

Received May 13, 2021, accepted July 14, 2021, date of publication August 2, 2021, date of current version August 10, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3101964

Kinect Validation of Ergonomics in Human Pick and Place Activities Through Lateral Automatic Posture Detection

ERNESTO ROCHA-IBARRA¹, MARVELLA-IZAMAR OROS-FLORES²,
DORA-LUZ ALMANZA-OJEDA², (Member, IEEE), GABRIEL-ARMANDO LUGO-BUSTILLO²,
ANDRES ROSALES-CASTELLANOS³, MARIO-ALBERTO IBARRA-MANZANO^{1,2}, (Member, IEEE),
AND JUAN CARLOS GOMEZ^{1,2}

¹Departamento de Arte y Empresa, DICIS, Universidad de Guanajuato, Salamanca 36885, Mexico

²Departamento de Ingeniería Electrónica, DICIS, Universidad de Guanajuato, Salamanca 36885, Mexico

³Department of Computing Science, University of Alberta, Edmonton, AB T6G 2R3, Canada

Corresponding author: Juan Carlos Gomez (jc.gomez@ugto.mx)

ABSTRACT In this paper we evaluate a system based on the Microsoft KinectTM sensor, aimed at the automatic detection of risk postures during human work activities. We first introduce a *pick and place* task, where three different lateral standing subjects move light cardboard boxes from the various levels of a bookcase to its top, and then putting them back to their original places. They repeat the task over several work cycles and we capture all their natural movements in a continuous way using Kinect, storing the joint positions and the color images. Secondly, from the joint positions, our system detects specific risk postures following the definitions of the Rapid Upper Limb Assessment (RULA) method. Finally, we compare the posture detections by our system with the baseline detections made by a panel of five experts who used the captured color images. In our study we find that the experts have problems to distinguish among some RULA postures during a work cycle because of the narrow detection margin and the difficulty to perceive if a limb reached a certain position; which is particularly true for the cases of wrist and neck. This leads to a larger false positive rate and to a lower general accuracy, with our system detecting postures that experts do not. After applying a $\pm 1^\circ$ of relaxation to our system, which is negligible for human perception, we are able to reach an accuracy of 0.93 in the comparison with the baseline. Our results show the suitability of Kinect for lateral risk posture detection in pick and place activities.

INDEX TERMS Ergonomics, Kinect, RULA, workplace MSDs.

I. INTRODUCTION

Musculoskeletal disorders (MSDs) are injuries that affect human body movements or the musculoskeletal system (i.e., muscles, tendons, ligaments, nerves, discs, blood vessels, etc.),¹ whose symptoms include pain, discomfort, numbness and tingling in the affected area. MSDs are known to be the single largest category of nonfatal workplace injuries that can lead to temporary or permanent disability, and are responsible for almost 30% of all worker compensation costs.² In the U.S., it is estimated that companies in 2011 spent about

50 billion USD on direct MSDs-related costs, with an average expenditure of 15, 000 USD per incident. Moreover, it is also estimated that the indirect costs can add up to five times the direct costs of MSDs.³

MSDs represent a trans-disciplinary phenomenon that affects workers in all types of industries, organizations and activities. There are several factors that influence MSDs at the workplace, including socio-demographic, psychological and physical [1]. Socio-demographic factors include age, education level, gender, smoking/drinking habits and work hours. Psycho-social factors include job demands, job control and job satisfaction. Physical factors include lifting heavy items, bending, static postures, vibrations, reaching overhead,

The associate editor coordinating the review of this manuscript and approving it for publication was Kin Kee Chow.

¹<http://ergo-plus.com/definition-of-musculoskeletal-disorder-msd/>

²<http://www.bls.gov/iif/oshsum.htm>

³<https://www.osha.gov/dcs/smallbusiness/safetypays/estimator.html>

pushing and pulling heavy loads, working in awkward body postures and performing the same or similar tasks repetitively.

Assessing the ergonomics of workers during daily work activities helps to control and take actions in order to minimize the likelihood of suffering a MSD. As such, there are several methods to measure the risk factors a person can be affected by during a work day. These methods can be generally classified into three types [2]: observational, direct and self-report. Observational methods [3]–[6], involve observing a person during the course of his/her work day. They have the advantage of being easy to use, cheap, and applicable to a large number of activities and a large number of participants. Nevertheless, postural data collection produces data that is not very accurate, given the high dynamics of real-world environments and the dependency on human expert evaluation, producing only broad results. On the other hand, direct methods [7]–[9] make use of instruments and sensors attached to the individuals in order to collect information about their postures and movements. Data collected is accurate, but the cost is high, and they are invasive and therefore hard to use in real-world scenarios. Additionally, there have been some attempts to put together techniques and devices from the two previous methodologies, in order to create more complete tools [10], [11]. In the two previous type of methods, a visualization of the worker from an external point is necessary, either by an expert or by a technological instrument [12]. In contrast, self-reports [13] are self-evaluations where workers write questionnaires or work diaries to report pain, postural discomfort or levels of effort. Self-reports can include paper surveys, or questionnaires through the web or a mobile application. Self-reports have some advantages, since they can be less expensive than direct and observational methods for evaluating large populations, or in conditions where the space of research is limited, or where privacy must be maintained [14]. Nevertheless, there is a low level of reliability with this type of methods because the perceptions of the workers on the postures are imprecise and the pain thresholds are subjective.

One strategy to help implementing observational methods in real-world dynamic environments consists of video recording or taking pictures of the workers' activities. Afterwards, experts can visualize the recorded images in slow motion and analyze the data in detail to assess the ergonomics. Using such devices, human experts can visually detect a number of postural variables reliably [15]–[17]. Nevertheless, this methodology still depends on highly-skilled human experts to discover the risk patterns, which is a slow process that can be costly.

More modern methods for ergonomic assessment try to automatize the process of detecting diverse postural variables by using cameras/sensors arrays, and applying computer vision and movement/joint detection [18]. It is noteworthy to mention the recent use of low-cost sensors, such as Microsoft KinectTM [19]–[23] as a good alternative to more expensive devices [7], [24], [25]. Nevertheless, most studies where such technology is used focus on strictly controlled experiments,

where subjects are instructed to perform specific (sometimes fixed) postures to capture data, and then evaluate methods to detect those postures automatically [2], [26]–[28].

In this paper we propose to enrich the studies in ergonomics in work environments by assessing the use of Kinect for automatic risk evaluation of postures in human work activities. In order to conduct the assessment, we first use Kinect to collect postural data of three lateral standing individuals performing the same task independently. With Kinect capturing the joint positions and the images from the subjects when performing the task. For this study, we designed a *pick and place* task that consisted of repetitive short work cycles where each individual had to move small light cardboard boxes from the various levels of a three-level bookcase to its top, and the other way around. Pick and place is an activity that many workers in different industries conduct on a daily basis. For example, in the Bajío region in Mexico, where there are installed several car assembly factories and many metal mechanic manufactures, the companies indicate the pick and place activities as some of the most predominant in their work routines.

In our work, we collected the postural data with Kinect for the whole activity in a continuous dynamic way. That is, the subjects did not stop to simulate a specific posture but performed natural movements. After the capture, with the joint positions we used an in-house developed framework to analyze them automatically and evaluate the ergonomic risk in each posture, following the definitions of the observational method RULA (Rapid Upper Limb Assessment) [5]. We focused on the analysis of the upper arms, lower arms, wrists, neck and trunk, since these are the most relevant body parts for RULA, and the most convenient to observe using Kinect. Finally, similar to other works in the literature [2], we use the observational evaluation made by five human experts, that analyzed the corresponding color images captured by Kinect, as the baseline to compare the results of the automatic evaluation.

We have two research hypotheses:

- 1) It is possible to detect lateral standing ergonomic postures using the Kinect joint data that is captured in a continuous dynamic way for a pick and place activity.
- 2) It is hard for the experts to distinguish among some RULA postures for a subject during a work cycle, in particular the cases of wrist and neck.

