail. The proposed system consists of a binocular eye-tracker (Pupil Labs) for tracking the eye gaze and a Kinects V2 (Microsoft) used to capture the hand movements as shown in Fig. 1.

2.3.1 Eye-tracker: The head-mounted eye-tracker used in the proposed system is produced by Pupil Labs (Berlin, Germany). It is a binocular eye-tracker which is configured with a world camera recording the subject's field of vision and two eye cameras capturing the subject's eye movements. The headset is lightweight and adjustable which has various configurations to meet the need of diverse applications. The sampling rate of the world camera is from high resolution capture at 30 Hz to low latency (5.7 ms) 120 Hz and the field of view (FOV) of the lenses can be selected from 60° or 100°. The resolution of World Camera used in the proposed system is 1280×720 at 60 fps and the FOV of the lenses is 100°.

To ensure the robust tracking performance of eye-tracker, all the camera should be in focus with a good FOV of examinee's eyes. Then a mapping between the pupil and eye gaze coordinates is established by calibration. Manual marker calibration is selected to fit the moderate distances and wide FOV in the proposed system. Pupil's algorithms automatically detect the participant's pupil and estimate the eye gaze position with the 3d detection and mapping mode. Pupil capture saves the world video stream and all corresponding gaze data. The gaze accuracy of eye-tracker used in the proposed system is $0.60 \pm 0.08°$.

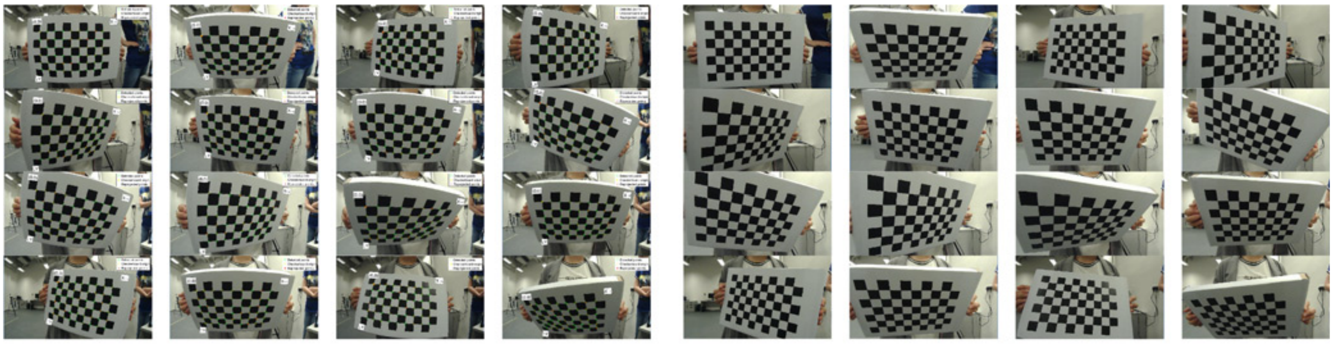
2.3.2 Microsoft Kinect: An Xbox Kinect sensor and a Kinect adapter for Windows are assembled to achieve the development of Microsoft Kinect for Windows V2 (Kinect V2). There is a depth sensor and a colour camera in the Kinect V2 sensor. The 1080p colour images are obtained by the colour camera at the sampling rate of 30 Hz. The depth sensor uses the method of time of flight to obtain depth information. The depth image is captured at the sampling rate of 30 Hz with a resolution of 512×424 pixels and the FOV of it is $70 \times 60°$es.

Unlike other body detection algorithms [15], Kinect uses the depth-image-based human skeleton tracking algorithm developed by the Microsoft Cambridge Research Institute to identify and capture the entire action without any props. It can simultaneously track six objects and 25 skeletal joints per object in a range of 0.5–4.5 m. What is especially noteworthy is that no marker is needed during skeleton detection.

