Development of Duplex Eye Contact Framework for Human-Robot Inter Communication

SHAYLA SHARMIN[1], **(Graduate Student Member, IEEE),**
MOHAMMED MOSHIUL HOQUE[1], **(Senior Member, IEEE),**
S. M. RIAZUL ISLAM[2], **(Member, IEEE),**
MD. FAZLUL KADER[3], **(Senior Member, IEEE),**
AND IQBAL H. SARKER[1], **(Member, IEEE)**

[1]Department of Computer Science and Engineering, Chittagong University of Engineering and Technology, Chattogram 4349, Bangladesh
[2]Department of Computer Science and Engineering, Sejong University, Seoul 05006, South Korea
[3]Department of Electrical and Electronic Engineering, University of Chittagong, Chittagong 4331, Bangladesh

Corresponding authors: Mohammed Moshiul Hoque (moshiul_240@cuet.ac.bd) and S. M. Riazul Islam (riaz@sejong.ac.kr)

ABSTRACT Establishing eye contact is the fundamental key to begin any interaction between human-human and robot-human. Two approaches are available to develop an eye contact mechanism for robot-human interaction, such as simplex and duplex. The two most critical tasks: gaze crossing and gaze awareness, are prerequisite to implementing an active eye contact mechanism in any approach. However, most past robot-human interaction studies implemented a gaze crossing function to develop eye contact in the simplex mode where a robot holds for the human to initiate the communication. However, implementing gaze crossing alone is inadequate to create an active eye contact episode; the gaze awareness function also essential to achieve. This paper aims to develop a mechanism of duplex eye contact for robot-human inter-communication satisfying both functions. This work proposes a conceptual model of a duplex eye contact mechanism considering two cases: human initiative (where the human starts communication with the robot) and robot initiative (where the robot starts the communication with the participant) to achieve a duplex eye contact mechanism. Moreover, a simple robotic system is developed consisting of four software constituents: face detection module, gaze detection and tracking module, gaze awareness module, and robot response and control module to implement the conceptual model of duplex eye contact. Several preliminary experiments are performed to extract necessary cues for designing the duplex eye contact mechanism's behavioural protocol and present their results to show the usefulness of extracted cues. Moreover, the robotic framework results in a scenario (*e.g., reading the book*) with the proposed duplex eye contact mechanism are presented. The results show that the proposed scheme achieved 92% and 86% accuracy for human initiative case and robot initiative case, respectively in making eye contact.

INDEX TERMS Human-robot interaction, duplex eye contact, social robots, gaze awareness, face detection, evaluation.

I. INTRODUCTION

Setting up eye contact is one of the notable primitive abilities to institute any interplay in robot-human or human-human communications. Eye contact offers a notable function in regulating face-to-face interaction and in initializing any conversation [1]. It is the foundation of developmental

The associate editor coordinating the review of this manuscript and approving it for publication was Tao Liu.

harbinger to more arduous gaze functions such as joint attention and language understanding [2], [3]. Moreover, it consequences in superior information recall of the conversation [4], and participants have to establish eye contact to start any social conversation and sustained [5]. Psychological surveys illustrated that eye contact enrich the feeling of interest, affection, trust, engagement and solicitation in one another [6], [7]. The central function of gaze in human-human interaction is to regulate the flow of conversation [8]. Setting up duplex eye

contact is one of the essential functionalities to be schemed and invoked in social agents including robots. Gaze producing and gaze awareness are the two critical functions to forge a gainful eye contact event for human-human or human-robot communications. The eye contact established when both parties perceive their gaze by looking at each other's face or eyes simultaneously [8]. Moreover, it often assumed that the gaze producing function (i.e., looking at each other) is adequate to set up eye contact. However, psychological surveys revealed that the gaze producing function alone is not adequate; gaze awareness should also invoke to establish an actual eye contact behavior [9], [10]. Gaze awareness or gaze-responsive behaviour is also a vital component in duplex eye contact scheme for creating the feeling of eye contact [11]. Both interacting partners should produce appropriate gaze responsive behaviours to interpret each other response. Nonverbal behaviours such as smiling or eye blinks is considered persuasive cues when both parties in face-to-face [12], [13].

Our primary goal is to develop a duplex eye contact mechanism for human-robot intercommunication (HRI), satisfying gaze crossing and gaze awareness tasks using non-verbal behaviours. A robot that brings eye contact with a human is a significant capability to be introduced in social robots. In recent years, many social robots are trying to use in therapy, service, sales agent, or teaching [14], [15]. However, before starting any interaction, the robot must set up eye contact with the intended partner. In our work, we do not focus on a particular application of social robots; instead, we emphasize a fundamental social capability (i.e., establishing eye contact) necessary to initiate any interaction or conversation. Although there may be numerous circumstances, this paper considers a generic situation where the agents (i.e., robot and the human) are not facing each other at the beginning of interaction and the human is engaged in a task that does not occupy ample attention (i.e., reading a book). Under these constraints, we consider how the robot can behave to set up eye contact according to humans and robots' relative position. Visual stimuli by the robot's non-verbal behaviours cannot influence a human's attention where he/she cannot perceive the robot due to his/her posture. We do not contemplate such instances in this research. Many past studies considered the simplex eye contact approach in the passive mode where the human faces the robot initially, and the robot wait for the human to begin the interaction [16]–[21]. However, this simplex behaviour may not work in all circumstances, and there are many circumstances in reality that demand duplex eye contact. A previous system implemented the duplex eye contact mechanism in the HRI framework, which utilized a flat-screen display as the robot's head to manifest computer graphics-based smile appearance as gaze awareness. Nevertheless, a flat-screen is impractical to realize a robot's face. Moreover, most previous systems used the gaze crossing function alone with the costly and complex robotic platform to implement eye contact mechanism. The proposed work presents a conceptual model for duplex eye contact

where both the robot and human can initiate the eye contact event, concerning both gaze crossing and gaze awareness tasks.

By the adoption of human-human interaction, HRI scheme should also make sure of eye contact behaviour with the implementation of gaze crossing and gaze awareness functions [22]. Providing social robots with lifelike sociable competencies that stimulate the impression of a much smarter and instinctual interaction, ensuring a high degree of contentment to communicating humans [23]. Thus, the significant concerns in our research are: (i) How to design an HRI framework that can perform the duplex eye contact task? (ii) How to design discreet cues for a robot to execute the gaze crossing when interacting partners are not facing each other due to their spatial positioning? (iii) How does the robot respond when the human wants to interact with it? (iv) How the robot vouches if the human is looking at it against its actions? (v) How can the robot display gaze awareness when it has ensured gaze crossing with the human? To address the above mentioned concerns this work proposes an HRI framework of duplex eye contact by considering two situations; (i) Robot Initiative Case (RIC) and (ii) Human Initiative Case (HIC). The proposed framework performs several activities or actions in RIC and HIC to make eye contact between humans and robots. Our work's significant contributions are listed below:

- Develop a conceptual model of duplex eye contact considering two cases: human initiative and robot initiative.
- Develop a robotic platform having 04 software modules (such as FDM, GDTM, GAwM and RRCM) to verify the conceptual model's effectiveness.
- Investigate the actions (i.e., cues) of the human and robot to design suitable cues for gaze crossing and gaze awareness functions.
- Propose a robotic system's behavioural protocol to perform eye contact. Perform several preliminary experiments to assess the proposed methods' functionality when the interacting partners are not facing each other due to their spatial arrangement.
- Evaluate the proposed framework's performance in a particular scenario where the intended participant involved in a low attention absorption task.

The remaining of the paper is arranged following: sub section I(A) represents related work. A brief description of the conceptual design of the duplex eye contact mechanism is outlined in Section 2. Section 3 provides the suggested framework with a detailed explanation. Section 4 describes the overall robotic platform with the description of hardware and software modules. Sections 5, 6, and 7 discussed the preliminary studies. Section 8 states the evaluation experiment of the proposed framework with the analysis of findings. In section 9 there is an analysis of all the experiments including the preliminary studies have been discussed. Finally, the paper concludes with a summary in Section 10.

A. RELATED WORK

Performing gaze crossing and ensuring gaze awareness are two essential prerequisites to establish active eye contact in any interaction. In human-human conversation surveys, few studies have conducted regarding how a human cross with the others' gaze to begun a conversation besides the rudimentary facts that the people halt a specific extent [24], initiate the conversation with a reception [25], and organize oneself in a spatial arrangement [26]. Few recent studies focused on human's response to the robot's eye contact behaviour in interactive jobs. These studies accustomed to perform the robot's gaze functions are usually either not founded on people gaze pattern or not responsive to the human mate's behaviours [27]–[34]. Gaze behaviour characterized as a necessary means for interaction and coordination in human conversations. The previous study on human conversation has investigated how people involve in gaze coordination [35], [36]. This exploration is generally biased, looking at each participant's gaze in segregation, and did not grab the knotty coordinate patterns in which mates' gaze behaviours interplay. Based on the past HRI survey, the eye contact mechanism can be categorized in two ways: simplex and duplex [8].

In the simplex method, the intended partner initiates the eye contact event. Few robotic systems have developed where the robot hold on for the human to begin the eye contact process. Various past HRI systems have examined welcoming behaviour to start eye contact event at a public distant [37]–[41]. These systems developed to emit certain welcoming words for beginning the communication with the participant. A small number of frameworks endeavoured to encourage people's engagement by using vocal signals [16], [17] and recognizing urging behaviour [18]. Majority of these frameworks does not examine how robots should act to the interacting partner for establishing eye contact. Few robotic systems have prepared with the ability to stimulate participants for setting up the eye contact event using nonverbal actions such as physical position and gaze [19], approaching path [42], standing footing [21], and tracking behaviours [20]. These systems supposed that the interacting participant faces the robot and aims to interact with it; nevertheless, in actuality, this surmise may not possess consistently. Robots may hold on for a participant to start a conversation, and utilizing speech assuredly captures other participant's attention, comprising the intended participant. Though such a passive approach can function in a few circumstances, numerous situations demand a robot to exploit a better active technique [43]–[46].

Indeed, a robot that approached participant and starts a conversation proactively by setting up eye contact should realise to simulate more life-like than a robot holds on for its interacting partner [47]. Few robotic frameworks tooled with the capacity to start conversation proactively with the participant. Satake et al. [48], [49] developed a framework that facilitates a robot to approach people proactively by estimating his/her routes in a social space. Their system uses voice signal to initiate the conversation, but the robot's attempt failed when the target person is busy with another person. Mitsunaga et al. [50] utilised Robovie-IV framework, which rambles in an office space and searches the interacting partner. Nevertheless, the engagement phase is passive since the robot needs to hold on for the identified participant to get closer. Performing approaching behaviour by a robot is not a simple task since it behaviour requires to be affirmed non-verbally preemptively; unless the approached participant might not realise that the robot is addressing her/him or would be overwhelmed by the robot's ill-mannered break. People performs this nicely with the eye gaze [25], [51]. However, these schemes failed to detect the participant's gaze and his/her body dictate, which are the essential parameters to determine whether the participant has acknowledged to the robot's beacon or not. Numerous robotic schemes have designed to obtain eye contact utilising the gaze crossing employment [52]–[55]. These schemes assumed to institute eye contact with the people by shifting cameras toward him/her faces. A stuffed-toy robot used by Yonezawa et al. [56] which can stimulate a favourable impression by efficient aid of eye contact effects with shared attention. An eye-gaze process for interactions has been developed in humanoid robots to indicate allies on their roles in interaction [55]. Most of these researches directed on the gaze crossing element solely to design eye contact skill of social robots and creating gaze-awareness mechanism was lacking absolutely.