In our study, we indeed find that for the human eye it is hard to distinguish correctly among some RULA postures during a work cycle for the wrist, neck and trunk, because their detection margins are narrow or because it was difficult to perceive if the limb reached a certain position. The difficulty to perceive some postures is observed when analyzing the images captured by Kinect; the observability rate of postures by the expert panel; and the false positive rate (a false positive is a posture detected by Kinect but not observed by the expert panel). We thus apply in our system a relaxation angle in the detection margin on the outputs of Kinect, producing

an increase in the general accuracy by allowing a softer match between automatic and human detections. With only $\pm 1^\circ$ of relaxation, that is negligible for human perception, we reached an accuracy of 0.93 when matching the Kinect detections with the experts evaluation.

Our experiments present encouraging results for the lateral detection of ergonomic postures for pick and place activities; reinforcing the suitability of the use of Kinect for the study of ergonomics in work environments.

The rest of the paper is organized as follows: in Section 2, we briefly review the relevant literature; in Section 3, we describe the general RULA method and the simplified version we used in this study; in Section 4, we present and explain our methodology; in Section 5, we show and discuss the experimental results; and in Section 6, we conclude the paper with a general overview and possible future research directions.

II. LITERATURE REVIEW

In order to reduce health issues and high costs related with bad postures, several methods have been developed over the years to record and detect those bad ergonomic postures during work activities. The first methods were manual and observational, such as RULA [5], REBA (Rapid Entire Body Assessment) [3], OWAS (Ovako Working posture Assessment System) [29], and EAWS (Ergonomic Assessment Worksheet), to name but a few. With these methods, an expert uses a working sheet and follows the workers to record every occasion in which they execute an awkward or bad posture. More modern methodologies use recorded videos and images to help experts to detect problematic postures visually [15], [16].

The most recent observational methods are based on using video cameras, computer vision and sensors in order to automatize the ergonomic postures detection process. Among the sensor-based methods, Kinect [30] has gained popularity because of its low cost, portability and user-friendly interface. In [31], the author used Kinect to capture relative 3D coordinates of four 0.10 m cubes in a range of 1 to 3 m, and compared the estimated postures with the ones of the Vicon system. He established that, with some corrections, Kinect can be used as a portable 3D motion capture system to perform ergonomic field assessments. In [32], the authors used Kinect for posture analysis and compared the results with those of an observational method. Their results showed that Kinect could be incorporated into an effective and efficient system to help with injury prevention, and made it clear that the use of Kinect significantly reduced costs when compared to the use of experts and manual evaluation. In [33], the authors performed real-time evaluations of assembly line operations, using Kinect for 3D motion detection and RULA as the posture evaluator. Their system produced an alert when the operator took a risk posture or could become injured. They encountered challenges in determining voxel size as a control factor for the accuracy and performance of joint angle calculations; in particular, they had problems to track the

movements of the wrist. In [34] and [35], the authors developed a system based on Kinect, RULA and REBA, and used it to detect postures of seated workers in order to prevent the syndrome of the office workers, caused by sitting too long in front of the computer. Additionally, the system used a data mining classification for detection of prolonged sitting, and it gave an alert to the user when it detected unhealthy postures. In [2], the authors used Kinect to extract 3D information of persons performing specific predefined postures, and validated its use to assess their ergonomics. Their system estimated scores for OWAS and compared the system scores with the ones provided by human experts, which used videos and photographs to produce an evaluation. They found that the sensor position angle affects the precision in posture classification. In [36], the authors developed a system that used Kinect to extract features of depth images and trained a classifier with such features to categorize postures based on EAWS. The system calculated a score for EAWS and compared it with the one calculated manually. It was applied by adjusting the height of a work table when the worker's posture was uncomfortable. In [37], the authors built a system that analyzed human movements from a sequence of depth images acquired using Kinect. The system calculated the information of the dynamic movement of the body and at the same time classified each movement. In [19], the authors proposed a method to correct the output data of Kinect to see if there were occlusions that could produce errors when detecting the body joints. The method used prerecorded postures to help with the correction, and calculated a RULA score using both the corrected and the uncorrected data. The results were compared with those obtained using the Vicon camera in a work environment with and without obstruction, and found that the corrected data produced lower detection errors. In [24], the authors presented a tool called K2RULA to detect RULA postures using the depth camera of Kinect. They conducted experiments using 15 predefined fixed postures performed by an actor in a controlled scenario. The authors validated their tool in two ways; first, by comparing the K2RULA scores with those obtained with the BTS SMART-DX 5000 optical motion capture system, finding a statistically perfect match, and second, by comparing the system scores with those obtained by a RULA expert rater, finding again a statistically perfect match.

Our study differs from others in the literature in two main aspects. First, in the scenario, where we are considering a worker that is situated in front of a machine/furniture performing a repetitive task. In this case, it is not possible to locate the sensor in a frontal view, that is the common consideration in most of the related works, because of the obstruction caused by the machine/furniture. Locating the sensor to capture the subjects in a lateral view, helps to avoid such obstruction; considering that the precision of Kinect in frontal and lateral views is similar as shown in [38], where the authors track upper body joints with Kinect and with a BTS multi-view stereo system as the ground truth, finding

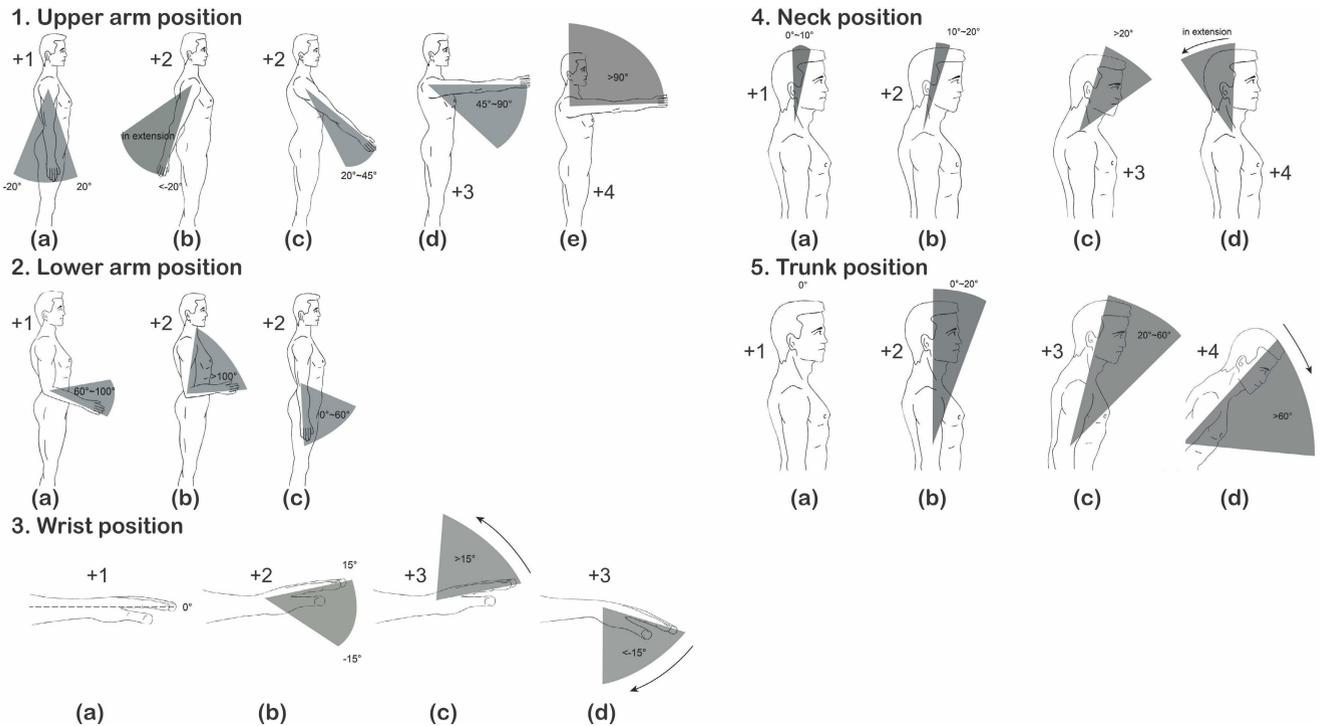


FIGURE 1. Body parts, postures and points in RULA considered in this study.

a precision of 0.25 degrees for frontal and lateral views for several joints.