3 Data pre-processing

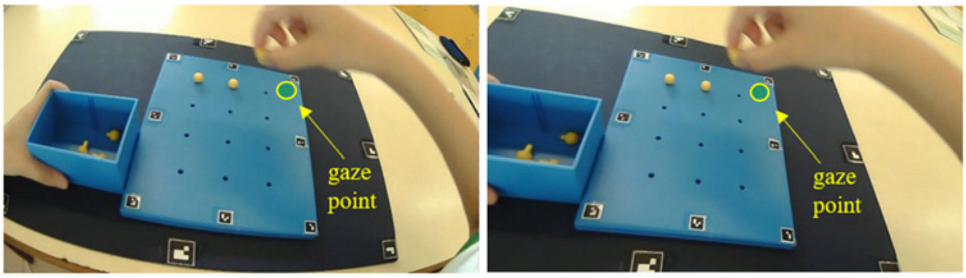
3.1 Camera de-distortion

There is a distortion in the world camera of eye-tracker. The distortion [16] like the situation that the line becomes a curve seriously impacts the accuracy of eye gaze position. Therefore, the distortion correction [17] is needed to get a more precise result and the template method is used to do the camera de-distortion here.



a *b*

Fig. 2 *Camera de-distortion using the template method*
a Calibration patterns
b Corrected results



a *b*

Fig. 3 *Example of corrected result*
a Unprocessed image
b Corrected image

The principle of the template method is as follows. A pre-made template (usually black and white grid) is needed firstly. Then take multi-shots of the template from different angles as shown in Fig. 2*a*.

Calculate the parameters using the feature points extracted from patterns on the basis of the camera distortion model. A radial distortion model is shown in formulae (1) and (2) and tangential model is in formulae (3) and (4)

$$x_1 = x_0(1 + k_1r^2 + k_2r^4 + k_3r^6), \tag{1}$$

$$y_1 = y_0(1 + k_1r^2 + k_2r^4 + k_3r^6), \tag{2}$$

$$x_1 = x_0 + [2p_1xy + p_2(r^2 + 2x^2)], \tag{3}$$

$$y_1 = y_0 + [p_1(r^2 + 2y^2) + 2p_2xy]. \tag{4}$$

Here, $[x_1, y_1]$ represents the ideal coordinate without distortion and $[x_0, y_0]$ is the coordinate of the actual image. These five parameters k_1, k_2, p_1, p_2, k_3 are the distortion factors of the camera that need to be determined in the camera calibration. Finally, the image is corrected using the resulting parameters of the world camera as shown in Fig. 3.

In Fig. 3, the green dot in the yellow circle is the eye gaze point. It is obvious that straight lines in the real world appear as straight lines in the corrected image.

3.2 Hand detection

The detailed information of hand movements is important in fine tasks. However, Kinect fails in this for it can only detect some rough action of the hand. OpenPose [18, 19] is used to detect the skeleton joints of the hand. OpenPose is an open source library used to realise real-time multiuser key-points detection with deep learning. Compared with Kinect tracking 25 key points of the whole body, OpenPose is much more fruitful especially in the finger detection and tracking. For example, Kinect perceives that a person is raising a hand, and OpenPose can actually detect that this person is pointing to something in the same scene. OpenPose can detect 21 key points of one hand, so it is well suitable for detecting fingers in fine motor tasks. The results of detecting the half-body and hands with OpenPose in the fine motor tasks are shown in Fig. 4.

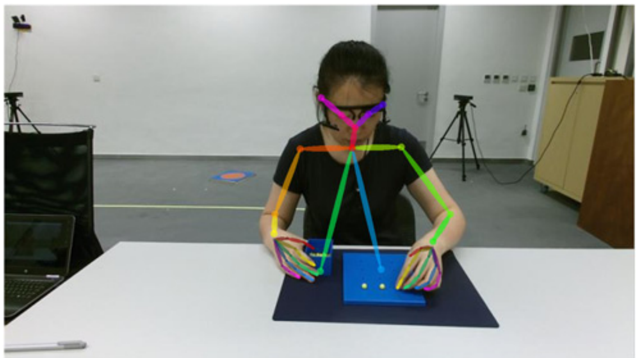
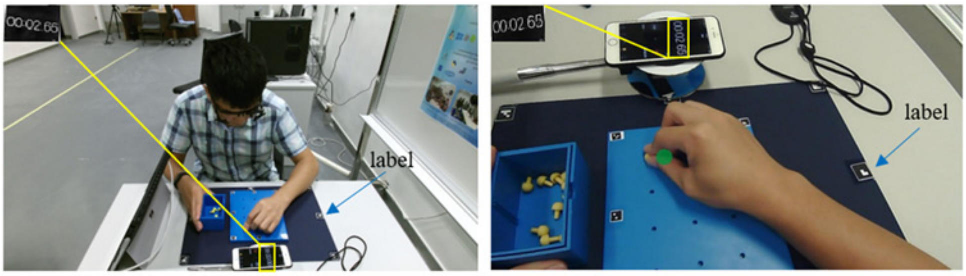