Few simplex eye contact frameworks employed both gaze crossing and gaze-awareness. Hoque et al. [12], [57] proposed an eye contact system in which the participant's attention captured by turning a robot's head for performing gaze crossing and use eye blinks for creating gaze awareness. A design explained the opening a conversation process with the target participant in diverse viewing conditions where a robot was capable of meeting the gaze with him/her by attracting attention through head-turning, head shaking, and greeting words [58], [59]. This method employed CG images to perform eye blinks as the gaze awareness capacity. Huang and Thomaz [60], [61] adopted the Simon robot to assemble the awareness capacity. The Simon flash its ear when it hears an announcement. Nevertheless, Simon did not applied ear blinks as the gaze awareness function instead apply to produce communication consciousness. Yoshikawa et al. [10] employed a robotic framework to construct the active gaze behaviours. This framework revealed that the active gaze mechanism strengthens the feeling of oneself being gazed at the robot. Nevertheless it is unexplored how the robot acts gaze-awareness behaviour to the reacted participant. Mehlmann et al. [62] used Nao robot which can set up a mutual gaze and respond to verbal cues with the gaze. Phyo et al. [63] highlighted on nonverbal action of people where the robot examines whether the individual is beckoning the robot by shaking hand. All of these previous HRI studies

concentrated on developing simplex eye contact schemes where the robot should begin the eye contact event. In actuality, any agent (i.e., the robot or the human) may initiate eye contact process in collaborative work. Thus, HRI framework should design in such a way so that it can capable to work not only in simplex but also in the duplex way.

Very few studies have focused on designing the duplex or bi-directional eye contact in employing both gaze crossing and gaze awareness components. Andrist *et al.* [22] presented a mechanism of duplex gaze in a virtual agent. A virtual character coordinates the production and the exposure of gaze cues in this system with the participants. Miyauchi *et al.* [8] proposed a duplex eye contact scheme. In their system, the robot produces a smiling expression as the gaze-awareness following meeting the gaze. This scheme employed a flat-screen digital display as the robot's head and presented 3D computer graphic (CG) images to imagine smile appearance. A flat-screen is as unusual and unrealistic as a face. A fundamental feature is the face appearance's geometry; the half-spherical appearance lets the public look at the face into a 180° extended sphere. Imitating the eyes' geometry, which persists the utmost significant component in the face, aids understand the robot's glance, enhancing interplay [64].

Many works such as [8], [19], [42], [65] consider detecting the human's frontal face as she/he is approaching the robot to initiate the interaction. However, it may happen that although the camera gets a frontal face, the gaze direction is not towards the robot that leads to failing to establish gaze crossing. We introduced a gaze detection and tracking mechanism after detecting a human's face to help a robotic system set up gaze crossing more effectively. Most previous studies has been suggested the simplex eye contact approach either in HIC ([19], [20], [42], [43]) or RIC ([48]–[50]).The proposed work presents a duplex eye contact mechanism considering both HIC and RIC. Many past works [12], [65] does not consider any gaze awareness function concerning the human. Our approach considered detecting smile expression as human's gaze awareness. This approach will help the robot understand the human's willingness as we know a smiling face shows a positive response for further interaction [66]. As a gaze awareness, the proposed framework select head nodding with eye blinks for the robot, which is more human-like than blinking ear [60], [61] blinking eyes [65] or projecting smile on a flat-screen [8]. In addition to these, all the robotic frameworks used in past HRI investigations were systematically complicated and costly to develop, assemble, and manage. This work proposed a duplex eye contact framework taking into account past systems' weaknesses. Moreover, to verify the proposed framework effectiveness, a simple robotic platform is constructed that is cheaper, easy to set up and maintain than the existing systems.

II. CONCEPTUAL DESIGN OF DUPLEX EYE CONTACT
The central purpose of gaze in people interaction is to regulate the meta-communication. Numerous robotic systems such as

Robita [35], Robovie [67], and Cog [52] have employing gaze for meta-communication. To implement active eye contact in the human-initiative case (HIC) or the robot initiative case (RIC), both parties (i.e., the human (H) and the robot (R)) should notice that they are gazing at each other and intent to initiate interaction, which ensures gaze awareness. It resembles that people can execute eye contact if they gazing at each other's face (i.e., gaze crossing) [9]. Gaze crossing plays a crucial role in any conversation having a significant influence on social behaviour. Gaze crossing used as a synchronizing signal, as people look at each other while talking or listening, and this phenomenon is also given feedback to the participants on any particular points [8], [12].

Gaze awareness action is obliged to establish firm eye contact as this cue helps the participants to understand the attention response of each other [8], [65]. To express awareness, people use verbal or non-verbal actions. Nonverbal action is considered a week action and is powerful when both parties are face to face. Eye blink, smiling, nodding head are widespread movements of humans to show awareness when the distance is small along with the condition face to face [12]. Verbal action considered as a firm stroke such as by calling name or using some reference terms and can be applied in both long and short distance communication. However, in that situation, if there are more than one person, confusion and annoyance may be raised [68]. Figure 1 shows the abstract view of eye contact process with essential constituents: gaze crossing and gaze awareness. Figure 1 (a) indicates that the gaze crossing establishes when the human (H) and the robot (R) are staring at each other face or eyes. Fig. 1 (b) demonstrates that both R and H notices each other's gaze after meeting the gaze and display gaze awareness with a smiling expression.
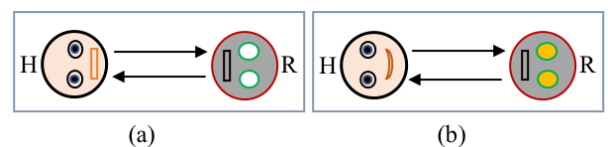

FIGURE 1. Abstract process of eye contact: (a) Gaze crossing, and (b) Gaze awareness.

People typically turn their head or face to whom they wish to interact since the head turning is regarded as the utmost requisite signal to attract the attention of other [55]. If the intended participant is unattended, he/she repeatedly attempts with the identical signal or with more powerful one (e.g., shaking hand, waving the head, bodily motions, or uttering speech). The social robots also adopt the identical protocol as humans in an actual HRI situation. Where a talker or hearer is staring is conceivably a dominant signal about attentiveness and intention in social interaction and offline communication [69]. Therefore, if the recipient perceived attraction by the robot's action, he/she will shift approaching it, which will obtain confronting arrangement (i.e., crossing gaze). A number of psychological investigations reveal that
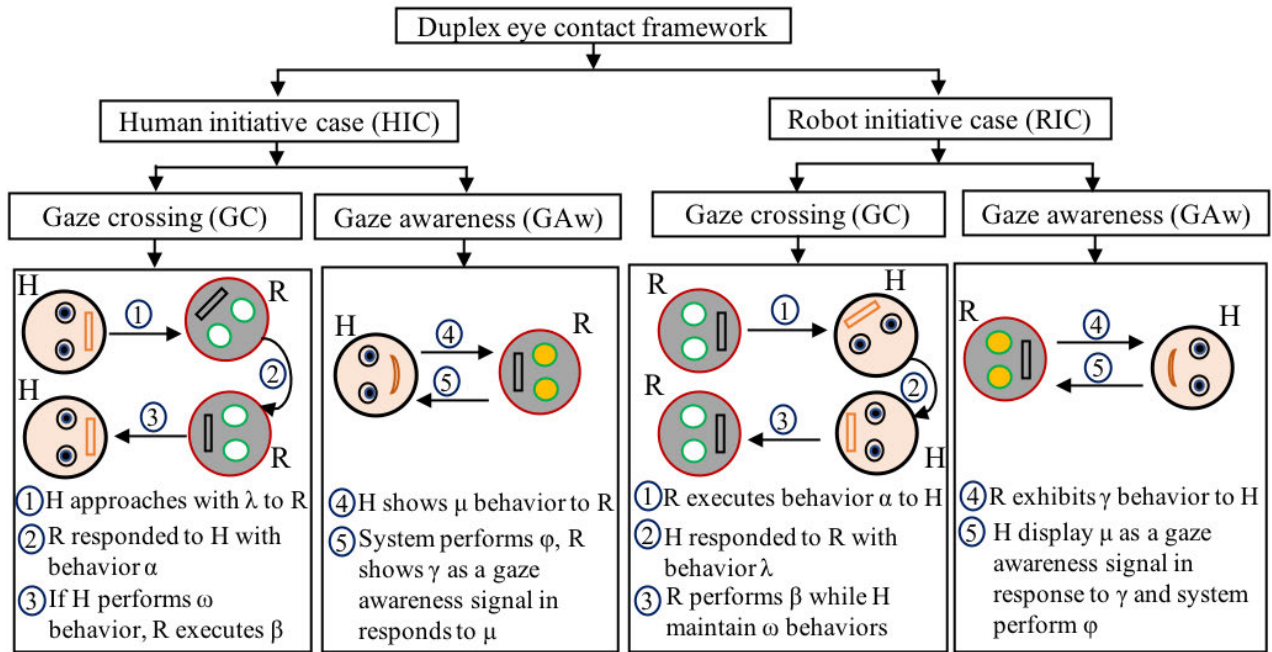
FIGURE 2. Conceptual process of duplex eye contact framework.

gaze crossing effort merely is not be adequate to build an active eye contact [9]. Yoshikawa *et al.* [10] also noticed that merely gazing is not adequate all the times for the robot to show people feels that they are being looked at it. Hence, each interaction agent should understand each other respond and perform suitable gaze awareness to each other after meeting the gaze. Exhibition of gaze awareness is an essential function for the recipient agent to create the feeling of attentional response.

Based on the preceding discourse, we can hypothesize that the duplex eye contact framework should consider HIC and RIC. Both cases perform two fundamental active eye contact behaviours: gaze crossing (GC) and gaze awareness (GAw) consecutively. Figure 2 shows the conceptual process of the duplex eye contact mechanism. Performing duplex eye contact, both agents (R and H) must exhibit explicit behaviours and respond competently to understand each other in each case. That means R and H play a set of behaviours $R = \{\alpha, \beta, \gamma, \phi\}$ and $H = \{\lambda, \omega, \mu\}$. These behaviours used to exhibit GC and GAw functions in both RIC and HIC. In the proposed RIC scheme, a set of the behavioural parameter of R such as $\alpha = \{$head turn, *and/or* head shake$\}$ is used to attract the attention of H. If H is looking at R by displaying $\lambda = \{$head and gaze turn toward R$\}$, it is assumed that he/she noticed R's communicative intention. If H maintains $\omega = \{$keep looking toward R$\}$, it performs a set of functions $\beta = \{$frontal face detection, gaze detection & gaze tracking$\}$. Thus, GC is established successfully. The GC constituent initializes the GAw by exhibiting the $\gamma = \{$head nod with eye blinks$\}$ behaviour by R. The system expects that H shows a responsive behaviour to R with $\mu = \{$smiling$\}$. The gaze awareness function completes by detecting the smile expression (ϕ), ensuring eye contact.

Again, in the proposed HIC scheme, it is assuming that H approaches R initially by looking at it to initiate an interaction. Thus, H may display behaviour $\lambda = \{$head or gaze turn toward R$\}$ to convey his/her communicative intention. R executes behaviour $\alpha = \{$head turn$\}$ in response to H. If H maintains $\omega = \{$keep looking toward R$\}$, the R exhibits set of behaviours $\beta = \{$frontal face detection, gaze detection, & gaze tracking$\}$ which ensure to set up GC. After successful completion of GC, H shows a gaze awareness behaviour $\mu = \{$smiling$\}$. R detects H's smile expression (ϕ) and exhibits gaze responsive behaviour to H by $\gamma = \{$head nod with eye blinks$\}$. Thus, the gaze awareness function is performed, which ensures eye contact.

III. PROPOSED APPROACH OF DUPLEX EYE CONTACT
Our work assumed that in any cases, H and R are not facing each other initially, and the H is occupying his/her current job. Therefore, in the case of RIC, R should turn its head as the primary signal to attracts its human partner's attention toward R. On the other hand, in HIC, R should direct its head to H to set up gaze crossing when H is trying to start communication. A state diagram in Fig 3 describes the duplex eye contact method. The state variable S_T defines either the human initiative case (HIC) or the robot initiative case (RIC) which defined as $S_T \in \{HIC, RIC\}$. The signals obtained in the proposed method are listed in Table 1.