The second difference, is that we collected the postural data for a whole activity in a dynamic way, with the subjects performing continuous natural movements, while in most of the works in the literature they captured fixed postures performed by actors in very controlled and specific scenarios, which reduces the applicability of the methods.

III. RULA

RULA [5] is a well-established method used to evaluate the risk level of a worker suffering an upper limb injury, based on an analysis of the body posture, muscle use, load weight, task duration and task frequency. For the analysis, the method uses a worksheet, where it first divides the human body into two groups, A and B. Group A corresponds to upper arms, lower arms and wrists, and group B to neck, trunk and legs. Second, it evaluates the individual postures of each body part and rates them in a scale from one to four, depending on the body part, with one being the most comfortable posture, and the individual scores are recorded on the worksheet. Third, the method considers extra postures that affect the individual scores, such as if the shoulders are raised or the wrist, neck or trunk are twisted. Fourth, it uses two tables to assign two aggregate scores between one and nine to group A and group B independently, depending on the individual scores of all the group’s body parts. Fifth, it considers two other factors for each group: muscle use and weight load. Muscle use adds one

point if the posture is held for more than ten minutes or if the posture occurs more than four times per minute. Weight load (a weight that is loading by a worker) adds from one to three points depending on the weight the worker is lifting/moving. Finally, RULA uses another table to assign a global score of one to seven.

In this study, we have chosen RULA because it is easy to implement and understand, and it evaluates different risk factors of workers in their environment quickly but concisely. Additionally, it is a recommended method when someone wants to have a general overview of the workers’ ergonomics before introducing a more complete and expensive ergonomic plan. Furthermore, the information obtained from RULA is sufficient to detect possible injuries produced by incorrect postures or repetitive movements executed during the work cycle.

An additional reason for using RULA is that it evaluates postures that are feasible to capture and analyze using an automatic observational approach based on Kinect; this makes it possible to compare automatic and human assessments. For practical reasons, in this study we did not use the complete version of the RULA method. We did not consider wrist, neck or trunk twists, because it was not feasible to capture them in our context. Also, the load was not considered because it could not be detected directly using Kinect. Finally, the legs were not detected because RULA only considers them if at any point in the task the person discharges the weight of his body on a single foot, which was not relevant in our study given the design of the task. Nevertheless,

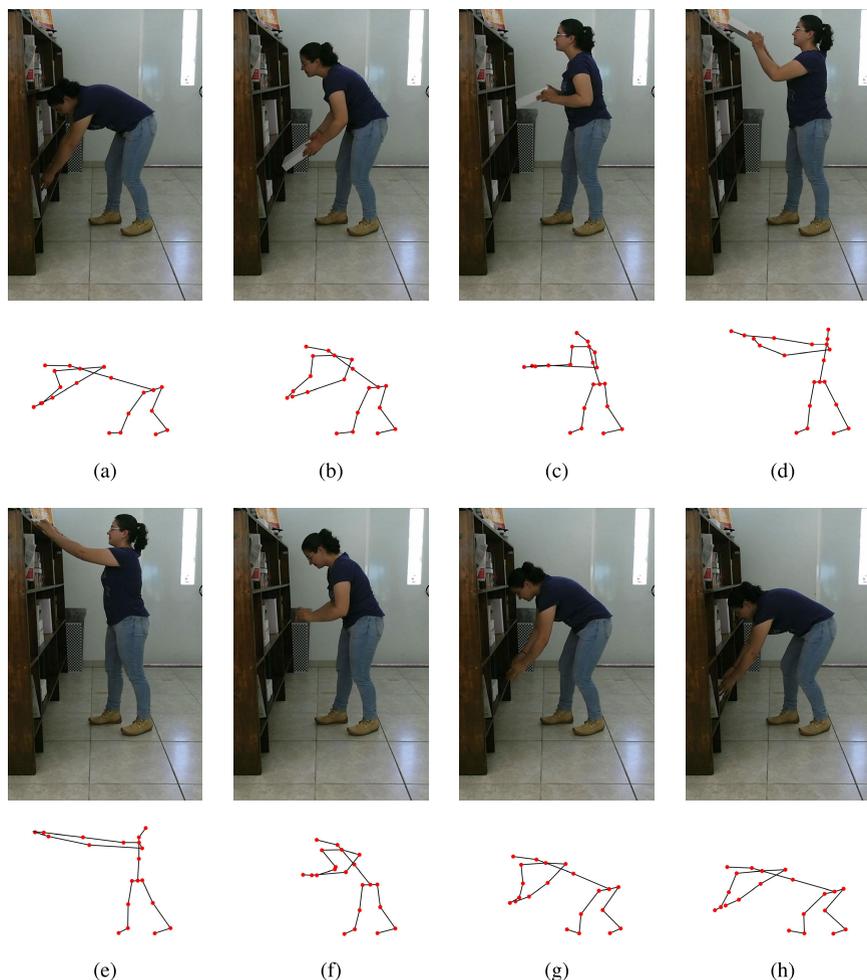


FIGURE 2. A work cycle of the performed pick and place activity. The first and third rows show some of the color images acquired, and the second and fourth rows depict the corresponding joint positions of the subject on the scene.

the impact in the final RULA score of the twists in neck or trunk, or the leg rising, is minimal.

We thus focused on analyzing the postures of the upper arms, lower arms, wrists, neck and trunk, considering the different angles defined by RULA to determine specific postures. Fig. 1 shows the body parts that we considered in this study for analysis, together with the different postures of each body part and the corresponding points in RULA for each posture.

IV. MATERIALS AND METHODS

A. EXPERIMENTAL SETUP

We conducted our experiments inside a room with natural light where three subjects standing in lateral view performed cycles of a pick and place task, and we used a Kinect One™ sensor to capture the postures of the subjects. We selected subjects of different height, weight and sex to ensure diversity, as shown in Table 1.

The Kinect was located approximately at 2.9 m in front of the scene and over a table at 0.5 m from the floor.

TABLE 1. Physical features of the subjects.

Sex	Age	Height (m)	Weight (kg)
Male	24	1.74	74
Male	29	1.83	80
Female	24	1.59	70
Ave/std	25.6(±2.9)	1.72(±0.12)	74.6(±5.0)

We determined the position of the sensor experimentally, in order to capture the complete body and extremities of each subject.

The designed task consisted on repetitive short work cycles where each subject had to move nine empty cardboard boxes measuring $0.25 \times 0.21 \times 0.05$ m, from the various levels of a three-level bookcase to its top, and the other way around. The levels were located at three different heights from the floor: 0.24, 0.66 and 1.07 m respectively, with each level containing three boxes, and the top of the bookcase was at 1.5 m. The subjects were located at 0.60 m from the bookcase.

In our design, a work cycle consisted on moving one box from a given level to the top of the bookcase or to move



FIGURE 3. Joints detected in the human body by Kinect.

one box from the top to its original place. Each of the three subjects performed 18 work cycles, moving the nine boxes to the top and then putting them back. The subjects performed the tasks in a natural continuous and dynamic way, without stopping in any particular position. We recorded the work cycles at a rate of 10 frames per second to capture color images and joint positions using the Microsoft Kinect Development Kit and a desktop PC with 16 GB in RAM, one Corei7-3960x 3.30 GHz microprocessor and Windows 10. We acquired an average of 40 images per work cycle per subject, for a total of 2,160 images.

Fig. 2 illustrates some of the color images acquired during one of the work cycles. The images below depict the corresponding joint positions of the subjects on the scene. Note that, images (a) and (h) show the subject at the initial and final postures, respectively, of the work cycle. Thus, a complete work cycle consisted of a subject taking the box from one of the three levels, putting it on top of the bookcase and moving his hands back to the initial posture just before being ready to pick up another box.