Fig. 4 *Detection results of the half-body and hands with OpenPose*

3.3 Data fusion

The eye gaze point and hand skeleton joints are required to merge into one picture for further visual-motor coordination analysis. There may be a question that why not use OpenPose to detect the hand directly from the picture obtained by the eye-tracker and



a *b*

Fig. 5 *Example of the synchronisation*
a Picture captured by Kinect
b Picture captured by eye-tracker

thus data fusion is not needed. However, the movement ranges of some examinees are so large that their hands fall out of the FOV of the eye-tracker.

It is obvious that extracting the eye gaze point into the colour image of Kinect is the best choice because the colour image includes the complete experimental scene and examinee's body. A synchronisation is required firstly due to the different sampling rates between eye-tracker (30 Hz) and Kinect (60 Hz). The current approach is to place a timer in the area which can be photographed by both eye-tracker and Kinect. Then match every image taken at a lower sampling rate with the corresponding image captured at a higher rate according to the timer. One example of the synchronisation is shown in Fig. 5.

The number of the timer can be extracted and identified. In Fig. 5, the timers display the same time in the two pictures and the time is 00:02:65, which presents these two pictures are captured at the same moment.

The approximate 180° opposite FOV between eye-tracker and Kinect brings a great challenge to the data fusion [20]. Fortunately, there are many labels on the blue pegboard and the black mat which can be used for data fusion as shown in Fig. 6. The picture captured by eye-tracker is denoted as Picture A and the one captured by Kinect is donated as Picture B for convenience. Four or more labels both in Picture A and Picture B are selected and recorded as a_1, a_2, a_3, a_4 and b_1, b_2, b_3, b_4, where $\{a_1, a_2, a_3, a_4\} \in \text{Picture } A$, $\{b_1, b_2, b_3, b_4\} \in \text{Picture } B$, and a_i and b_i are the same label in the scene. Then the transformation matrix between $\{a_1, a_2, a_3, a_4\}$ and $\{b_1, b_2, b_3, b_4\}$ is calculated and referred as $\boldsymbol{T}$. The corresponding eye gaze position in Picture B can be obtained by the formula (5) as shown below.

$$[u, v] = [x, y] \times T, \tag{5}$$

where $[x, y]$ is the eye gaze coordinate in Picture A, and $[u, v]$ is the corresponding position in Picture B. Eye gaze position in Fig. 5*a* matched in Fig. 5*b* is shown in Fig. 6. The red circle is the eye gaze point matched well. The tiny error in the matching result can be considered as a systematic error for the moment.

4 Results

4.1 Group differences at baseline

Actually, the complete data of 20 participants (3 with DCD, 4 at risk and 13 typically developing children) were obtained finally, due to some hardware malfunction. We divided those 20 participants into two groups. One is the TP group (consists of the 13 typical developing children), and the other is the R&D group (consists of 4 children at risk and 3 children with DCD). The performance scores and basic anthropometric variables of these two groups are shown in Table 1. Obviously, the Movement Assessment Battery for Children - Second Edition (MABC-2) total score of the TP group is far higher than that of the R&D group. No differences are

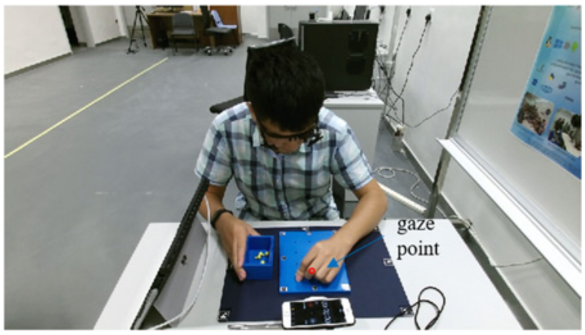


Fig. 6 *Result of eye gaze position in Fig. 5a matched in Fig. 5b*

found with regard to age and height. The difference in weight can be explained by the research that DCD is related to obesity to some degree [21].