A. RIC IN EYE CONTACT
For the case of RIC, the robot exhibits head movements depending on the S_T as $M_S \in \{HT, HS, HN, 0\}$, where $HT \in \{0, ht\}$ with *ht* indicates pre-defined head-turning action within 0° to 35° where $ht \in \{0, \alpha_1\}$, $HS \in \{0, hs\}$ denotes the pre-defined head-shaking action where
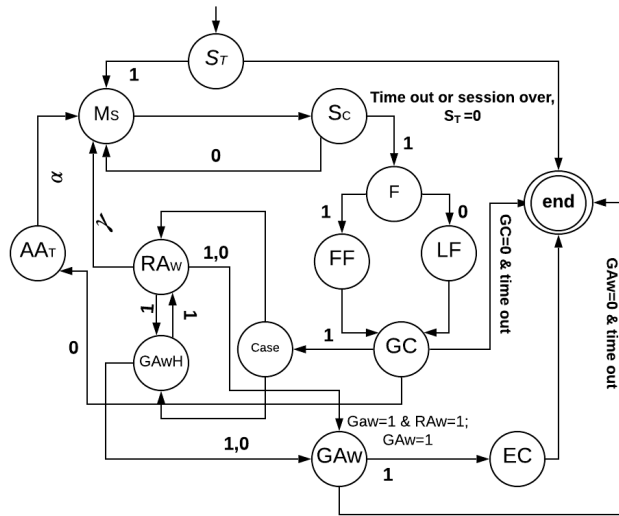
FIGURE 3. Proposed approach of duplex eye contact.

TABLE 1. Signals used in the proposed method.

Attributes	HIC (human initiative case)	RIC (robot initiative case)
$S_C \in F : FF/LF$ (Scene (face): frontal/lateral face)	FF=1; detects frontal face	FF/LF =1; detects face
GC (gaze cross)	GC=1; confirms gaze cross	GC=0 ; attracts human; GC=1 confirm gaze cross
GA_WH (gaze awareness H)	$GA_W H$=1; confirms smile	$GA_W H$=1; confirms smile
M_S (motion state)	HT: turns head to fix gaze; HN: nods head; -	HT: turns head to fix gaze; HS: shakes head; HN: nods head
RA_W (robot awareness)	HN	HN
AA_T (attention attraction)	-	HS
GA_w (gaze awareness)	$GA_w H$==1 then RA_w=1; confirm gaze awareness	RA_w=1 then $GA_w H$=1; confirm gaze awareness
EC (Eye contact)	GC=1, GA_W=1; confirm eye contact	GC=1, GA_W=1; confirm eye contact

$hs \in \{0, \alpha\}$ is the $\pm 20°$ back and forth from the initial position, and 0 indicates no movement. $HN \in \{0, hn\}$ is the head nodding action. 0 indicates no movement and $hn \in \{0, \gamma\}$ denotes the head-nodding within $0°$ to $20°$. The $S_T = 1$ when the system gets on its predefined position by turning head ($ht = \alpha_1$); thus, it stops moving ($M_S = 0$). If the system gets a human face ($S_C = 1$) then the next steps will be continued from this predefined position of the robot; hence, the robot will not turn to its initial position. R tries to attract H by shaking its head at the initial point, where ($\alpha = 1$) is to attract the participant and nods its head once to show gaze awareness ($\gamma = 1$) to human.

R usually turns it's head initialy toward the interacting partner and institutes shaking its head once (if necessary) to capture his/her attention at the M_S state. The human detection state, $S_C \in \{F, 0\}$ analyzes the face position, $F \in FF, LF$ which analysis frontal face (FF) and lateral face (LF). If there is any face, the robot will check for the next steps to perform the whole operation. If $LF = 1$ the system will consider there is no gaze cross ($GC = 0$), thus it will jump to AA_T phase. On the other hand, having been detected the frontal face, the R checks the GC state, which refers to whether two parties cross their gaze or not. $GC = 1$ if two participants cross their gaze; otherwise, the robotic head tries to attract him/her. To recapitulate, if ($GC == 0$), the R enters AA_T attention attraction

state to attract the robot by shaking ($\alpha = 1$) operation. After the shaking operation, the robot will again check the GC states. The attention attraction and robot awareness showing are performed by generating motion action through M_S state. Finally, when ($GC == 1$) R shows awareness in RA_w where R nods ($\gamma = 1$) it's head at M_S state and blink its eye at the same time showing by a projector to display gaze awareness. In the meantime, the system checks for gaze awareness at GA_wH state, considering human is not reluctant for the rest of the conversation. $GA_wH = 1$ indicates the human participant shows gaze awareness. When both R and H shows gaze awareness activities ($RA_w == 1 \&\& GA_wH == 1$; then $GA_W = 1$) the system ensure the gaze awareness; that complete the full trial of robot initiative case which means EC established.

Besides, the robot holds four seconds after commencing each effort for making eye contact. It is shown that quietness of more than four seconds turn into embarrassment since they entail a rift in the thread of intercommunication [70]. Hence, the attempt will be considered as failed attempts after the time is up.

B. HIC IN EYE CONTACT

To embark on, in HIC, prior to the showing awareness, H and R is facing each other. Noteworthy, only by showing responsive behaviour within 2 seconds after crossing gaze R considers the situation as a human initiative case; otherwise, it is considered as negative feedback in human-human interaction [70]. First, the robot is in its regular motion from the primary pose to predefined position searching for the participant. This motion state or M_S contains four parts of the motion of the robot. It is already mentioned that $S_C = 1$ means it will continue the next steps from this predefined position; thus the tilt motor will stop, $M_S = \alpha$ stands for shaking the head, and $M_S = \gamma$ represents nodding head. Therefore, in the beginning, R turns its head and stops at the predefined position and will get a human ($S_C = 1$). It will remain at this place for further operations because H and R are in face to face position, thus $F = FF$ in the human detection state ($S_C \in \{F, 0\}$; $F \in FF, LF$. The value of the state, S_C is 0 if no face is detected, in this case, $ht = 0$, which means the robot will turn to its initial position again. However, continuing the process, R checks for gaze crossing after detecting a face. The R checks the $GC \in \{0, 1\}$ state which refers to whether two parties cross their gaze or not. Here, $GC = 1$ means to the two parties crosses their gaze. At this point, R ensures whether H smiles back to it within a certain time in the $GA_wH \in \{0, 1\}$ state. If $GA_wH = 1$, RA_w shows R's responsive behavior by nodding its head once ($\gamma \in \{0, 1\}$) and blinking its eyes projected by a pocket projector, here, $\gamma = 1$ stands for nodding operation where 0 represents no operation. These last two actions define that both parties understand each other actions, which ensures gaze awareness (GA_W) which confirms eye contact (EC) has been established and thus human initiative case has been established.
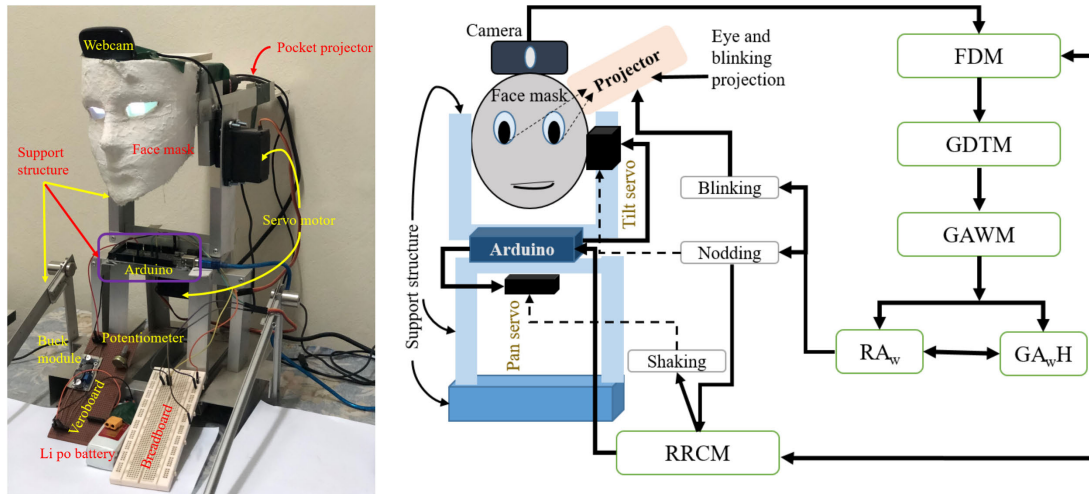
FIGURE 4. Robotic platform (a) prototype robotic head with hardware compositions, and (b) an overview of the head with key software.

IV. ROBOTIC PLATFORM

We designed a robotic head (as a platform) to conduct the human-robot inter-communication experiments and implement our conceptual design of duplex eye contact on this platform. Figure 4 illustrates the schematic layout of the developed robotic system. Following subsections explain the development process of the robotic platform with hardware configuration and software modules.

A. HARDWARE COMPOSITION

The robotic head consists of an LED pocket projector (3M, MPro150), a spherical 3D mask, a webcam (Logitech C525, HD 720p), and tracing paper (of 50-gram weight for the apparent projection of robot eyes). The pocket projector and 3D mask mounted on a supporting structure made of aluminium angles and sheets. Two servo motors (S8503 CYS) and a webcam also attached on this support structure. These servo motors perform the various head motions such as head turning, head shaking or head nodding. The webcam detects the frontal face of the interacting partner and his/her smile expression. The LED projector is placed on the face mask's rear position and projects computer graphic (CG) eye images on the face mask to generate eye blinks. Fig. 4 (a) shows the hardware components used to develop the robotic platform.

In order to establish a intercommunication among the various hardware constituents of the platform, there is a typical USB link between the multi-purpose computer (Windows 10, 64bit) and Arduino (MEGA 2560, 16 Mhz). The Arduino receives the serial command from the computer. The micro-controller controls the rotation and speed of the servo motors to produce head movements. Three U-shape supports and a base made of aluminium constructed where the first support can fold at the middle. This support used to carry the whole structure. The second support fixed with a servo motor and third support is attached with the shaft of this servo motor

to perform the movement. A servo motor (SM1) moves ±90° horizontally to create the pan movement of the head. Another servo motor (SM2) fixed with the upper frame at the left side of the frame to produce the tilt movement. A 3D spherical face mask is appended with the shaft of SM2 and placed inside this frame. SM1 and SM2 are wired to an Arduino board. The computer sends serial data to Arduino which can turn on or off the servo motors. A command with the angular values (in degrees) send to the Arduino which turns the servo motor at a specific position. To warrant stable running, the pan and tilt movements of the head suited by controlling the speed of servo motors in several experimental trials. Table 2 represents the key characteristics of the designed robotic head specified by the empirical observations.

TABLE 2. Characterises of robotic head.

Items	Features
Horizontal head movements	0°to 90°(right) and 0°to -90°(left)
Vertical head movements	0°(initial) to ± 20°
Blinks rate	20 per minute
Eye tracking distance (maximum)	2 feet
Gaze tracking accuracy	93.71%
Face detection accuracy	96.79%
Gaze awareness cue	head nod with eye blinks

B. SOFTWARE COMPONENTS

The proposed robotic platform comprises of four key software components: FDM (Face Detection Module), GDTM (Gaze Detection and Tracking Module), GAwM (Gaze Awareness Module), and RRCM (Robot Response and Control Module). The RRCM module governs the head motions based on the output of other modules. The states described in section III are controlled by these software modules. Figure 4 (b) illustrates integration of software modules with hardware.
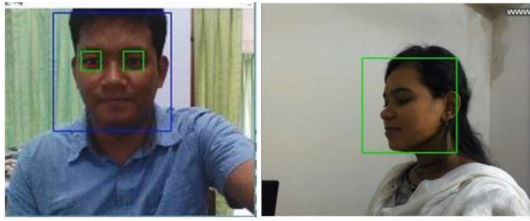
FIGURE 5. Output of the FDM module (a) frontal face detection, (b) lateral face detection.

1) FACE DETECTION MODULE (FDM)

If an agent (human or robot) wants to start an interaction, it should be looking at the interacting partner by turning face or gaze. If the human is staring at the robot in any case, it judges that she/he is interested or responded to communicate to it. The robot needs to be adjusted its angular orientation to ensures the gaze crossing. In that situation, the FDM module utilizes the forehead webcam to identify his/her face. The FDM uses cascaded classifiers based upon AdaBoost and Haar-like features to detect the human face [71]. This module works on grey-scale images captured by the webcam. The detector returns corner coordinates (x, y) of the given image including the height (h), and width (w). Basically, the state S_C that is detecting human's frontal (FF) and lateral (LF) face is run by this module.Fig. 5 shows the outcomes of the FDM module.