B. JOINT ANGLES COMPUTATION

Fig. 3 depicts the location of the 25 joints detected by Kinect on a human body. Kinect provided the 3D location (x,y,z) of each joint. Using these locations we obtained the angles used by our system to evaluate the RULA postures. Fig. 4 illustrates the specific joints and the directions of the angles used to calculate the postures considered in this study, in accordance with RULA as specified in Fig. 1. We evaluated a total of 18 postures for the upper arms (five postures), lower arms (three postures), wrists (four postures), neck (four postures) and trunk (four postures).

The system used joints 8, 9 and 16 to measure the upper arm postures. It measured arm postures (see Fig. 4)

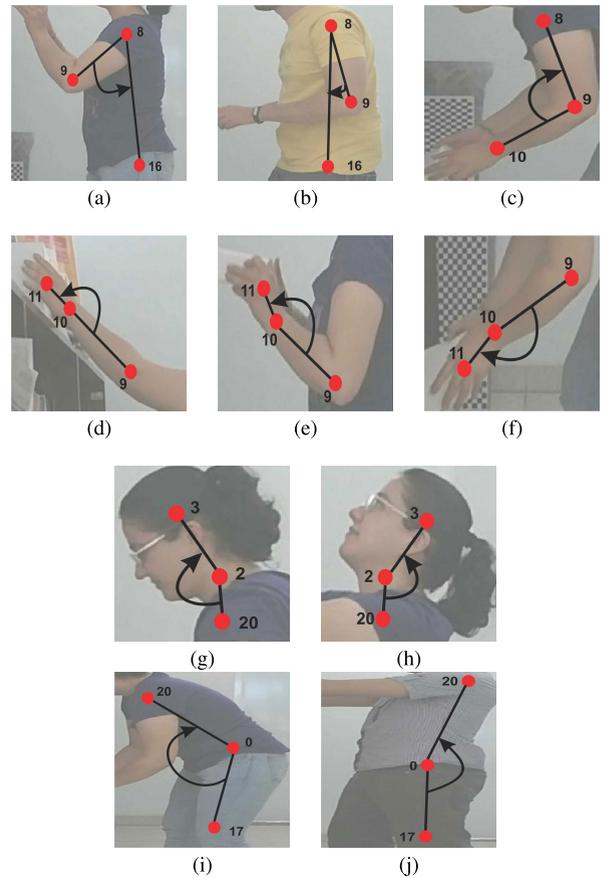


FIGURE 4. Joints and angles used with Kinect for detecting the considered RULA postures.

1.a ($0^\circ - \pm 20^\circ$), 1.c ($20^\circ - 45^\circ$), 1.c ($45^\circ - 90^\circ$) and 1.d ($> 90^\circ$) considering the angle counterclockwise, as illustrated in Fig. 4a. For the in-extension posture 1.b ($< -20^\circ$) of the arm, the system considered the clockwise angle to locate the arm posture, as shown in Fig. 4b. To locate the lower arm posture, the system used joints 8, 9 and 10 with a clockwise angle (see Fig. 4c). Using such joints, it located 3 postures of the lower arm: 2.a ($60^\circ - 100^\circ$), 2.b ($0^\circ - 60^\circ$) and 2.c ($> 100^\circ$).

The system used joints 9, 10 and 11 to locate the wrist postures 3.a, 3.b, 3.c and 3.d (see Fig. 4d-f). It measured the posture 3.a (0°) by evaluating if the points were colinear. It measured the posture 3.d calculating the complementary angle of the counterclockwise angle.

The system computed the neck postures using joints 3, 2 and 20 measuring clockwise, as shown in Fig. 4g and Fig. 4h. Finally, the system used joints 17, 0 and 20 to calculate the last postures corresponding to the trunk.

C. EXPERT EVALUATION

A committee composed of five experts visualized the recorded color images captured with the Kinect sensor, and separately evaluated the same postures as our system. The experts are researchers with postgraduate degrees in

engineering or related areas, and trained to identify postures as specified in the RULA worksheet.

We obtained a total of 2,160 recorded images that were analyzed by the experts, with an average of 40 images per work cycle per subject. The experts made their evaluation independently of each other. For each subject, we obtained five worksheets reporting all the risk postures observed during a work cycle per subject. In each worksheet it was recorded a 1 in case the posture was observed, and 0 otherwise. The final evaluation was obtained based on a consensus of the five expert results, yielding a unique worksheet for each subject evaluated, indicating whether a posture was observed or not. The evaluations were made per posture, per subject, per work cycle, for a total of 1,134 evaluations. The results of the expert committee were used as a baseline to compare and evaluate our system results.

A deeper analysis of the expert committee results was performed to know the real percentage of votes when a body part position was observed or not during a work cycle. Considering five expert results, then six events could define the observability of a specific posture: 1) all experts agreed to observe the posture, 2) 4 of 5 experts agreed, 3) 3 of 5 agreed, 4) 2 of 5 agreed, 5) only 1 observed the posture, 6) nobody observed the posture. Events 1 and 2 labelled the posture as “observable” with a high percentage of votes. In contrast, events 5 and 6 labelled the posture as “not observable” (abbreviated as “Not obs.”) with a high percentage of votes. Finally, events 3) and 4) group the cases “observable” or “not observable” with a minimal difference of votes, and this case is labelled as “unsure”.

Table 2 illustrates the observability rate of all the postures by the expert committee during the evaluation of all the work cycles of the three subjects. The numerical values denote the rates of “observable”, “not observable” and “unsure” per posture. The values of each rate goes between 0 and 1, and the sum of the values of the three rates per posture is 1. High values in a column mean the experts agree about they observe, not observe, or were unsure about a posture. Values in the unsure column means there is a degree of uncertainty.

In the case of the upper arm, we have in all postures a high degree of agreement for the observable or not observable cases, which means there is a low degree of uncertainty about the observability (or not) of the postures. Postures 1.a, 1.c, 1.d and 1.e were observed most of the times, while posture 1.b was not observed most of the times. Only posture 1.a has a slight degree of uncertainty. The same degree of agreement and certainty for observability occurs with the lower arm postures 2.a and 2.b. Posture 2.c has a median agreement for the unobservable case, but there is a significant value in the unsure column.

In the case of the wrist, we have three postures with a high degree of agreement; postures 3.a and 3.b were observed (3.b. with a certain degree of uncertainty) and posture 3.e was not observed. Postures 3.c and 3.d are more uncertain, because there are high values in the unsure columns, specially for posture 3.c.

TABLE 2. Observability rate of postures by the expert committee.

Posture	Observable	Not obs.	Unsure
Upper Arm			
1a	0.741	0.130	0.129
1b	0.000	0.907	0.093
1c	0.870	0.111	0.019
1d	1.000	0.000	0.000
1e	0.926	0.000	0.074
Lower Arm			
2a	0.982	0.000	0.018
2b	0.982	0.000	0.018
2c	0.204	0.519	0.277
Wrist			
3a	1.000	0.000	0.000
3b	0.815	0.019	0.166
3c	0.111	0.444	0.445
3d	0.074	0.593	0.333
3e	0.000	1.000	0.000
Neck			
4a	0.944	0.000	0.056
4b	0.407	0.222	0.371
4c	0.019	0.759	0.222
4d	0.611	0.259	0.130
Trunk			
5a	0.074	0.444	0.482
5b	0.852	0.000	0.148
5c	0.537	0.389	0.074
5d	0.333	0.667	0.000

For the neck, there is a posture, 4.a, with a high degree of agreement that it was observed. Posture 4.c was not observed but with a certain level of uncertainty. Postures 4.b and 4.d were observed in some cases and not observed in others, but both have a degree of uncertainty, specially for posture 4.b.

For the trunk, we have a posture, 5.b, that was observed with a slight degree of uncertainty. Posture 5.a has a high value in the unsure column and thus a high degree of uncertainty. Postures 5.c and 5.d have values in the observed and not observed columns, but low values in the unsure column, that would mean that in some cases the postures were observed and in others they were not. In this case, that was expected, since for the designed task, only in some cases (for the boxes at the bottom levels of the bookcase) the subject would have to bend to such degrees.