4.2 Correlation of visual motor and hand movement

Pegs are the target of interest in this task. Actually, pegs are usually pinched by thumb and index finger when placing pegs and the thumb is especially essential to pick up pegs for most examinees who are not disabled. Therefore, the changes of the thumb tip's position can stand for the hand movements. We find that the action of placing one peg can be divided into two steps: pick up a peg from the box and insert it into the pegboard. There are some special patterns of visual-motor integration. The eye gaze of the subject fixates on the blue box to search for a proper peg to pick up before the hand touches the box and when a peg is picked up, the gaze of the subject transfers to the target hole before inserting the peg into it. Thus, it is significant to study the correlation of gaze point and thumb tip, which can explain how examinees' hand and eye work together in placing pegs.

The fusion data of two children (one with DCD, and the other not) in placing three pegs are extracted from their complete tasks of placing pegs. The external conditions of these two children are almost the same, such as gender, height, and weight. The normalised coordinates of eye gaze and thumb tip of these two children are shown in Fig. 7 in which Figs. 7*a* and *b* present the horizontal coordinate and vertical coordinate of normal child, respectively, and Figs. 7*c* and *d* are for a child with DCD.

In Fig. 7, the dotted lines presenting the position of thumb tip have obvious cyclical changes with three peaks and three troughs. The peaks of the curves (red dotted line) imply that the thumb is on the board and the troughs indicate in the box. In fact, the fingers of these two children were on the board preparing to do the test at the beginning. When received the signal, their fingers began to move to the box to get a peg, and then moved to the board to complete the inserted action. They repeated three times like that. Obviously, the trends of the dotted lines are consistent with the fact mentioned above. Then the thumb movement is used as a benchmark to analyse the eye gaze position. The solid lines stand for the eye gaze position. The trend of eye gaze curves of DCD child is almost synchronised with his thumb, and sometimes eyes move a little earlier than a thumb. While the normal child's eyes change more flexible than his thumb. The peaks and troughs of his eye gaze are regularly in advance a lot to the corresponding peaks and troughs of the thumb all the time. These indicate that the DCD child's hand and eyes move to the target almost at the same time, which gives a clumsy feeling to us. Normal child's eyes always move to the next target prior to hands. This complex control of the eye and hand movement coordination brings efficient using of time. Therefore, normal child completes the task faster than a child with DCD.

It is noteworthy that since the hole inserted through the peg is determined by the examinee himself every time and the vertical coordinates of the holes in the bottom two lines of the pegboard are nearly the same as that of the box, the change of the gaze point's coordinate is not obvious if the peg is inserted into the holes in the last two rows as shown in Fig. 7*d*. In fact, the last two pegs of DCD child are inserted into the holes in the last two rows. Fortunately, the changes of eye gaze's horizontal coordinate are unaffected, because the box and the board are placed on different sides and there is an obvious difference in their horizontal coordinates which can be used for quantitative analysis.

The correlation coefficient between the movements of eyes and hand is calculated using the following formula to give a quantitative illustration

$$R(x, y) = \frac{\mathrm{Cov}(x, y)}{\sqrt{\mathrm{Cov}(x, x)\mathrm{Cov}(y, y)}}, \tag{6}$$

where $\mathrm{Cov}(x, y)$ is the covariance matrix of the matrix consisted of x and y. The larger the absolute value of $R(x, y)$ is, the higher correlation of x and y is. A negative value of $R(x, y)$ means negative correlation and a positive value is a positive correlation. The

Table 1 Performance scores and basic anthropometric variables of those two groups