2) GAZE DETECTION AND TRACKING MODULE (GDTM)

The output of the FDM sends to the GDTM. It checks whether the human's eye is looking towards the robot or not. Finding the frontal appearance alone may not justify that the human is staring at the robot. In numerous circumstances, it is common to come-up a state where R detects the frontal face of the interacting partner; however, she/he is staring at another focus. To surmount this difficulty, we employed GDTM to identify the eye gaze of the human. This module identify and tracks the direction of the human's gaze. If H's eyes gazed through R, the result delivers to RRCM for confirming the endowment of gaze crossing. After detecting eye the system detect the eyeball. This eye detection has been showed in Fig. 6.
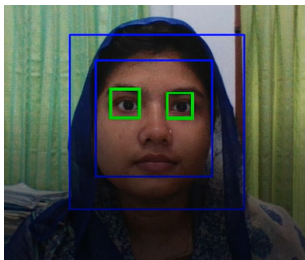

FIGURE 6. Detection of gaze of the human.

Algorithm 1 describes the process of eye-ball detection in the eye image, which ensures whether H is gazing at R or not. Fig 7 illustrates the output of each step of

---

Algorithm 1 Gaze Detection and Tracking Algorithm

Input: Detected face image
Output: Eye ball detection
foreach *Frame with face* **do**

Find eye region using haar features
Select only one eye portion
Convert the image into grayscale
Find edge using canny edge detector
Detect eyeball using the circular Hough transform algorithm
Calculate the position of the eyeball from the boundary of the eye region
if *position of the eyeball is* $\approx$ *center* **then**
Eye is looking forward direction

GDTM. Noteworthy, the state GC is checked in this module. We have computed the time complexity of Algorithm 1, which becomes $O(n^3)$.The complexity of finding the eye region from the face using Haar feature (of Viola-Jones algorithm) and convert it to grey-scale is linear which gives $O(n)$. The complexity of Canny edge detection and Hough transform are $O(mnlogmn)$ and $O(n^3)$ respectively. The complexity of selecting one eye and finding the distance of eyeball from the boundary is $O(1)$. Thus, for each frame complexity,

$$C = (O(n) + O(mnlogmn) + O(1) + O(n^3),$$

We can ignore the $O(n), O(1), and O(mnlogmn)$ complexity due their smaller contributions compared to $O(n^3)$. So, the complexity of each frame is $O(n^3)$ and the total complexity for f frame becomes, $C = O(f * n^3)$. We analyzed the recorded interaction videos to calculate the time needed for detecting gaze. Gaze detection time was computed by counting the time between after detecting the face and before starting the gaze awareness cue. Approximately 0.15 seconds needed to detect the gaze.

a: EYE REGION DETECTION

The human eye searched from the detected face using Haar-like features at the upper half of the face region [71]. This eye region is separated by examining the face region $[y : y + h/2, x : x + w]$, where x and y denote the corner coordinates of the face region including the width (w) and height (h). This returns the x_e, y_e, h_e, w_e values.

b: ONE EYE SELECTION

In the face-to-face orientation, since both eyes directed on the same focus, then detecting one eye (left or right) is enough to ensure gaze crossing. Detecting one eye also reduces the computational cost. Therefore, we choose one eye to find out the eye-ball position (Fig. 7(a)). In order to detach one eye, area of the eye-ball region will crop using the following coordinates,

$$one_{eye} = (x_e, y_e)(x_e + \frac{w_e}{2}; , y_e + h_e)$$

where x_e, and y_e represents the coordinates of the cropped eye vicinity. h_e and $w_e/2$ indicates the height and the width of the cropped eye.

c: GREY SCALE CONVERSION

The extracted portion of eye converted to grayscale image (fig. 7) (b)) using the technique of luminance [72]. Equation 1 used for the grey-scale conversion.

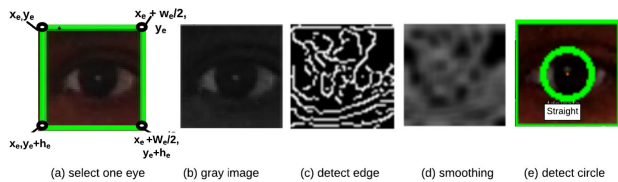$$g_{Luminance} = 0.299 \times R + 0.587 \times G + 0.114 \times B \quad (1)$$

FIGURE 7. Processing outcomes of GDTM.

(a) select one eye (b) gray image (c) detect edge (d) smoothing (e) detect circle

d: DETECTING EDGE

Edges are those positions of the images where perimeter or boundary is founded. The luminance of the pixels in an image is changed when it finds an edge. The properties and values of an edge calculate concerning the neighbour pixels, which is a vector variable. If the value of the grey level is similar to the grey level value of neighbourhood pixels, there is no edge considered. If there is a huge difference with the neighbourhood pixel, then decided that there is an edge. The canny edge detector is used (where $\sigma = 1.4$ and kernel size (5×5)) to detect the edges [73]. Fig. 7 (c) indicates result of edge detection.

e: SMOOTHING ROI

This step performs the smoothing of images by reducing the noise. Smoothing is done by blurring operation [73] in which convolution used with a kernel of low-pass filter. The averaging of the blurred edges are acquired by the convolution operation of the image to the normalized box filter. It merely average all pixels within the kernel area and alternates the principal components. Fig. 7 (d) shows the result of the smoothing operation.

f: CIRCLE DETECTION

The circle denotes the eye-ball in the eye image and a circle can be is illustrated as Equation 2.

$$(x - x_{center})^2 + (y - y_{center})^2 = r^2 \quad (2)$$

where (x_{center}, y_{center}) denotes the center coordinate, and r means the radius of the circle.

21HT circle detection algorithm is used for detecting the circle in the image [74]. Fig. 7 (e) illustrates the detected circle in the eye image.

g: CALCULATE THE POSITION OF THE EYE BALL

The centre of the eyeball is recognised concerning the eye region using the Hough transform [74]. If the eyeball is positioning at the centre of the image that will consider as directly looking to the camera, otherwise it will consider as reluctant in starting communication. Therefore, a result will forward to the control module as an indication of the gaze crossing set up. Having been detected the distance from the four boundaries to the centre allow the system to determine whether it is in centre or not. The distance from the upper, lower, left, and right boundary to the centre are denoted as disU, disD, disL, and disR respectively. The Equation 3 determines these distances from the position of the eyeball.

$$dis_i = \sqrt{(x_r - x_i)^2 + (y_r - y_i)^2} \quad (3)$$

where (x_r, y_r) denotes the center of the circle. and $x_i = x_e \text{ or } x_e + w_e/2$, $y_i = y_e \text{ or } y_e + h_e$. The haar feature sends the four coordinates of the eye region and it is already mentioned that only one eye has been selected for the further calculation. So, the four coordinates of the eye region are (x_e, y_e), $(x_e, y_e + h_e)$, $(x_e + w_e/2, y_e)$, $(x_e + w_e/2, y_e + h_e)$ (figure 7(a)) where h_e and $w_e/2$ indicates the height and the width (w_e indicates the width of two eyes, hence, for one eye we take the value of $(w_e/2)$. At this point, the distance from the four boundaries to the centre would be the following: $disU = \sqrt{(y_r - y_e)^2}$, $disD = \sqrt{(y_r - y_e + h_e)^2}$, $disL = \sqrt{(x_r - x_e)^2}$, and $disR = \sqrt{(y_r - x_e + w_e/2)^2}$
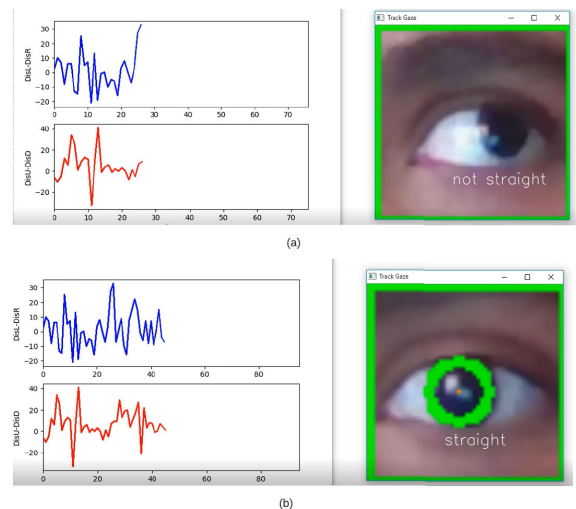

FIGURE 8. Results of GDM (a) H is not looking at R (b) H is looking at R.

The conditions $disU \approx disD$ and $disL \approx disR$ indicates that the participant is gazing directly toward the robot. Fig. 8 (a) depicts a case where the participant is not gazing at the robot, whereas Fig. 8(b) indicates a case in which the participant looked at the robot.

3) GAZE AWARENESS MODULE (GAwM)

The GAwM recognise the facial expression of the human during the interactions. The purpose of this module is to confirm
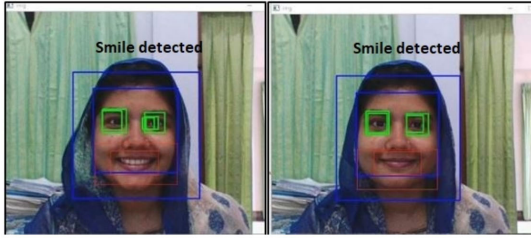
FIGURE 9. Detection of gaze responsive behavior of the human.

FIGURE 11. Responsive cues for robot: (a)-(b): snapshot of head turn (c)-(e) snapshot of head shake.

that the participant is not gazing toward the robot voluntary but also she/he understood its action. The GDTM sends the results of eyeball detection to activate GAwM. It is clear that the state GA_wH is controlled by this module to detect human response. The GAwM detects the smile of the human as gaze awareness signal using haar features. The GAwM detects the smile expression as the gaze awareness behaviour of the human and sends this result to RRCM to execute the head nods with eye blinks as the gaze awareness behaviours of the robot. There is a variation of human's eyes and mouth regions when she/he is laughing. We used the Viola-Jones algorithm [71] to detect the face, and gaze. We used a method introduced by Deniz *et al.*, [75] to detect the smile expression. This method used Viola-Jones cascade classifier. The cascade classifier is trained to detect a smile expression by superimposing the images with smiling faces over 2436 positive and 3376 negative images. The Viola-Jones algorithm uses haar-like features to detect facial properties and so smile. Haar cascade is used for smile detection, which returns the red coloured rectangle area of the mouth. The coordinates of the detected face $(x, x + w; y, y + h)$ get from FDM. These coordinates are sent to GAwM to detect the smiling. Haar-cascades are classifiers, a series of filters used to detect smile features by superimposing predefined patterns over smiling face segments and used as XML files. These filters are applied one after another to detect a smiling face through its features. The cascade is a series of filters that will apply one after the other to detect a smiling face through its feature. Fig. 9 shows the outcome of smile detection. Fig. 10(a) shows some haar features applied to detect smile on a face and fig. 10(b) shows a detected smiling face.
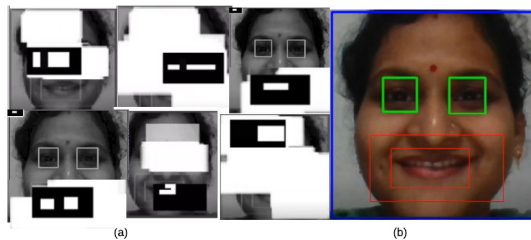

FIGURE 10. (a) haar features used to detect smile (b) Smiling face detected.

4) ROBOT RESPONSE AND CONTROL MODULE (RRCM)
By using the servo motors (SM1 and SM2), the robot performed all physical motions with the appropriate control
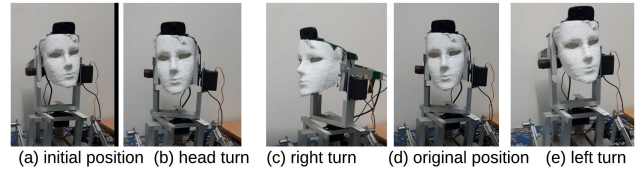
signal arriving from the different modules. In the current implementation, the robot can execute specific behaviours during HRI at M_S state such as the head-turning (HT), head shaking (HS), head-nodding (HN). At RA_W state the robot will perform head-nod (HN) and eye blinks (EB) to show awareness.