To better understand from where uncertainty comes and why the experts were unsure about observing some postures, we analyze the color images shown in Fig. 2, which are part of the ones used by the expert committee to do its assessment on the RULA postures. From the images, we can observe that for a human evaluator it would be difficult to perceive small angles ($< 10^\circ$), and therefore it would be hard to distinguish among some RULA postures during a work cycle or if a limb reaches a certain position. That apply in the case of postures 3.c, 3.d, 4.b, 4.c and 5.a.

We computed two measures of inter-agreement among experts: Pearson correlation coefficient, and Krippendorff's alpha coefficient. Table 3 shows the Pearson correlation

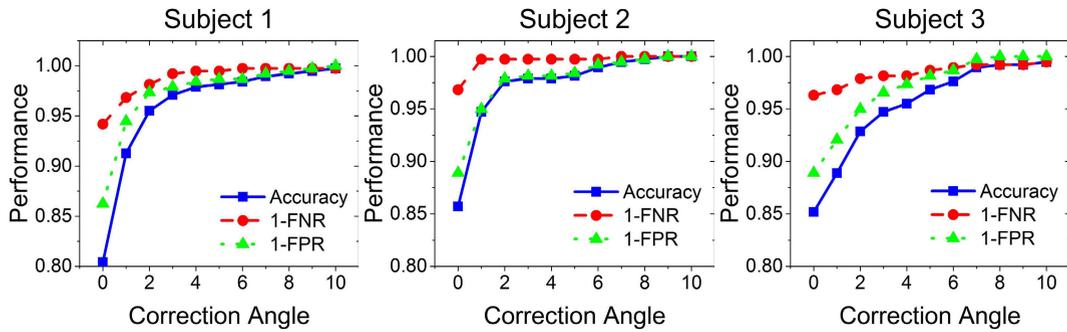


FIGURE 5. Performance of our system in terms of accuracy and complement of false positives and false negatives regarding the correction angle for subjects 1, 2 and 3.

TABLE 3. Correlations of the complete postures evaluation by the expert committee.

	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5
Expert 1	1.000	0.598	0.682	0.715	0.674
Expert 2	-	1.000	0.537	0.575	0.521
Expert 3	-	-	1.000	0.621	0.597
Expert 4	-	-	-	1.000	0.644
Expert 5	-	-	-	-	1.000

coefficients among expert evaluations, with an average correlation of 0.61. The Krippendorff’s alpha coefficient was 0.6. Both coefficients were computed considering every evaluation made by each expert. Both measures are similar and present a moderate level of agreement, showing the degree of uncertainty in the observations.

The advantage of our Kinect based system, is that Kinect obtain the location of each joint. With those locations, the computation of the formed angle was straightforward and with a higher precision than by simply estimating the angle with the naked eye.

V. RESULTS

We evaluated the accuracy of our system using the expert panel as the baseline. We defined the accuracy as the match between those postures observed by the experts and those detected by Kinect. A false positive is a posture detected by Kinect but not observed by the expert panel; and a false negative is a posture not detected by Kinect but observed by the experts. Similar to expert evaluation, Kinect detections were made per posture, per subject, per work cycle, for a total of 1,134 evaluations.

Fig. 5 illustrates the general accuracy of our system and the complement of the false positive rate (FPR) and false negative rate (FNR) for each subject. The first tick in the plot, indicates that our proposed strategy using Kinect produces a direct average accuracy of 0.83 compared to the expert panel. It can be observed that in all cases the FPR is higher than the FNR, which means that the most common error is a posture detected by Kinect but not observed by the expert panel. As mentioned in the previous section, some RULA postures are difficult to distinguish by the naked eye, generating uncertainty in the

final decision, and in such cases the expert panel produced a result by majority with a minimal difference of votes. On the other hand, Kinect measures the angles directly, and can detect postures with a higher precision than a human evaluator. For instance, posture 4.c requires an inclination of the neck of > 20° to be considered, but if the angle was of only 21° it could be hard to detect by the naked eye; in the case of Kinect, this value could be detected with better precision.

Considering the previous argument, to improve the accuracy of our method we considered a small relaxation of the margin in the upper and lower limit for detecting each posture with Kinect. This would make the Kinect detection flexible by simulating human visual perception. In Fig. 5, the rest of the ticks illustrates the effect of the relaxation angle (from ±1° to ±10°) in the performance. We observe that with a correction angle of ±1° (second tick) the average accuracy of the three subjects increased to around 0.93, and, as expected, we also see a general tendency to improve the performance when we increase the relaxation angle.

As described in subsection IV-A, the subjects’ activity was performed in 18 cycles. In cycles 1 to 9 some light cardboard boxes were placed over a bookcase, and during cycles 10 to 18 the boxes were put back in their original places. In a more detailed analysis, Fig. 6 depicts the system’s performance per work cycle using different relaxation angles. Each column of the figure illustrates the performance results for subjects 1, 2 and 3 respectively, and each row illustrates the results using a relaxation angle of 0°, ±1°, ±5° and ±10°, respectively. The accuracy without considering a relaxation angle varies from 0.7 to 0.9 for the three subjects along all cycles. Particularly, for subject 1, only work cycle 9 reached 0.9 of accuracy; for subject 2, work cycles 2 to 6 reached 0.9 of accuracy; and for subject 3, work cycles 1 to 6 obtained 0.9 of accuracy. With a relaxation angle of ±1°, for subject 1 the work cycles reached an accuracy between 0.85 and 0.95; for subject 2 between 0.9 and 0.95; and for subject 3 between 0.8 and 0.95. With a relaxation angle of ±5°, for subject 1 the work cycles reached an accuracy between 0.9 and 1 (in 12 cases); for subject 2 between 0.9 and 1 (in 12 cases); and for subject 3 between 0.85 and 1 (in 11 cases). Finally, with a relaxation angle of ±10°, almost