	TP group, mean ± SD	R&D group, mean ± SD
MABC-2 total standard score	54.8 ± 19.5	11.3 ± 5.9
Age, years	9.0 ± 0.8	9.3 ± 0.95
height, cm	138.8 ± 8.6	140.2 ± 5.7
weight, kg	19.2 ± 8.8	24.9 ± 8.8

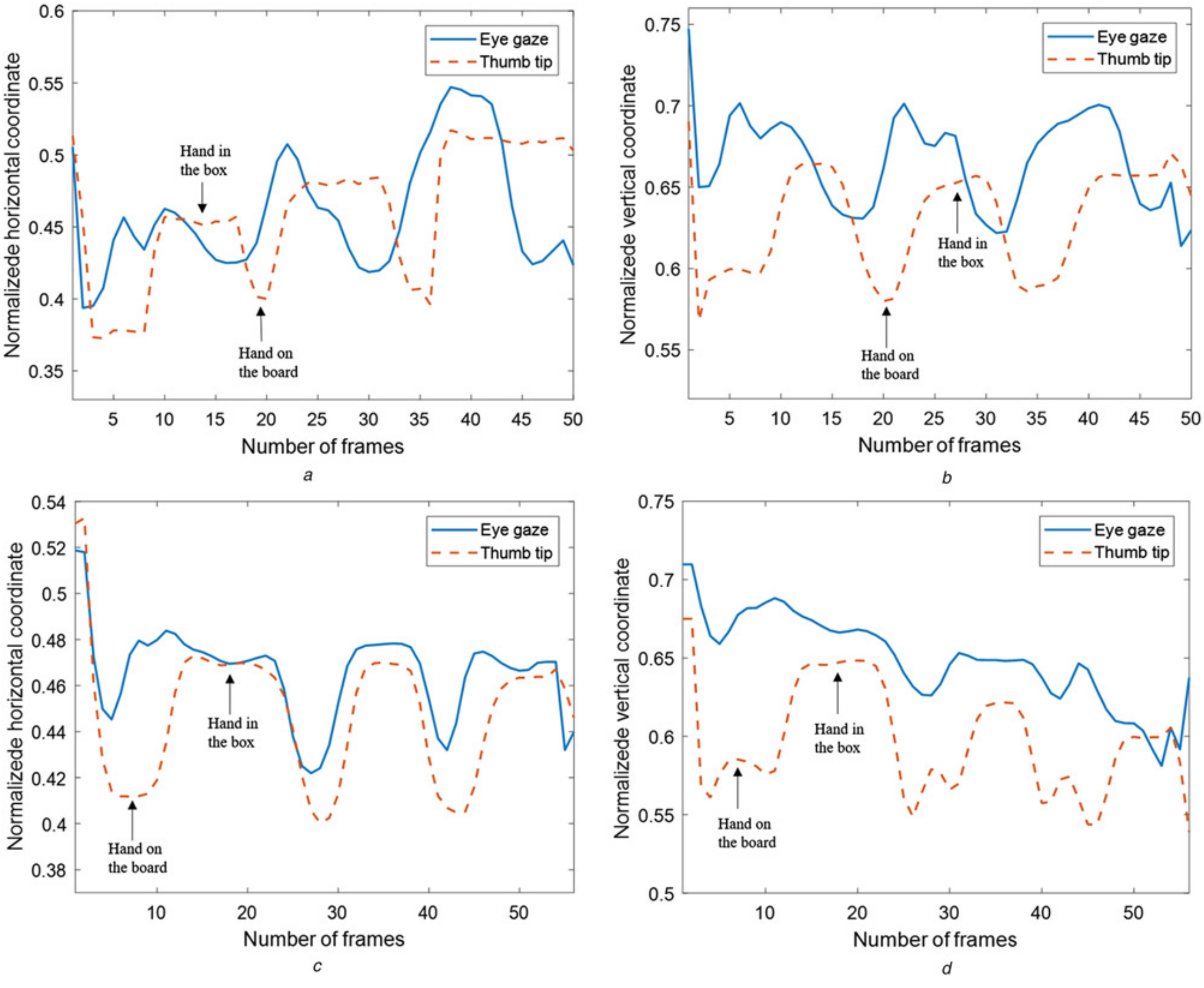


Fig. 7 *Normalised coordinates of eye gaze and thumb tip of two children (one with DCD and the other without DCD)*
a, *b* Horizontal and vertical coordinates of a normal child
c, *d* Horizontal and vertical coordinates of the child with DCD

doi: 10.1049/joe.2017.0405

correlation between eye gaze position and thumb tip position is given in Table 2.