- Head-turning (HT): This action is utilized to shifts the robot's head to the interacting participant from its original setting. We settled the pan speed of SM1 at 17°/second. The setting of the robot and the participant are settled. Thus, the robot requires to move its SM1 about 35° to ensure facing each other. Fig 11 (b) depicts HT action of the robot after changing its position toward the human from the original state (Fig 11 (a)).
- Head shaking (HS): The HS used to perform the waving action of the robot's head. This action designed by moving the head back and forth about ±20° from its original setting. That implies, the robot shifts once its head 20° left and 20° right. The speed of head-shaking is settled at 17°/second. The HS action performed by controlling the angular movements of SM1 and its speed.
- Head nodding (HN): Nodding head action can use as the responsive gaze behaviour of the robot. After detecting the participant's face and his/her eyeball, the robot performs the head-nodding actions. the robot shifts once its head 20° down and again lift its head 20° to get into the initial position. The speed of head-nodding is also settled at 17°/second. Figs. 12(a)–12(c) shows the snapshots of head-nodding action.
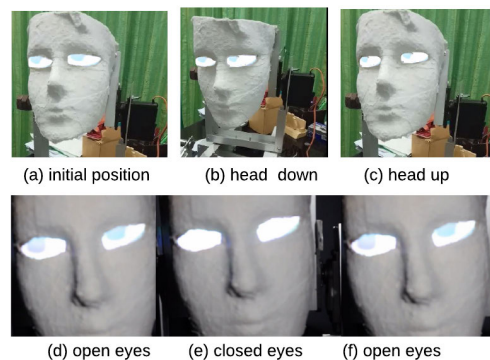

FIGURE 12. Responsive cues for robot: (a)-(c): snapshot of head nod (d)-(f) snapshot of an eye blink.

- Eye blink (EB): Eye blinks played by the prompt closing and opening of the eyelids of CG images which exhibited on the robot's eyes through the LED projector.

We fixed the robot to execute eye blinking at a rate of 1 blink/second. Fig. 12 (d)– 12(f) illustrates a few snapshots of blinking action.

- Eye blinks with a head nod (EB+HN): In case of showing awareness, we performed both HN and EB at the same time which follow the rules mentioned above.

The outcomes of the FDM, GDTM, and GAwM are sent to the servo motors through Arduino to perform various movements of head (such as turning, nodding, and shaking). The pan servo (SM1) uses to generate the head turning and shaking movements to attract the interacting partner's attention after detecting the face by FDM. The tilt servo (SM2) produce gaze responsive action by rotating its head once where this motion of the face mask is identical to the motion of human's neck movement up and down. Moreover, the robot display its gaze awareness by nodding its head once to let the human notice that it understand his/her gaze response. The actions of the RRCM depends on controlling the two motors by the signals (α_1, α, γ) mentioned in section III. The functions of RRCM module depends on the few predefined rules such as (i) at first R turns its head from initial position, then $\alpha_1 = 1$ (ii) if $FDM = 0$ then $\alpha_1 = 0$, R turns back to the initial position (ii) If *Face position=LF||FF && GDTM == false*; then $\alpha = 1$, (iii) In RIC, if *Face position=FF;GDTM == true* then $\gamma = 1$, on the other hand, in HIC if *Face position=FF;GDTM == true&&GAwM == true* then $\gamma = 1$,

Based on the values of α, β, and γ, the motors perform movement actions according to the rules (Eqs. 4-5).

$$SM1(x) = \begin{cases} Turns, & \text{if } \alpha = 1, \\ Shakes, & \text{if } \beta = 1, \\ Stop. & \text{otherwise} \end{cases} \quad (4)$$

$$SM2(x) = \begin{cases} Nods \text{ } once & \text{if } \gamma = 1 \\ No \text{ } movement. & \text{otherwise} \end{cases} \quad (5)$$

V. PRELIMINARY STUDY 1: TO VERIFY THE EFFECT OF ROBOT's TILT MOVEMENTS
The purpose of this study is to settle how much tilt movements (in angle) are appropriate for designing the head nod as a responsive gaze cue.

A. EXPERIMENTAL DESIGN
Prior to conducting the experiment, attendees were requested to seat on a chair stable in a predetermined location in a laboratory setting. Additionally, they asked to keep looking at the robot when it adjusted its head in such an orientation so that they could ensure face to face settings. Fig. 13 depicts the setting and a scene of the experimental environment.

Each trial commenced with the robot's gazing at the subject and terminated with the tilt actions of the robot. Before initiating the experiment, the subject was explained about the scope of the experiment. We designed the robot in three ways. That means the robot display its tile motions in there angular conditions: (i) at 10°, (ii) 20° and (iii) 30° respectively.
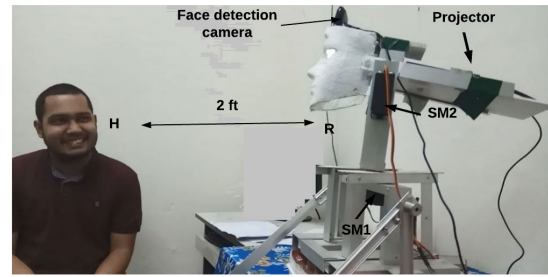

FIGURE 13. Experimental setting with a scene.

The subjects were requested to give a response to a questionnaire after interacting all the conditions. The responses are collected in terms of 1- to- 5 Likert scale where 1 denotes the less effective and 5 stands very effective. Noteworthy, the experiment was a within-subject design where the sequence of all experimental trials was counterbalanced.

- Evaluation on tilt movement: Your preference of the tilt movement suitable to design the nodding behavior.
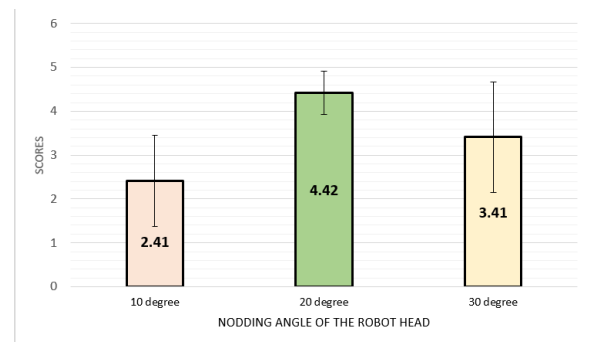

FIGURE 14. Average score of preference on three tilt motions. Error bars indicate the standard deviation.

B. RESULTS
Twenty human subjects were associated with the experiment. Their (12 female, 08 male) mean age was 21.9 years ($SD = 0.7$). All of them were undergraduate students in the engineering discipline. Figure 14 illustrates the subject's preferences on the tilt angle movements suitable for nodding. Repeated measures of analysis of variance (ANOVA) reveals a substantial difference among the conditions $F(2, 57) = 32.38, p < 0.05$. We performed multiple comparisons using Bonferroni technique which reveals a substantial differences between conditions for 10° vs 20° ($p = 8.2325e - 10 < 0.05$), 20° vs 30° ($p = 4.1039e - 07 < 0.5$), and 10° vs 30° ($p = 0.3288998 < 0.5$) respectively.

Results indicate that the tilt movement with a 20° angle achieved the highest score of $\mu = 4.42, SD = 0.49$) in all conditions, which revealed that this movement preferred most by the participants. Therefore, we will use 20° as the tilt movement to design the head-nodding cue of the robot.

VI. PRELIMINARY STUDY 2: TO VERIFY THE EFFECT OF NUMBER OF NODDING
Several actions have used in previous studies as the gaze responsive behaviour of the robot such as eye blinks [12],

CG smiling [8], and ear blinks [60]. In this work, we propose head nods as the responsive gaze behaviour of the robot. That means the robot displays the head nodding as responsive behaviour after confirming the human is looking at it with smiling. During nodding, the number of the nodding operation may also be a vital factor to convey gaze awareness behaviour. Therefore, a study was carried out to verify the effect of the number of nodding to design a significant head-nodding cue for robot initiative case with a face-to-face setting.

A. PARTICIPANTS

A total of twenty undergraduate students (12 female, 08 male) took part in the experiment. The mean age of the participators was 21.9 years ($SD = 0.7$), and they have no prior experience in any HRI experiments.

B. EXPERIMENTAL DESIGN AND PROCEDURE

The study was performed in a controlled environment. Head nodding considered as gestural information of positive feedback, to support questions, and to emphasize agreement with the conversational partner in face-to-face setting [76]. Therefore, a robot is prepared to play the head-nodding actions to create a feeling of gaze awareness among participants. We asked the participant to sit in a predefined position. The setup of this experiment is identical to the previous experiment (Sec. V). An experimental trial began with the establishing of gaze crossing between the participant and the robot. This gaze crossing process initiates by the robot looking at the participant (by turning its head), and she/he responds to the robot with a smile. The trial ended with a nodding action while the robot detected the smile of the participant. To detect the participant's face and his/her smile, a webcam attached on the robotic head. Before beginning the experiment, participants were demonstrated the objective of this experiment is to measure the appropriateness of an action of the robot to make them feel that it notice their looking respond. The robot nodded its head in response to the participant looking at it and observed how many times of nodding is enough to interpret the responsive gaze behaviour of the robot.

C. EXPERIMENTAL CONDITIONS

The participants were requested to observe the three conditions one after another: (i) the robot nods once, (ii) the robot nods twice, and (iii) the robot nods thrice. Noteworthy, the design of the experiment was a within-subject where the order of all experimental trials has been counterbalanced. Finally, after completing all the sessions in three conditions, the participant requested to provide the answer on the following questionnaire using a 1- to- 5 Likert scales (1 denotes for not effective, and 5 stands for very effective).

- Evaluation: Your preference for the number of nodding of the robot.

D. RESULTS

Fig. 15 shows that the participant's preferences for the number of nodding action of the robot. The result indicates that nodding once ($\mu = 4.3$, $SD = 0.59$) achieved a higher score than the nodding twice ($\mu = 2.8$, $SD = 0.6$), and nodding three times ($\mu = 1.92$, $SD = 1.44$).
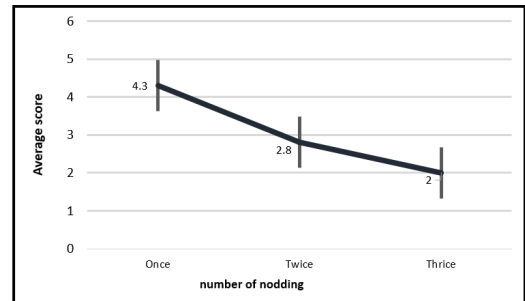

FIGURE 15. Mean values of subject impressions of different number of head nodding.

We conducted ANOVA analysis which shows a substantial difference among conditions ($F(2, 57) = 24.5142$, $p < 0.05$). Multiple comparisons with Bonferroni methods also reveals the significant means effect such as, nodding once vs twice: ($p = 0.0001028$; $p < 0.05$), nodding one time vs three times ($p = 1.4203e - 08$; $p < 0.5$). However, no substantial difference is noticed between the nodding twice vs thrice ($p = 0.0592380$). Thus, results indicated that nodding once is acceptable for the participant to understand as the responsive gaze behaviour of the robot.

VII. PRELIMINARY STUDY 3: TO VERIFY THE EFFECT OF GAZE AWARENESS CUES

In order to establish a perfect communication channel, both parties should understand each other actions explicitly. In HRI, it is easier for the human to interpret other cues, but for the robot, it is a quite tricky task. Moreover, the robot needs to be capable not only to identify the human's gaze awareness cues but also presents its gaze awareness cue explicitly so that the human partner can interpret easily. Several actions have been used to display the gaze awareness behaviour of the robot in previous studies such as eye blinks [12], [77], and graphics smiling [8]. The aim of this study is to assess the effect of three actions such as head nodding, eye blinks, and a combination of head-nodding with eye blink as the gaze awareness cue for the robot.

A. DESIGN AND PROCEDURE

The experiments performed in the laboratory setting. A total of 32 undergraduate students (14 male, 18 female, average age = 22.03 years, SD = 2.09) attended in this experiment. We requested participants to look around the robot randomly from a predefined sitting position. The robot is shaking its head once to capture the attention of the participant. We instructed the participant to look at the robot with a smile

if he/she notice shaking action. A USB camera fastened on the robot's head to identify the face of the participant and his/her smile. The robot displays gaze awareness cues according to the conditions after smile detection. Each participant attended all conditions one after another. A participant experienced three trials in each condition, and the average duration of each trial was approximately 60 seconds. All experimental trials were video recorded. The setup of this experiment was identical to the earlier experiment (as described in Section V).

B. EXPERIMENTAL CONDITIONS

To assess how the proposed gaze awareness cue affects the performance, we designed other two alternative cues for comparison. The experiment conducted as a within-participant scheme and the sequence of all trials was counterbalanced. We implemented the robotic head in three distinct styles, including the proposed technique as described in the following.