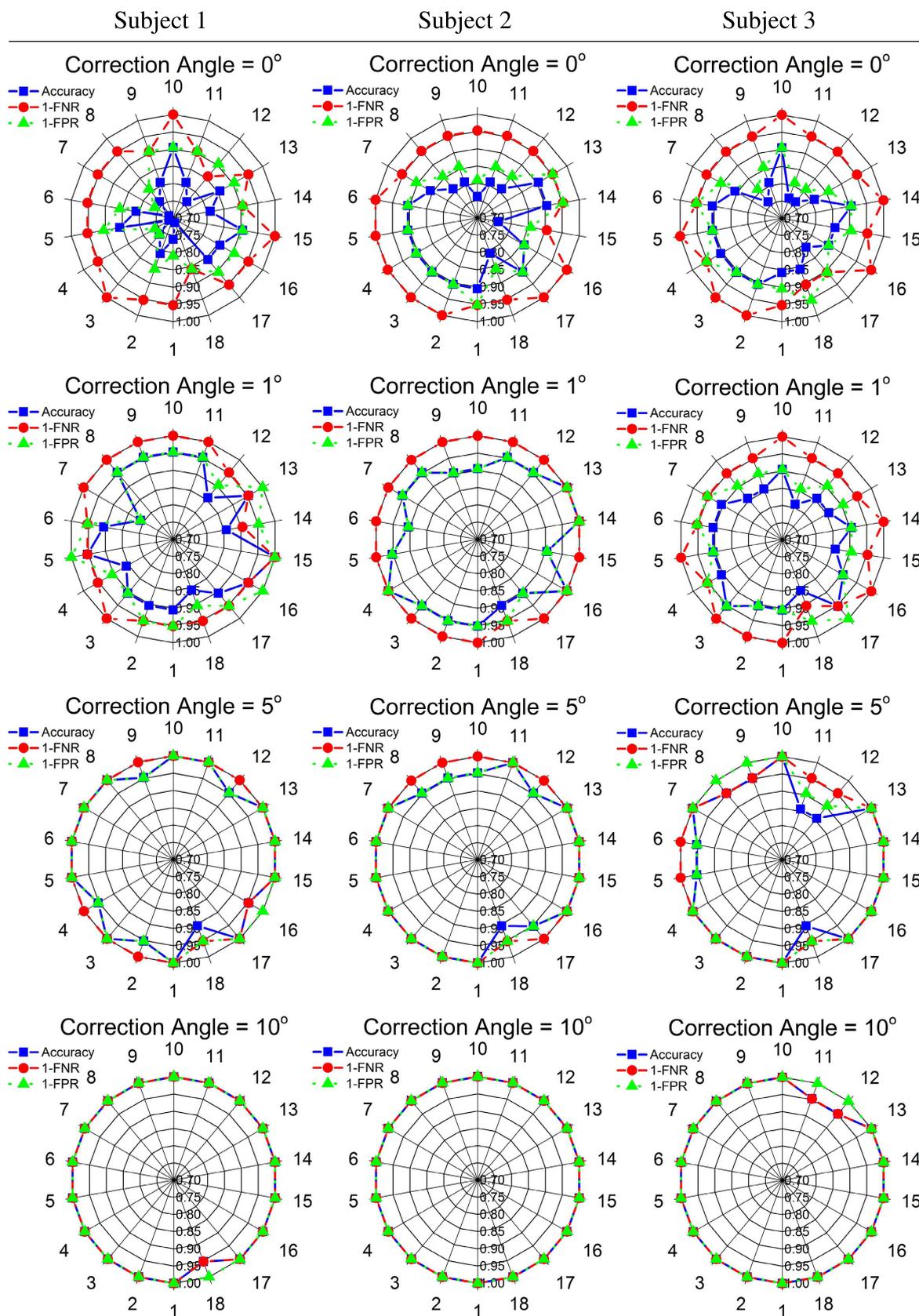


FIGURE 6. Performance of our system in terms of accuracy and complement of false positive and false negative rates regarding the correction angle and the work cycle for subjects 1, 2 and 3.

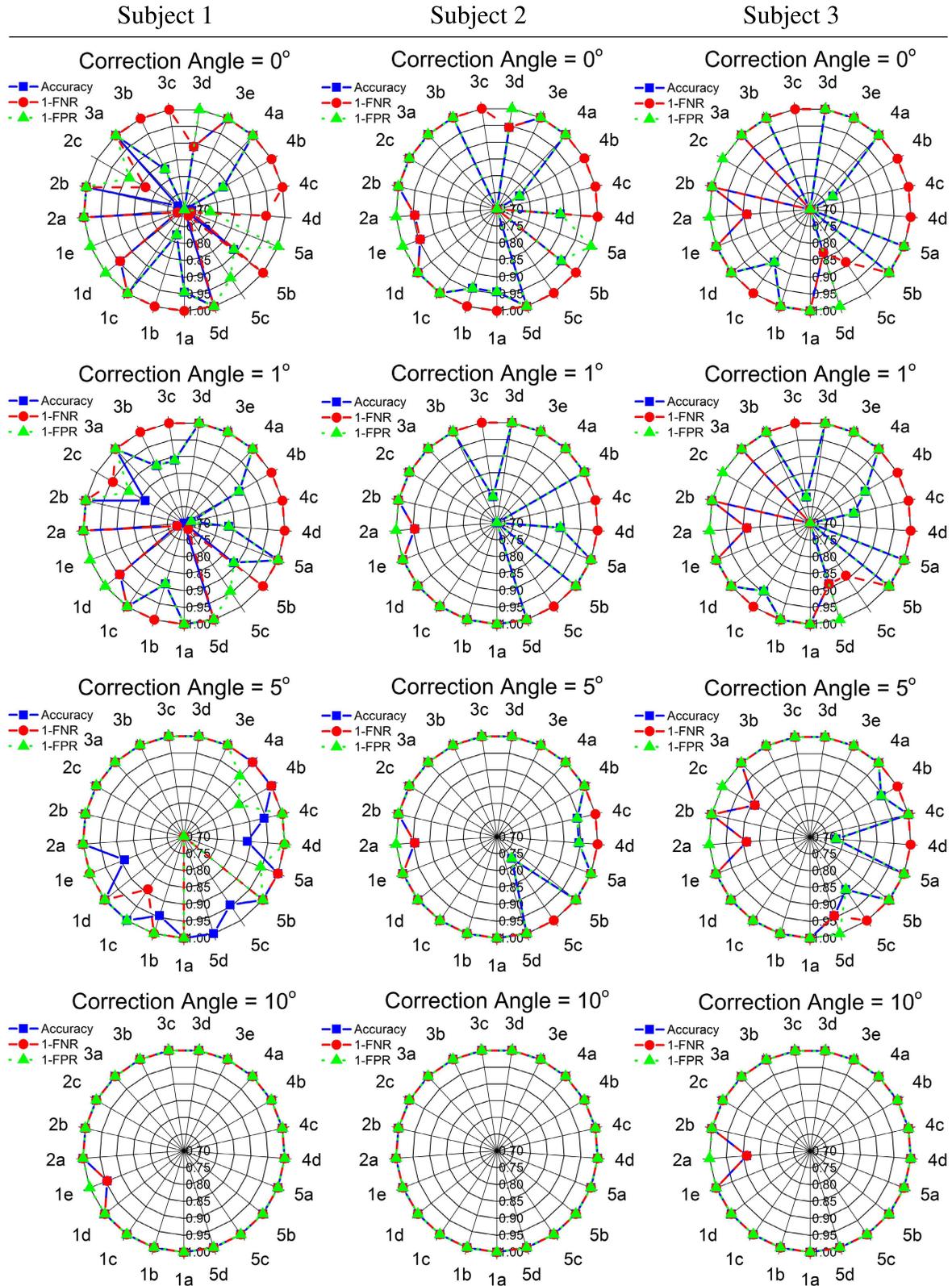


FIGURE 7. Performance of our system in terms of accuracy and complement of false positive and false negative rates regarding the correction angle and the RULA posture, as indicated in Fig. 1, for subjects 1, 2 and 3.

all work cycles of the three subjects reached their maximum performance. The exceptions were cycle 18 of subject 1 and

cycles 11 and 12 of subject 3. However, such cycles show false negative detections obtained from the expert panel;

which means that the posture was certainly detected using our system.

An analysis of the system's performance from a different perspective compares the average accuracy of detecting each posture along all cycles, considering the same relaxation angles as in Fig. 6. The system's performance regarding the observed postures per subject is summarized in the plots shown in Fig. 7. In the first row (without a relaxation angle), detection of postures with a high degree of uncertainty (see Table 2, 2.c (except subject 2), 3.b (only subject 1), 3.c, 3.d (except subject 3), 4.b, 4.c, 4.d, 5.a (except subject 3) and 5.b (except subject 3) show mid-values for each subject, which shows the difficulty to detect such postures by all the experts. In the second row, when using a relaxation angle of $\pm 1^\circ$, detection of postures 3.c, 3.d, 4.b, 5.a and 5.b improved notably. With a relaxation angle of $\pm 5^\circ$, practically all the postures improved their detection accuracy, except for posture 4.d in subject 3. With a relaxation angle of $\pm 10^\circ$ all the postures reached almost a perfect accuracy, except for posture 2.a in subject 3 and 1.e in subject 1. In those cases there are false negatives, which means the expert were confused about the exact position of the upper and lower arm for the particular subjects.

VI. CONCLUSION AND FUTURE WORK

In this study, we have presented a system to assess the use of Kinect in the automatic risk evaluation of postures in human work activities following the RULA specifications. We conducted experiments with a "pick and place" task, where three lateral standing subjects had to move nine empty cardboard boxes from the levels of a three-level bookcase to its top (during 9 work cycles), and then putting them back in their original places (with another 9 work cycles). We recorded the work cycles using Kinect at a rate of 10 frames per second, obtaining the joint positions and color images for the three subjects. Our system used the joint positions to detect the RULA postures for the upper arms, lower arms, wrists, neck and trunk. We compared the results of our system with the observational evaluation made by a committee of five experts that analyzed the color images captured by Kinect. The comparisons showed that Kinect is a suitable device for assessing the ergonomics of lateral standing subjects at work environments, with an average accuracy of 0.83 between the system and the expert committee without considering any relaxation in the margin of detection.