The absolute value of correlation coefficient of DCD child is larger than that of the normal child. Since there is a time difference between these two sets data of normal child on the basis of the fact that his eyes move earlier than his hand. At the same time, we notice that the correlation coefficient of the normal child's vertical coordinate is negative, which is due to the time difference causing that the trough of eye gaze curve is exactly right at the same moment with the peak of the thumb curve. So it seems that the vertical coordinates of eye gaze and the thumb tip of the normal child are negatively related.

To quantitatively analyse the different behaviours between DCD and normal children, the time difference of the eye gaze and hand movements is evaluated by the number of frames that eye gaze moves in advanced of the thumb as shown in formula (7)

$$N_{\mathrm{TD}} = \arg\max_{i}\left(\sum_{t=1}^{n} R(g(t), h(t+i))\right), \qquad (7)$$

where N_{TD} means the number of frames that eye gaze moves in advance of the thumb, n is the number of the total frames observed, $R(x, y)$ is the correlation coefficient of x and y which is defined in formula (6). $g(t)$ and $h(t)$ are the horizontal positions of the gaze point and thumb tip, respectively. The changes of correlation coefficient over the number of frames are shown in Fig. 8.

As can be seen from Fig. 8, the number of frames that eye gaze moves ahead of the thumb is the horizontal coordinate of the point which has the largest correction coefficient on the curve. So for the child with DCD, $N_{\mathrm{TD}}^{\mathrm{DCD}} = 1$, and for the normal child, $N_{\mathrm{TD}}^{\mathrm{normal}} = 4$. The exact values of time difference can be calculated using formula (8)

$$t = \frac{1}{f} \times N_{\mathrm{TD}}, \qquad (8)$$

Table 2 Correlation coefficient of eye gaze position and thumb tip position

	Child with DCD	Normal child
correlation of horizontal coordinate	0.6171	0.3091
correlation of vertical coordinate	0.2589	−0.0720

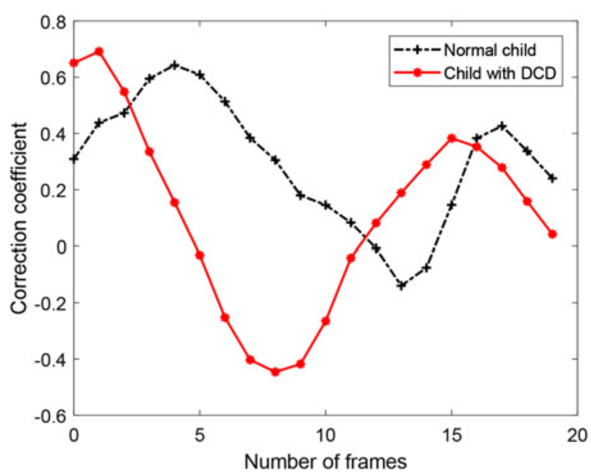


Fig. 8 *Changes of correlation coefficient over the number of frames*

Table 3 Visual movement speed and hand movement speed of the two groups

	TP group, mean ± SD	R&D group, mean ± SD
time to complete the task, s	25 ± 4.14	26.57 ± 4.24
visual movement speed, pixels/frame	0.229 ± 0.072	0.182 ± 0.080
hand movement speed, pixels/frame	0.195 ± 0.044	0.195 ± 0.056

where f is the sampling rate of Kinect and $f = 30$ Hz. So, $t^{\mathrm{DCD}} = 33$ ms and $t^{\mathrm{normal}} = 133$ ms.