- **Method 1 (Blink only)**: R blinks it's eyes after detecting H's face. The blinks cue is produced by projecting the CG images on the tracing paper laid on the robot's eye.
- **Method 2 (Nod only)**: R display nods once, after detecting H's face.
- **Method 3 (Blink+Nod) (Proposed)**: R blinks its eyes first one time and then nod head once, after detecting H's face.

C. EVALUATION

The attendees were requested to deliver assessments of all robots using a 1- to- 5 Likert scale in the questionnaire (1 denotes for the lowest and 5 for the highest). The questionnaire comprises four questions (Q1-Q4).

- (Q1) Did you think that the gaze awareness cue of the robot is reasonable?
- (Q2) Did you follow the gaze awareness cue of the robot?
- (Q3) Did you think that the action of the robot acknowledged to your response?
- (Q4) Did you feel that the behaviour of the robot is useful to create your feeling of gaze awareness?

TABLE 3. Participant's preferences on three gaze awareness cues.

Questions	Methods		
	Method 1 Mean (SD)	Method 2 Mean (SD)	Method 3 (Proposed) Mean (SD)
Q1	2.28 (0.94)	3.34 (1.05)	4.34 (0.599)
Q2	2.44 (1.19)	3.34 (0.99)	4.69 (0.52)
Q3	2.66 (1.42)	2.84 (1.25)	4.66 (0.73)
Q4	2 (1.27)	3.16 (1.03)	4.63 (0.54)

D. RESULTS

Table 3 illustrates the mean and standard deviation (SD) values of the attendee's assessment. This result shows that the proposed scheme achieved a higher score than the other schemes. For further investigation, a chi-square test used

to find statistically significant differences among the three methods. Scheffé test has been performed for the post-hoc analysis which declares the differences between the methods.

Concerning Q1, a statistical significant is revealed among three methods ($\chi^2(8, 32) = 50.96, p < 0.00001$) which indicates that all methods are not equally founded reasonable by the respondents. Scheffé test also shows there is a significant differences between pairs: M1 vs M3 ($t = 8.96, p < 0.01$), M2 vs M3 ($t = 4.34, p < 0.01$), and M2 vs M1 ($t = 4.64, p < 0.01$). In the case of Q2, there no statistically significant differences are found among three conditions ($\chi^2(8, 32) = 68.44, p < 0.00001$, significant at $p < .05$). Scheffé test also shows the significant differences for the pairs: M2 vs M3 ($t = 5.39, p < 0.01$), M1 vs M3 ($t = 9.36, p < 0.01$), and M1 vs M2 ($t = 3.77, p < 0.01$).

Concerning Q3, the respondents did not showed the similar feelings for all methods and chi-square test shows a significant differences ($\chi^2(8, 32) = 45.5, p < 0.00001$, at $p < 0.05$). Scheffé test also revealed that there is a significant difference between pairs such as M2 vs M3 ($t = 6.07, p = 2.01Xe - 08$) and M3 vs M1 ($t = 6.698, p < 0.01$) but no significant difference was found for the pair M2 vs M1 ($p < 0.05$). In the case of Q4, statistically significant differences are found among three conditions ($\chi^2(8, 32) = 76.4, p < 0.00001$ at level $p < 0.05$). Scheffé test also indicates significant differences between the pairs: M2 vs M3 ($t = 5.79, p < 0.01$), M3 vs M1 ($t = 10.35, p < 0.01$), and M1 vs M2 ($t = 4.56, p < 0.01$).

Although further investigation needed using more participants, the primary analysis of this study confirms that the combination of head nod and eye blinks are useful than other cues (i.e., only head or eye movements) to create the feeling of gaze awareness of the participants.

VIII. DUPLEX EYE CONTACT EXPERIMENTS

The central concern of this work is to design a duplex eye contact mechanism for HRI. In particular, the robot or the human can establish an eye contact process between each other using two modes: (i) robot initiative where the robot starts the eye contact event and (ii) human initiative where the human starts the eye contact event. Thus, we conducted two independent experiments to evaluate the proposed framework considering two modes.

TABLE 4. Cues extracted from preliminary experiments 1, 2, and 3.

Preliminary Exp.	Action	Extracted cue
#1	Nodding angle	20°
#2	Nodding times	Once
#3	Awareness action	Nodding with blinking operation

The robot used two signals or cues to perform gaze crossing functions: head-turning and head shaking. A combined signal is used to perform gaze awareness: head-nodding with eye blinks. Table 4 shows the signals extracted in experiments.

A. PROCESS OF DUPLEX EYE CONTACT

We assume that H and R are currently attending in their concerned tasks where they are not in face-to-face orientation. The proposed duplex eye contact framework functioning independently in two modes: HIC and RIC. Thus, selecting mode is the primary step of the framework, and this mode selection is made manually. Figure 16 depicts a flowchart illustrating the duplex eye contact process concerning HIC and RIC. In HIC, it is supposed that the H is approaching the R to begin an interaction. The R adjust its head orientation (if necessary) by turning to H, detects his/her face and eye gaze for ensuring the establishment of gaze crossing. If the H display a smile expression as gaze responsive behaviour, R detects this smile and display a head nod with eye blinks as a signal of gaze awareness. The system considers that eye contact is made successfully after completing R and H's gaze awareness signal.

In RIC, the system observes the H and detect his/her face. R first turns its head, holds (up to 2 seconds) for H's looking response and commences head shaking (if necessary) to capture H's attention. If H is staring at R by rotating his/her face, the system determines that he/she has responded to R's motions. It is recognized by detecting H's face and eye gaze in the camera images, ensuring face-to-face orientation (i.e., gaze crossing). If H is not looking toward R within 2 seconds, the system gives up to establish eye contact. After gaze crossing, R performs a head nod with an eye blink to display gaze awareness. H also shows a smile expression as a response to R's gaze awareness signal. R detects the smile expression, and the system considers that eye contact is established after ensuring R's and H's gaze awareness. The framework considers the case as a failure if it cannot detect H's face, gaze or smile and cannot generate head nod or eye blinks. The framework produces a beep sound in each mode to indicate eye contact's success (this is for experimental purpose).

B. EYE CONTACT EXPERIMENT IN HUMAN INITIATIVE CASE

The aim of this experiment is to assess the proposed mechanism when the human intends to develop eye contact with the robot (i.e., human initiative mode).

1) PARTICIPANTS

A total of 12 undergraduate students (two female and ten male) took part in this study who had no previous experience in HRI experiments. The mean values of their ages were 20.9 ($SD = 1.5$) years. Noteworthy, no remuneration was provided to the students.

2) DESIGN AND PROCEDURE

The study was carried out at the robotics lab, Chittagong University of Engineering and Technology, Bangladesh, where the developed robotic head settled on a table. The participant requested to interact with the robot from a predefined
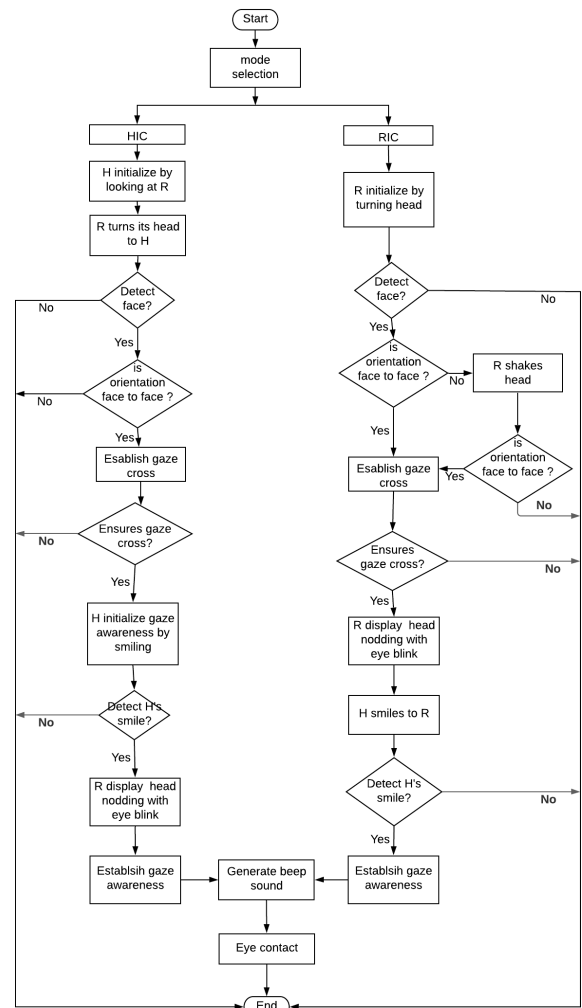

FIGURE 16. Flowchart of the proposed duplex eye contact framework.

position. Before starting the experiment, the experimenter manually adjusted the different parameters (such as pan, tilt, and zoom) of the head-mounted camera. Our primary intention was to let the participants assess different behaviours of the robot when she/he interested to make eye contact with it. Each participant interacted with the three methods, one after another, and each method consists of three trails. In the beginning, a demo behaviour is shown by the experimenter about the interaction protocol. The robot (in each condition) initialize eye contact process after recognizing the participant's face. All interactions are videotaped using a video camera. Fig. 17 illustrates the setting and a scene of the experiment.

3) EXPERIMENTAL CONDITIONS

To investigate the effect of the proposed system on performance evaluation, we compared it with two other methods. The design of the study had a within-participant, and the sequence of all trials was counterbalanced. Every individual interacted with the following three methods (M1-M3), one after another.
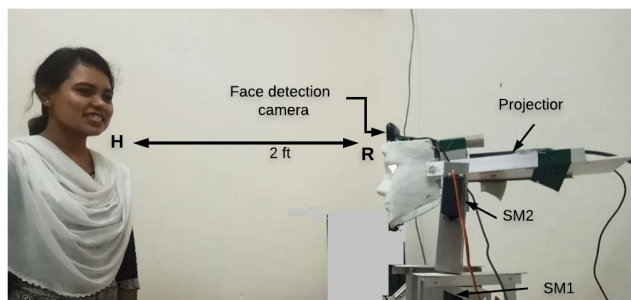
FIGURE 17. The setting of the experiment with a scene in human initiative case.

- **M1**: R is in a static initially. H approaches it by coming forward. R detects his/her face but did not turned its head from the initial position. If the human smiles, the robot detects it and blinks its eyes three times as the gaze awareness behaviour.
- **M2 [Proposed Method]**: R is in static initially. It turns its head towards H when she/he approaches to it which ensure gaze crossing. R detected H's smile and nod its head with blinking eyes as gaze awareness cues. Section III and Sec. IV describes this method in details.
- **M3**: R moves its head back and forth once initially. If the human approaches to the robot, it detects his/her face and stops moving. It adjusts its head to establish gaze crossing. The robot did not display any action (i.e., remain static) after detecting the participant's smile.

4) HYPOTHESES/PREDICTION

We proposed an HRI mechanism where the human feels that she/he set up eye contact with the robot when approaches to it. In order to perform the eye contact, the proposed method incorporated gaze crossing and gaze awareness components. While two alternative methods also intend to establish eye contact, which may lack some functions, but the proposed method employed these functions. Thus, our hypotheses argue that if the robot is successful in performing both gaze crossing and gaze awareness functions with appropriate cues, then it will establish more effective eye contact. In this regard, the proposed method can produce more meaningful interactions for making eye contact in the human initiative mode. Based on this concern, we assumed that the experiment would verify the following predictions (P1-P4).

- P1: Participants feel that the proposed method can understand his/her intention to interact with the robot.
- P2: The proposed method provides better gaze responsiveness to the participants.
- P3: Participant's feeling of making eye contact with the proposed scheme is better than others.
- P4: The proposed scheme outperforming the other two methods for the overall assessment.

5) EVALUATION MEASURES

We estimated the following qualitative measure in the experiment:

- **Impression of making eye contact**: After completion of all interactions, we proffered a questionnaire to collect the participant's impressions on a 1- to- 5 Likert scale. The questionnaire consisted of the following four items (Q1-Q4):
  - (Q1) Did you think that the robot crossed its gaze when you have approached it?
  - (Q2) Did you understand that the robot displayed appropriate gaze responsive behaviour to your approach?
  - (Q3) Did you think that the behaviours of the robot produced your feeling of establishing eye contact with it?
  - (Q4) How effective the mechanism was to make eye contact?