In our study we found that human visual perception is prone to more mistakes than the automatic system based on Kinect when deciding if a posture was observed or not. This is particularly true when the angle to determine the postures is within a small range; as is the case for postures of wrists and neck. Following this observation, we relaxed the margin used to consider a match between posture detections by the system and posture detections by the experts. When we relaxed the margin of detection by only $\pm 1^\circ$, we reached an accuracy of 0.93, and with a correction of $\pm 10^\circ$ we reached an almost perfect accuracy. The correction of $\pm 1^\circ$ is a small angle in

terms of human visual perception at the distance where we located the sensor.

Kinect sensor is thus a reliable device to consider for automatic detection of ergonomics when dealing with tasks similar to the one designed here, where the subject does not walk or move their feet constantly, and performs the task in a lateral standing posture. Similar tasks are common in assembly lines, workshops and warehouses.

Future studies should conduct experiments with our system in real-world environments, in order to detect possible failures and how to deal with them. A second type of study could explore the usability of Kinect when detecting ergonomics in more complex tasks, where other postures are performed, including the twist of wrists, neck and trunk, and the use of legs to reach or place certain objects. In order to reach this goal, we consider that the use of an array of sensors located at different places in the work environment could help by providing postural information from different angles. Finally, it would be interesting to study the effect of loads in the ergonomics, and idea for that would be to use a system that combines Kinect with some direct sensors.

REFERENCES

- [1] R. Pramitasari, S. Pitaksanurat, T. Phajan, and W. Laohasiriwong, "Association between ergonomic risk factors and work-related musculoskeletal disorders in beverage factory workers, Indonesia," in *Int. Seminar Workshop Public Health Action Building Healthy Community*, 2015, pp. 196–200.
- [2] J. A. Diego-Mas and J. Alcaide-Marzal, "Using Kinect sensor in observational methods for assessing postures at work," *Appl. Ergonom.*, vol. 45, no. 4, pp. 976–985, Jul. 2014. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0003687013002676>
- [3] S. Hignett and L. McAtamney, "Rapid entire body assessment (REBA)," *Appl. Ergonom.*, vol. 31, no. 2, pp. 201–205, 2000. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0003687099000393>
- [4] O. Karhu, P. Kansil, and I. Kuorinka, "Correcting working postures in industry: A practical method for analysis," *Appl. Ergonom.*, vol. 8, no. 4, pp. 199–201, 1977. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/0003687077901648>
- [5] L. McAtamney and E. N. Corlett, "RULA: A survey method for the investigation of work-related upper limb disorders," *Appl. Ergonom.*, vol. 24, no. 2, pp. 91–99, 1993. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/0003687093900808>
- [6] E.-P. Takala, I. Pehkonen, M. Forsman, G.-Å. Hansson, S. E. Mathiassen, W. P. Neumann, G. Sjøgaard, K. B. Veiersted, R. H. Westgaard, and J. Winkel, "Systematic evaluation of observational methods assessing biomechanical exposures at work," *Scandin. J. Work, Environ. Health*, vol. 36, no. 1, pp. 3–24, Jan. 2010.
- [7] A. Patrizi, E. Pennestrì, and P. P. Valentini, "Comparison between low-cost marker-less and high-end marker-based motion capture systems for the computer-aided assessment of working ergonomics," *Ergonomics*, vol. 59, no. 1, pp. 155–162, Jan. 2016, doi: [10.1080/00140139.2015.1057238](https://doi.org/10.1080/00140139.2015.1057238).
- [8] W. Umer, H. Li, G. P. Y. Szeto, and A. Y. L. Wong, "Identification of biomechanical risk factors for the development of lower-back disorders during manual rebar tying," *J. Construct. Eng. Manage.*, vol. 143, no. 1, Jan. 2017, Art. no. 04016080.
- [9] E. Valero, A. Sivanathan, F. Bosché, and M. Abdel-Wahab, "Musculoskeletal disorders in construction: A review and a novel system for activity tracking with body area network," *Appl. Ergonom.*, vol. 54, pp. 120–130, May 2016. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0003687015301204>
- [10] P. Mijović, E. Giagloglou, P. Todorović, I. Mačuzić, B. Jeremić, and I. Gligorijević, "A tool for neuroergonomic study of repetitive operational tasks," in *Proc. Eur. Conf. Cognit. Ergonom. (ECCE)*, 2014, p. 32.