4.3 *Activity-based outcomes*

To further analyse the data and give a statistical conclusion, the visual movement speed and the hand movement speed of every participant are calculated as follows:

$$S_k = \frac{\sum_{t=1} V_{t+1}}{N}, \quad (V_{t+1} > \mathrm{Th}). \qquad (9)$$

In the formula (9), S_k means the visual movement speed or hand movement speed of the k-th participant. $V_{t+1} = \sqrt{(x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2}$ is the instantaneous speed of the $(t+1)$-th frame. (x_t, y_t), (x_{t+1}, y_{t+1}) are the coordinate positions of eye gaze or thumb point in the tth frame and $(t+1)$th frame, respectively. Actually, the movement of the hand or eye gaze between most two frames is tiny. The value of the instantaneous speed of such two frames exactly capturing the moment that hand or eye gaze shifts between box and board is larger, which is meaningful to us. So, a threshold presented by Th is set to reserve the larger instantaneous speed and discard the smaller values. N is the total number of the instantaneous speed which is larger than Th. Here, Th = 0.1, which is counted from the statistics histogram. The statistics result of two groups' visual and hand movement speed is shown in Table 3.

As shown in Table 3, TP Group completes the task 1.57 seconds faster than the R&D group on average. But, the hand movement speed of those two groups is nearly the same, which indicates that children with DCD or at risk have the same flexible fingers as the typically developing children. Whereas, there is a significant difference on the visual movement speed of these two groups. The visual movement speed of the R&D group is 0.047 pixel/frame slower than that of the TP group. Furthermore, we find that TP group's eyes move 0.034 pixel/frame faster than their hands. As for R&D group, their eyes move a little bit slower than their hand. Compared with typical developing peers, children with DCD or at risk are lack of flexible eyes' movement which coordinates with the hand movement. This result is consistent with the study that there is some difference in the brain structure and function of children with DCD compared with the typically developing children [22], which affects the anticipatory planning and reduces the flexibility of some movement skills.

5 Conclusions

In summary, a markerless digitalised visual-motor integrated system is proposed to quantitatively analyse the coordination of eyes and hand movements of children with DCD in the fine motor tasks. The proposed system consists of an eye-tracker capturing the eye gaze of the subject and a Kinect obtaining the hand movements of the subject. The eye gaze position after distortion correction is matched into the colour picture from Kinect using a fusion algorithm. Then the key points of the hand are detected with OpenPose in the fusion images and thus the eye gaze point and the key points of the hand are integrated into an identical coordinate system for further analysis. The proposed system is used to monitor the behaviours of children in the fine task of placing pegs. The correlation coefficient of eye gaze position and thumb tip position is calculated to explore the pattern of visual-motor coordination. There is a conclusion derived from the current data that the visual movement speed of the children with DCD or at risk is 0.047 pixel/frame slower than that of the typical developing children to focus on the target of interest while their hand movement speed is almost the same, which gives us a new understanding of the DCD mechanism. Children with DCD look clumsy not because of the slow action of hands, but because they cannot capture and track the target flexibly (shown in the eyes). By the way, the proposed system cannot apply to children who wear glasses due to the head-mounted eye-tracker.

In the future, a more robust fusion algorithm will be developed to improve the matching accuracy. Also, more children will be invited to participate in the experiments and a statistical conclusion will be made, which needs the help of our collaborators because children with DCD are few. In addition, the proposed system is also meaningful for the monitoring of other diseases related to visual-motor coordination.

6 Acknowledgment

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project Reference No. UGC/FDS13/E02/16).

7 References

[1] Cermak S., Larkin D.: 'Developmental coordination disorder' (Delmar Thomson Learning, Albany, NY, USA, 2002)

[2] Yu J., Sit C.H., Capio C.M., *et al.*: 'Fundamental movement skills proficiency in children with developmental coordination disorder: does physical self-concept matter?', *Disabil. Rehabil.*, 2015, **38**, (1), pp. 45–51

[3] Henderson S.E., Sugden D.A., Barnett A.L.: 'Movement assessment battery for children-2 second edition (movement ABC-2)' (Pearson Assessment, London, UK, 2007)