6) RESULTS

The result of the questionnaire shows that the proposed method always get the highest score from the participants. Fig. 18 shows the participants' responses to Q1-Q4. We used a chi-square test to calculate the statistical differences among three schemes. Besides, Scheffé test has been performed for the pairwise comparisons which declares the differences between the methods.
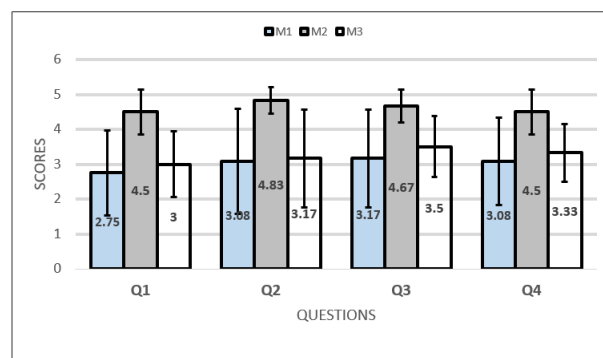

FIGURE 18. Participant's impression on Q1, Q2, Q3, and Q4. The error bars denotes the standard deviation values.

Concerning the fact of understanding the gaze crossing behaviour, the chi-square result of Q1 shows that there is a significant differences among the methods ($\chi^2 = 22.2, p = 0.0047$). As turning the head is the fundamental cue to convey the communication intention of interacting partners, the proposed method achieved a higher score ($\mu = 4.5, SD = 0.65$) than the methods M1 ($\mu = 2.75, SD = 1.22$) and M3 ($\mu = 3.00, SD = 0.95$). Scheffé test also shows that there is a significant difference between pairs: M1 vs M2 ($t = 4.4, p = 0.0005$) and M2 vs M3 ($t = 3.78, p = 0.0026$), However, no significant difference is observed between M1 vs M3. This result shows that the proposed method creates the better feeling among the participants of understanding human's approach. Thus, prediction 1 is verified.

In the case of perceiving the gaze responsive behavior, the chi square analysis shows that there is a significant differences among the methods ($\chi^2 = 18.125, p = 0.02$;

significant at $p < 0.05$). Pair-wise comparison with Scheffé test indicates the significant differences between pairs: M1 vs M2 ($t = 3.41, p = 0.007$) and M2 vs M3 ($t = 3.24, p = 0.01$). However, there is no significant difference is observed between M1 vs M3. The results revealed that the participants can easily understand the actions of the proposed robot as gaze awareness behaviours. Hence, prediction 2 is verified.

Concerning Q3, Chi-square analysis shows a significant differences among the three methods ($\chi^2 = 18.67, p = 0.016$; significant at $p < 0.05$). Pair-wise comparison using Scheffé test shows there is a significant difference between M1 vs M2 ($t = 3.89, p = 0.001$ at $p < 0.05$) and M2 vs M3 ($t = 3.03, p = 0.01, p < 0.05$). Nevertheless, no significant difference was reported for M1 vs M3. The results depicted that the feeling of eye contact with the proposed method is a better way than the other two methods. Thus, the prediction 3 is verified.

Concerning the overall evaluation, a statistically significant difference is found ($\chi^2 = 23.19, p = 0.003$; significant at $p < 0.05$) from chi-square test. Multiple comparisons with Scheffé test shows that the result of M1 vs M2 ($t = 3.49, p = 0.0056$) and M2 vs M3 ($t = 2.88, p = 0.02$) is significant at $p < 0.05$. However, there is no significant difference was reported for the pair M1 vs M3. This result indicated that the performance of the proposed method outweighs the other two. Therefore, the prediction 4 is verified.

C. EYE CONTACT EXPERIMENT IN ROBOT INITIATIVE CASE
This experiment aims to assess the proposed mechanism in the robot initiative mode (i.e. while the robot intends to develop eye contact with the human).

1) PARTICIPANTS
A total of 12 participants (8 males, and 4 females) were associated with the experiment. All of them are undergraduate students of a public university of Bangladesh and their average age was 20.92 ($SD = 1.32$). Participants had no previous experience to interact with the robot. There is no remuneration paid for the participants.

2) DESIGN AND PROCEDURE
To simulate the interaction, we consider a scenario 'reading the book'. That means the human involves to a task (i.e., reading the book) and the robot tries to make eye contact with him/her. The robot did not play any movement during the first 60 seconds of the interplay. The robot placed on the table and the parameters of the head-mounted camera manually adjusted so that it can track the face of the participant. To produce the stimuli, the robotic head programmed in three different conditions. During the reading, the participant experienced one stimulus at a time. In each condition, the robot attempted to interact in three trials. Fig. 19 shows some scenes of the experiment during interaction with the robot.

The robot attempts to capture the attention of the interacting partner by its head motions when s/he reads the book. We asked the participant to look at the robot when she/he
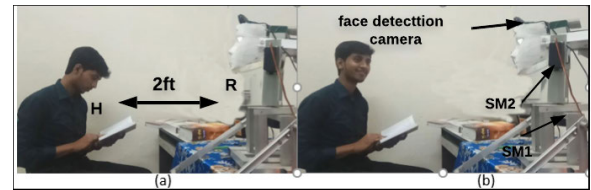

FIGURE 19. Setting and scenes of the experiment (a) H busy with reading the book; R tries to attract H (b) H perceived R's action; H smiles and in response R nods back with blinks.

feel attracted by its motions. The turning angle of the robot adjusted in such a way so that face-to-face orientation ensured between them. That is means the gaze crossing will establish when the participant and robot are looking at each other. If the participant looks at the robot within 2 seconds, it detects his/her face and gaze. After detecting the gaze, the robot display is gaze awareness behaviours and complete eye contact process. Before starting the session, participants were presented a demo behaviour by the experimenter about the interaction protocol. In any trial, if the participant did not attend toward the robot within 2 seconds following the robot's motions, then it regarded the trial as a failure and ended the session. All interaction sessions were videotaped to analyze the behaviours of the participants.

3) EXPERIMENTAL CONDITIONS
The success of establishing eye contact between the robot and the human depends on their orientation and the nature of the task they engaged [12]. On the other hand, capturing someone's attention (to initiate eye contact process) depends on the intensity and nature of the action played by the robot. A mild action may be acceptable to win people attention in some cases, but most circumstances demand intense action. Based on this consideration, we design the proposed robot with both head turn and shaking actions to gain the attention of the interacting partner. That means, the robot usually use head-turning action to attracts the human's attention. However, it commences head-shaking if it fails to capture his/her attention by the head-turning action. We proposed an HRI framework for making eye contact by employing gaze crossing and gaze awareness components. Therefore, it is essential to distinguish the proposed approach with others that lack or weak eye contact functionalities. In order to assess the suitability of the proposed scheme, we programmed the robotic head with three alternatives for comparison. Every participant interacted the following four approaches, one after another.

- Method 1: The robot turns its head toward the participant. If the participant looks at the robot within 2 seconds, it detects his/her face and gaze. After detecting the gaze, the robot blinks its eyes once.
- Method 2: The robotic head turns to the participant. If she/he looks at the robot within 2 seconds, it detects his/her face and gaze. After detecting the gaze, the robot nods one time.

- Method 3: The robot shakes its head and ends shaking by looking at the participant. If the participant looks at the robot within the expected time frame, it blinks eyes once.
- Method 4 [Proposed]: The robot customarily turns head toward the participant and originates shaking its head (if necessary) to capture his/her attention. If the participant stares at the robot, it blinks its eyes with head nods one time. The details design of each cue and operating procedure of this robot illustrated in Sections III and IV.

4) HYPOTHESES AND PREDICTIONS

Turning the head toward the interacting partner considered the basic action to convey the communication intention in HRI studies [12]. However, it might be tough to attract people's attention merely by this action. Especially, when the target participant engaged in high attention absorption task or the interaction partners are not facing each other. On the other hand, some cues create a better feeling of establishing eye contact than other cues and can convey more meaningful information to the interacting partner about its intention. Thus, the following hypotheses (H1-H4) would be checked by the experiment.

- (H1): Participants recognize that the proposed robot is better at initiating interaction by attracting their attention.
- (H2): Participants feel that the proposed robot communicates its gaze awareness behaviour more effectively.
- (H3): Interacting participants understand that the proposed robot generates a better impression of making eye contact.
- (H4): The proposed approach outperforms the other three approaches for the overall assessment.

5) EVALUATION MEASURES

The design of the study had a within-participant, and the sequence of all trials was counterbalanced. After completion of all interactions, we proffered a questionnaire to collect the participant's opinions on a 1- to- 5 Likert scale. The questionnaire contained the following four items:

- **Initiating interaction**: Did you think that the behaviour of the robot gained your attention to it?
- **Impression on gaze awareness**: Did you feel that the robot conveyed its gaze awareness behaviour explicitly during the interaction?
- **Feeling of making eye contact**: Did you think that the robot transmitted the feeling of establishing eye contact?
- **Overall evaluation**: How useful is the method to establish eye contact?

6) RESULTS

We used a chi-square test to determine the statistically significant differences among the four methods. Figure 20 shows the mean values participant's responses to Q1, Q2, Q3, and Q4. From the figure, it is clear that the proposed method always get the highest scores.
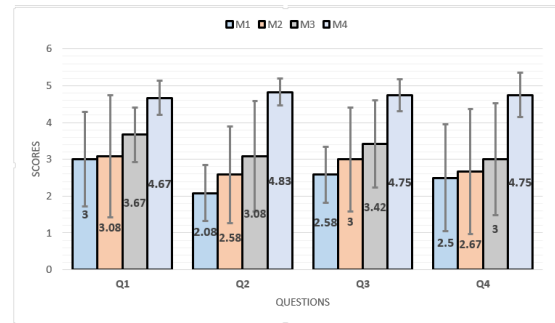

FIGURE 20. Participant's score on different methods. The error bars denotes the standard deviation values.

Concerning to initiating the interaction, the chi-square test result shows that there is no significant difference among the four methods ($\chi^2 = 28.15, p = 0.005$) as all methods gained attention of the participants. However, multiple comparisons using Scheffé test revealed significant differences between pairs M1 vs M4 ($t = 3.43, p = 0.014$) and M2 vs M4 ($t = 3.26, p = 0.022$). Apart from these, there are no significant differences between pairs M3 vs M4, M1 vs M2, M2 vs M3, and M1 vs M3. In addition, the proposed method achieved the higher scores ($\mu = 4.67, SD = 0.47$) than the methods M1 ($\mu = 2.75, SD = 0.76$), M2 ($\mu = 3.08, SD = 1.65$) and M3 ($\mu = 3, SD = 1.29$). This result signify that the proposed method with the head turning and shaking actions preferred by the participants to capture their attention for initializing interaction. Thus, hypothesis 1 is verified.

In the case of impression on gaze awareness, chi square test represents a significant difference among the methods ($\chi^2 = 34.3, p = 0.0006$, significant level $p < 0.05$). Multiple comparison with Scheffé test depicted the significant difference between pairs: M1 vs M4 ($t = 5.95, p = 8.47Xe - 06$), M2 vs M4 ($t = 4.86, p = 0.003$), and M3 vs M4 ($t = 3.78, p = .006$). Nevertheless, no significant difference was reported for the pairs: M1 vs M2, M1 vs M3, and M2 vs M3. This results showed that the participants perceived the proposed method is better to display the gaze awareness behaviours. Therefore, hypothesis 2 verified.

Taking concern about the feeling of making eye contact, chi-square analysis reveals that there is a significant differences among the four methods ($\chi^2 = 25.4, p = 0.042$, significant level at $p = 0.01 < 0.05$). Multiple comparison using Scheffé test shows a significant difference between pairs: M1 vs M4 ($t = 4.97, p = 0.0002$), M2 vs M4 ($t = 4.02, p = 0.003$), and M3 vs M4 ($t = 3.06, p = .035$). However, there is no significant difference was reported between pairs M1 vs M2, M2 vs M3, and M1 vs M3. The results revealed that the proposed method produces the better feeling of making eye contact than other methods. Therefore, the hypothesis 3 verified.