- [11] Y. Z. Chong, S. J. Yong, and J. Yunus, "Development of cost-effective human motion analysis system for children with movement disorders," in *Proc. IEEE Conf. Biomed. Eng. Sci. (IECBES)*, Dec. 2014, pp. 607–611. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7047575>
- [12] G. Li and P. Buckle, "Current techniques for assessing physical exposure to work-related musculoskeletal risks, with emphasis on posture-based methods," *Ergonomics*, vol. 42, no. 5, pp. 674–695, May 1999. [Online]. Available: <http://www.tandfonline.com/doi/abs/10.1080/001401399185388>
- [13] G. C. David, "Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders," *Occupational Med.*, vol. 55, no. 3, pp. 190–199, May 2005. [Online]. Available: <http://academic.oup.com/ocmed/article/55/3/190/1420798/Ergonomic-metho%ds-for-assessing-exposure-to-risk>
- [14] L. H. Barrero, J. N. Katz, and J. T. Dennerlein, "Validity of self-reported mechanical demands for occupational epidemiologic research of musculoskeletal disorders," *Scand. J. Work, Environ. Health*, vol. 35, no. 4, p. 245, 2009.
- [15] P. Coenen, I. Kingma, C. R. L. Boot, G. S. Faber, X. Xu, P. M. Bongers, and J. H. van Dieën, "Estimation of low back moments from video analysis: A validation study," *J. Biomech.*, vol. 44, no. 13, pp. 2369–2375, Sep. 2011. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0021929011005008>
- [16] P. Coenen, I. Kingma, C. R. L. Boot, P. M. Bongers, and J. H. van Dieën, "Inter-rater reliability of a video-analysis method measuring low-back load in a field situation," *Appl. Ergonom.*, vol. 44, no. 5, pp. 828–834, Sep. 2013. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S000368701300029X>
- [17] A. M. Genaidy, R. J. Simmons, L. Guo, and J. A. Hidalgo, "Can visual perception be used to estimate body part angles?" *Ergonomics*, vol. 36, no. 4, pp. 323–329, Apr. 1993. [Online]. Available: <http://www.tandfonline.com/doi/abs/10.1080/00140139308967889>
- [18] S. Pinzke and L. Kopp, "Marker-less systems for tracking working postures—Results from two experiments," *Appl. Ergonom.*, vol. 32, no. 5, pp. 461–471, Oct. 2001. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0003687001000230>
- [19] P. Plantard, H. Shum, and F. Multon, "Ergonomics measurements using Kinect with a pose correction framework," *Digit. Hum. Model.*, Montreal, QC, Canada, Tech. Rep. hal-01332716, Jun. 2016, p. 8.
- [20] P. Plantard, H. P. H. Shum, and F. Multon, "Filtered pose graph for efficient Kinect pose reconstruction," *Multimedia Tools Appl.*, vol. 76, no. 3, pp. 4291–4312, Feb. 2017, doi: [10.1007/s11042-016-3546-4](https://doi.org/10.1007/s11042-016-3546-4).
- [21] A. Scano, M. Caimmi, M. Malosio, and L. M. Tosatti, "Using Kinect for upper-limb functional evaluation in home rehabilitation: A comparison with a 3D stereoscopic passive marker system," in *Proc. 5th IEEE RAS/EMBS Int. Conf. Biomed. Robot. Biomechtron.*, Aug. 2014, pp. 561–566.
- [22] T. Steinebach, E. H. Grosse, C. H. Glock, J. Wakula, and A. Lunin, "Accuracy evaluation of two markerless motion capture systems for measurement of upper extremities: Kinect v2 and Captiv," *Hum. Factors Ergonom. Manuf. Service Industries*, vol. 30, no. 4, pp. 291–302, Jul. 2020.
- [23] A. Humadi, M. Nazarahari, R. Ahmad, and H. Rouhani, "In-field instrumented ergonomic risk assessment: Inertial measurement units versus Kinect v2," *Int. J. Ind. Ergonom.*, vol. 84, Jul. 2021, Art. no. 103147.
- [24] V. M. Manghisi, A. E. Uva, M. Fiorentino, V. Bevilacqua, G. F. Trotta, and G. Monno, "Real time RULA assessment using Kinect v2 sensor," *Appl. Ergonom.*, vol. 65, pp. 481–491, Nov. 2017.
- [25] K. Woolford, "Defining accuracy in the use of Kinect v2 for exercise monitoring," in *Proc. 2nd Int. Workshop Movement Comput.*, Aug. 2015, pp. 112–119. [Online]. Available: <http://dl.acm.org/citation.cfm?doid=2790994.2791002>
- [26] T. Allahyari, A. S. Samani, and H.-R. Khalkhali, "Validity of the Microsoft Kinect for measurement of neck angle: Comparison with electrogoniometry," *Int. J. Occupational Saf. Ergonom.*, vol. 23, no. 4, pp. 524–532, 2016.
- [27] A. P. G. Castro, J. D. Pacheco, C. Lourenço, S. Queirós, A. H. J. Moreira, N. F. Rodrigues, and J. L. Vilaça, "Evaluation of spinal posture using Microsoft Kinect: A preliminary case-study with 98 volunteers," *Porto Biomed. J.*, vol. 2, no. 1, pp. 18–22, Jan. 2017. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S2444866416300939>
- [28] P. Poli, S. B. Healy, and D. P. Dee, "Pose estimation with a Kinect for ergonomic studies: Evaluation of the accuracy using a virtual mannequin," *Sensors*, vol. 15, no. 1, pp. 1785–1803, 2015. [Online]. Available: <http://www.mdpi.com/1424-8220/15/1/1785/>
- [29] G. B. Scott and N. R. Lambe, "Working practices in a perchery system, using the OVAKO working posture analysing system (OWAS)," *Appl. Ergonom.*, vol. 27, no. 4, pp. 281–284, Aug. 1996. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/0003687096000099>
- [30] Z. Zhang, "Microsoft Kinect sensor and its effect," *IEEE MultiMedia*, vol. 19, no. 2, pp. 4–10, Feb. 2012. [Online]. Available: <http://ieeexplore.ieee.org/document/6190806/>
- [31] T. Dutta, "Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace," *Appl. Ergonom.*, vol. 43, no. 4, pp. 645–649, Jul. 2012. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0003687011001529>
- [32] C. C. Martin, D. C. Burkert, K. R. Choi, N. B. Wieczorek, P. M. McGregor, R. A. Herrmann, and P. A. Beling, "A real-time ergonomic monitoring system using the Microsoft Kinect," in *Proc. Syst. Inf. Design Symp. (SIEDS)*, Apr. 2012, pp. 50–55. [Online]. Available: <http://ieeexplore.ieee.org/document/6215130/>
- [33] H. Haggag, M. Hossny, S. Nahavandi, and D. Creighton, "Real time ergonomic assessment for assembly operations using Kinect," in *Proc. UKSim 15th Int. Conf. Comput. Modelling Simulation*, Apr. 2013, pp. 495–500. [Online]. Available: <http://ieeexplore.ieee.org/document/6527468/>
- [34] P. Paliyawan, C. Nukoolkit, and P. Mongkolnam, "Office workers syndrome monitoring using Kinect," in *Proc. 20th Asia-Pacific Conf. Commun. (APCC)*, Oct. 2014, pp. 58–63.
- [35] J. Kim, N. H. Lee, B.-C. Bae, and J. D. Cho, "A feedback system for the prevention of forward head posture in sedentary work environments," in *Proc. ACM Conf. Companion Publication Designing Interact. Syst.*, Jun. 2016, pp. 161–164.
- [36] J. Krüger and T. D. Nguyen, "Automated vision-based live ergonomics analysis in assembly operations," *CIRP Ann.*, vol. 64, no. 1, pp. 9–12, 2015. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0007850615000542>
- [37] M. Devanne, H. Wannous, S. Berretti, P. Pala, M. Daoudi, and A. Del Bimbo, "3-D human action recognition by shape analysis of motion trajectories on Riemannian manifold," *IEEE Trans. Cybern.*, vol. 45, no. 7, pp. 1340–1352, Jul. 2015.
- [38] S. Giancola, A. Corti, F. Molteni, and R. Sala, "Motion capture: An evaluation of Kinect v2 body tracking for upper limb motion analysis," in *Wireless Mobile Communication and Healthcare*, P. Perego, G. Andreoni, and G. Rizzo, Eds. Cham, Switzerland: Springer, 2017, pp. 302–309.



ERNESTO ROCHA-IBARRA was born in Guanajuato, Mexico, in 1978. He received the bachelor's degree in industrial relations and the master's degree in organizational development from the University of Guanajuato, in 2010 and 2012, respectively, and the Ph.D. degree in economics sciences from the Autonomous University of Baja California, Mexico, in 2016. He is currently a Professor with the Arts and Enterprise Department, Engineering Division, University of Guanajuato. His research interests include ergonomics, productivity, and human capital. He was a recipient of the Academic Excellence Medal of the Organizational Development Program granted by the University of Guanajuato, in 2012.



MARVELLA-IZAMAR ORÓS-FLORES received the B.Sc. degree in computational systems and the M.Sc. degree in electrical engineering with a focus on posture recognition from the University of Guanajuato, Mexico, in 2015 and 2018, respectively.



DORA-LUZ ALMANZA-OJEDA (Member, IEEE)

received the B.Sc. degree in electronic engineering and the M.Eng. degree in electric from the University of Guanajuato, Salamanca, Mexico, in 2003 and 2005, respectively, and the Ph.D. degree from the University of Toulouse III-Paul Sabatier, Toulouse, France, in 2011. She is interested in embedded vision for controlling autonomous robots and real-time systems. She is currently a full-time Professor with the Electronics Engineering Department, University of Guanajuato.



GABRIEL-ARMANDO LUGO-BUSTILLO received

the B.Sc. degree in electronic engineering from the University Dr. Rafael Beloso Chacín, Maracaibo, Venezuela, in 2013, and the M.Sc. degree in electrical engineering from the University of Guanajuato, Salamanca, Mexico, in 2017. He supported the development of the algorithm and programming for the acquisition of data, in a project that incorporates cameras with infrared light technology. He is interested in multimedia and digital signal processing.



ANDRES ROSALES-CASTELLANOS received

the B.Sc. degree in electronics engineering from the National Experimental University of Tachira, San Cristobal, Venezuela, and the M.Sc. degree in electrical engineering from the University of Guanajuato, Salamanca, Mexico, in 2017. He is currently pursuing the Ph.D. degree in transportation engineering with the University of Alberta. He has contributed with several projects involving sensors for different applications, in which he maintains a strong interest at University of Alberta.



MARIO-ALBERTO IBARRA-MANZANO (Member, IEEE) received the B.Eng. degree in communication and electronic engineering and the M.Eng. degree in electric from the University of Guanajuato, Salamanca, Mexico, in 2003 and 2006, respectively, and the Ph.D. degree (Hons.) from the Institut National des Sciences Appliquées, Toulouse, France, in 2011. He is interested in digital design on FPGA for image processing applied on autonomous robots and real-time systems. He is currently a full-time Professor with the Electronics Engineering Department, Universidad de Guanajuato.



JUAN CARLOS GOMEZ received the M.Sc. degree in astrophysics and the Ph.D. degree in computer science from INAOE, Mexico, in 2002 and 2007, respectively. He was a Post-doctoral Researcher with KU Leuven, Belgium, from 2008 to 2009 and from 2011 to 2015, and has been with ITESM, Mexico, since 2010. His research interests include machine learning, data mining, evolutionary computation, and information retrieval, where he has published several peer-reviewed articles. He is currently a full-time Professor with the Departamento de Ingeniería Electrónica, Universidad de Guanajuato, Mexico, and a member of the Sistema Nacional de Investigadores (SNI level 1), Mexico.

...