[4] Blank R., Smits E.B., Polatajko H.J., *et al.*: 'European academy for childhood disability (EACD): recommendations on the definition, diagnosis and intervention of developmental coordination disorder (long version)', *Dev. Med. Child Neurol.*, 2012, **54**, (1), pp. 54–93, doi: 10.1111/j.1469-8749.2011.04171.x

[5] Bruininks R., Bruininks B.: 'Bruininks–Oseretsky test of motor proficiency' (NCS Pearson, Minneapolis, MN, USA, 2005, 2nd edn)

[6] Zhang H., Guo F., Zhang M., *et al.*: 'Human motion correction and representation method from motion camera', *J. Eng.*, 2017, pp. 370–375, doi: 10.1049/joe.2017.0159

[7] Jin K., Jiang M., Kong J., *et al.*: 'Action recognition using vague division DMMs', *J. Eng.*, 2017, pp. 77–84, doi: 10.1049/joe.2016.0330

[8] Li S., Li B., Zhang S., *et al.*: 'A markerless visual-motor tracking system for behavior monitoring in DCD assessment'. Asia-Pacific Signal and Information Processsing Association Annual Summit and Conf., Kuala Lumpur, Malaysia, 12–15 December 2017

[9] Gaul D., Mat A., O'Shea D., *et al.*: 'Impaired visual motor coordination in obese adults', *J. Obesity*, 2016, pp. 1–8, doi: 10.1155/2016/6178575

[10] Heiz J., Barisnikov K.: 'Visual-motor integration, visual perception and motor coordination in a population with Williams syndrome and in typically developing children', *J. Intell. Disabil. Res.*, 2016, **60**, (10), pp. 945–955

[11] Nicola K., Watter P.: 'Visual-motor integration performance in children with severe specific language impairment', *Child: Care Health Dev.*, 2016, **42**, (5), pp. 742–749

[12] Baglioni V., Neri V., Silvestri P.R., *et al.*: 'Motor ability and visual-motor integration in children affected by tic disorder', *Prev. Res.*, 2013, **2**, pp. 22–26, doi: 10.11138/pr/2013.2.1.022

[13] Essig K., Dornbusch D., Prinzhorn D., *et al.*: 'Automatic analysis of 3D gaze coordinates on scene objects using data from eye-tracking and motion-capture systems'. Proc. Symp. on Eye Tracking Research and Applications, March 2012, pp. 37–44, doi: 10.1145/2168556.2168561

[14] Miles C., Wood G., Vine S., *et al.*: 'Quiet eye training facilitates visuomotor coordination in children with developmental coordination disorder', *Res. Dev. Disabil.*, 2015, **40**, pp. 31–41, doi: 10.1016/j.ridd.2015.01.005

[15] Xia D., Li S.: 'Rotation angle recovery for rotation invariant detector in lying pose human body detection', *J. Eng.*, 2015, pp. 160–163, doi: 10.1049/joe.2015.0032

[16] Kim T.: 'Analysis on the characteristics of camera lens distortion', *Indian J. Sci. Technol.*, 2016, **9**, (35), doi: 10.17485/ijst/2016/v9i35/101776

[17] Shih S., Hung Y., Lin W.: 'When should we consider lens disorder in camera calibration', *Pattern Recognit.*, 1995, **28**, (3), pp. 447–461, doi: 10.1016/0031-3203(94)00107-w

[18] Cao Z., Simon T., Wei S., *et al.*: 'Realtime multi-person 2d pose estimation using part affinity fields'. Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2017, pp. 7291–7299

[19] Simon T., Joo H., Matthews I., *et al.*: 'Hand keypoint detection in single images using multiview bootstrapping'. Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2017

[20] Ming A., Ma H.: 'Region-SIFT descriptor based correspondence between multiple cameras', *Chin. J. Comput.*, 2009, **31**, (4), pp. 650–661

[21] Faught B., Demetriades S., Hay J., *et al.*: 'Does relative body fat influence the movement ABC-2 assessment in children with and without developmental coordination disorder', *Res. Dev. Disabil.*, 2013, **34**, (12), pp. 4433–4438

[22] Visser J.: 'Developmental coordination disorder: a review of research on subtypes and comorbidities', *Hum. Mov. Sci.*, 2003, **22**, (4–5), pp. 479–493