Having been analyzed the overall evaluation, a significant effect was found for Q4 ($\chi^2 = 26.37, p = 0.009$) using chi-square test among the four methods. Multiple comparison using Scheffé test reveals that there are significant differences between pairs: M1 vs M4 ($t = 3.81, p = 0.0053$), M2 vs M4
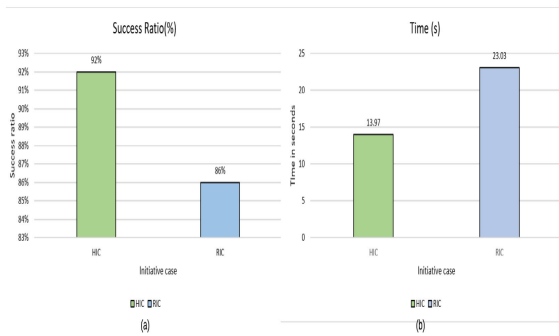
FIGURE 21. Overall evaluation of HIC and RIC: (a) Success ratio(%) (b) Average time (s).

($t = 3.53, p = 0.011$), and M3 vs M4 ($t = 2.97, p = .04$). However, there is no significant differences were found for the pairs: M1 vs M2, M1 vs M3, and M2 vs M3. This results indicates that the proposed method seems interesting to the participants and they found it more effective than the other methods for making eye contact. Thus, the hypothesis 4 is verified.

D. HIC VS RIC

For making successful eye contact both parties should understand the communicative intention of each other and behaves appropriately. In addition to that both agents (human or robot) interpret each other signals or cues explicitly. In order to make a comparative sketches between human initiative and robot initiative approaches, we measured the Success Ratio (SR) and Average Time (T_{avg}).

SR can be determined as the ratio between the number of times that the agent performed the gaze awareness activity successfully (A_{GA_w}) and number of times the agent attempted ($A_{Attempt}$) (Eq. 6).

$$SR = \frac{A_{GA_w}}{A_{Attempt}} \times 100\% \tag{6}$$

In HIC and RIC experiments, each participants interacted thrice with the proposed method. Thus, it is observed a total of (($12(participants) \times 3(actions) = 36$ interactions in each case. Figure 21 (a) shows *SR* of the proposed method in HIC and RIC. Concerning the success ratio, the result indicates that the proposed system performed better in human initiative case (92%) that than the robot initiative case (86%). The robot sometimes failed to display an appropriate response because of the false detection rate which happened in both cases. In addition, the participants sometimes missed the robot attention attraction action in RIC. Thus, the HIC shows a better result than RIC because there is no other horizontal movement of the robot head after the head-turning action.

We calculate the average time (T_{avg}) from the successful eye contact episodes by investigating experimental videos. The T_{avg} denotes the ratio of total time (in seconds) elapsed for making eye contact in all sessions to the total number of eye contact sessions. A session time counted from the starting to ending of an eye contact event in any case. Figure 21 (b)

illustrates T_{avg} for HIC and RIC. Result indicates that HIC requires in average 13.97 seconds whereas RIC requires a bit higher average time of 23.03 seconds. In RIC, the robot waits delivering its head turning and shaking actions which requires a bit extra time to establish gaze crossing than HIC.

IX. DISCUSSION

Indeed eye contact or gaze contact is the primary contributor for initiating a social interaction and indicates the degree of rapport along with proximity, topic affinity, and amount of positive expression [78]. Gaze contact reveals that the participants are willing to continue any interaction. In order to make meaningful gaze contact, both gaze crossing and gaze awareness are important [12]. In a human-robot interaction scenario, any agent (i.e., robot or human) can start the conversion. Thus, this work focused on developing a duplex eye contact for human-robot intercommunication.

We developed a conceptual model of duplex eye contact method, including both the human initiative case and the robot initiative case. To verify the conceptual model's effectiveness, we have constructed a simple, easy maintaining robotic platform consisting of 04 software modules FDM, GDTM, GAwM, and RRCM. The FDM detects the human's face, GDTM identifies the gaze direction, and GAwM recognizes the smile expression as a gaze awareness cue. Moreover, the robotic framework can turn, shake, and nod its head controlled by the RRCM.

Several cues have extracted by experiments to design the behavioural protocol of the robot. Preliminary experiment 1 performed to extract the robotic head's tilt movement angle. It observed that 20° tilt movement is better than 10° and 30° movements. Experiment 2 revealed that nodding once is more acceptable than nodding twice and thrice for the participant to understand the robot's responsive gaze behaviour. Experiment 3 confirmed that the combined signals such as nodding head with blinking eyes are more useful than only head nodding or eye blinks signal to create a feeling of gaze awareness.

The conceptual model is implemented in our developed robotic framework after extracting the suitable cues for gaze crossing and gaze awareness. Two experiments have been performed concerning two cases (human initiative and robot initiative) to evaluate the proposed duplex eye contact mechanism's performance in a particular scenario (such as "reading a book"). Evaluation results show that the proposed framework performs its functions satisfactory in both cases. However, the proposed system is achieved higher accuracy in HIC (92%) than RIC (86%).

A. SELECTIVE FRIENDLY EYE CONTACT APPROACH

In duplex eye contact, any agent (the human or robot) can grab the attention of its interacting partner first as a prerequisite of eye contact event. We intended to design an HRI framework by which an agent can set up eye contact with an interested agent while avoiding catching other agent's attention as much as possible. Therefore, the interacting agent should reckon the

present state of its partner and employ a suitable cue. This concept is the basic design of our system. The agent should begin with a weaker action to prevent capturing other agent's attention as much as viable besides the target agent and utilize more decisive action only when the light action flops. Based on the past findings of psychology and human-robot interaction researches, we selected the head-turning as the weakest and primary cue to grab someone's attention [55]. The system would use the head-shaking cue as a second attempt if the primary cue failed. Several experiments have affirmed that the proposed system useful to actualize such an HRI framework that can grab a designated agent as selectively as viable.

B. GAZE AWARENESS MODALITY

The previous survey confirmed that establishing gaze crossing alone inadequate to begin any intercommunication event. Gaze awareness also a significant component in making meaningful eye contact event [1], [55]. Blinking, nodding and smiling cues strengthening the feeling of being gazed at, and these cues can be utilized to transmit a feeling much successfully in interpreting human's cognitive or sociable behaviours. In many situations, the nodding cue used to denote acknowledgement or consent. However, only blinking or nodding action may fail to create a stronger feeling of eye contact being established. Compound eye blinks with nodding actions can create a stronger feeling of making eye contact. The previous study performed by Hoque et al. [12] used eye blinks as gaze awareness modality. We have experimented with the three actions: eye blinks only, head-nodding only and eye blinks with nodding. Experimental outcomes have affirmed that eye blinks with nod cues of the robot proven effective in relaying to the human as the gaze awareness modality.

C. FUTURE CHALLENGES

There are various concerns have not covered in the current implementation. Few of these addressed in the below:

1) LIMITED PROXIMITY AND VIEWING ANGLE

The current system functioning well in a limited distant between the human and the robot. The performance of the system degraded while the distance between the camera and the participant increased or even failed for greater proximity due to the constraint of camera focus. Moreover, the current system considered that the interacting parties should be in the central/peripheral field of views. Although these might be true in a few cases, more situations occur where the intended agent stay in out of the field of view. Thus, the system should capture the whole field of view (360°) to interact in any viewing conditions. The success of making eye contact depends on establishing gaze crossing. The robot ensures to establish gaze crossing with the human when detecting his/her frontal face and gaze within its field of view. Thus, the robotic framework successfully establishes the gaze crossing when the human's face angle within ±30°. Mennesson et al., [79]

also showed that the face detection rate considered poor with the face angle > ±30°.

The proposed framework may fail to establish eye contact in several scenarios. For example, the robot's camera cannot receive the frontal face within 30° due to the human's looking response (i.e., face angle) (Fig. 22 (a)). The robot may also fail to develop contact when the human looks at the robot by his/her gaze turning. The system should detect the face first for identifying the gaze. It cannot detect gaze without detecting the face and hence failed to establish gaze crossing. Fig. 22(b) depicts a failure case when the human is looking at the robot by turning gaze.
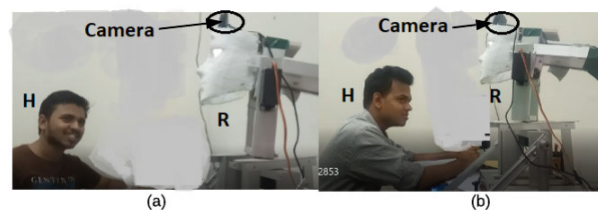

FIGURE 22. Some unsuccessful cases (a) the camera failed to detect H's face within 30° (b) H is looking at the R by turning his gaze alone.

2) GENERALIZABILITY

More studies should be introduced for multi parties scenario in different ages. We evaluated the developed scheme in a particular scenario where only one human can interact while she/he involved a common attention absorption task. Moreover, all the evaluations have performed in controlled environments. Thus, the generalizability of the scheme is limited. The more intense technique need to be employed to make eye contact with the intended human in the multiparty setting. More experiments should be conducted to investigate the dynamics of spectators where humans are involved in high attention absorption activities.

3) CONSTRAINED CUES

The system constrained within the four cues such as heading turn, head shaking, head nod, and eye blinks. Nevertheless, in reality, these cues are not sufficient in all situations. The robot should use other physical cues or voice cue depending on the situations. However, in multiparty scenario, voice cue certainly attracts the other's attention too. Thus, a more subtle technique should explore and design suitable cues based on the situations. As examples, waiving hand, touch or going nearby to the target person can be applied usefully instead of voice.

4) EMBODIMENT

In order to assess the proposed system, this work presented a robotic platform which composes of a head mask with eye blinks. A robot with the full-body embodiment or anthropomorphic appearances should undoubtedly affect the interacting patterns of HRI. Thus, the current framework may be installed in humanoid or social robot to investigate the usability and performance.

5) PERSONAL OR CULTURAL DISCREPANCIES

The personal and cultural factors may affect the human's impression regarding the behaviours of robots. The robots require to cope with their process of eye contact with their partner's traits. Moreover, humans also understand the capability of their robotic partners. In future, it should explore the impacts of robot's behaviour in various cultural or societal contexts. In addition to that, how robots work in multiple languages and communicate to humans concerning various demographic and personality characteristics should also address.

6) TECHNICAL CHALLENGES

Due to the deficiency of computer vision technology, recent humanoid or social robots produce much-restricted interactivity concerning social behaviours and cognitive functionality. The proposed framework can recognize, tracks and understand some responses of one participant (such as smile, and looking at the robot). However, the robot should comprehend more participants responses and behaved accordingly to adopt with the real-world scenarios. Moreover, developing instantaneous or competent interactivity into humanoids will demand interpretation of hybrid cues (a combination of verbal and non-verbal) and their production.

X. CONCLUSION

The principal aim of this research is to develop a duplex eye contact scheme for social robots using nonverbal actions. To fulfil this aim, we have introduced a conceptual model of duplex eye contact process by bearing in mind two scenarios: human initiative and robot initiative. A low-cost and less complicated robotic framework is developed in this work to implement the conceptual model and to verify its effectiveness. To design the behavioural protocol of the proposed framework, several effective cues or actions are extracted from several preliminary experimental studies. Evaluation results with human participators showed the effectiveness of the proposed framework. A significant amount of technical and methodological challenges remained bottlenecks while designing the duplex eye contact framework. Future improvement should include an automatic pan-tilt zoom camera and pan-tilt unit instead of servo motors for smooth operation and better performance. More sophisticated techniques can be used for face/gaze detection. Furthermore, to be more realistic, the system can be extended to work with peripheral or out of the field of viewing situations in the multiparty scenario. The facial expression recognition, full-body embodiment, and other social cues (such as physical touch and hand waving) can be included for further improvements.

REFERENCES

[1] M. Argyle, *Bodily Communication*. New York, NY, USA: International Universities Press, 1990.

[2] T. Striano, V. M. Reid, and S. Hoehl, "Neural mechanisms of joint attention in infancy," *Eur. J. Neurosci.*, vol. 23, no. 10, pp. 2819–2823, May 2006.

[3] T. Farroni, E. M. Mansfield, C. Lai, and M. H. Johnson, "Infants perceiving and acting on the eyes: Tests of an evolutionary hypothesis," *J. Experim. Child Psychol.*, vol. 85, no. 3, pp. 199–212, Jul. 2003.

[4] C. Trevarthen and K. J. Aitken, "Infant intersubjectivity: Research, theory, and clinical applications," *J. Child Psychol. Psychiatry*, vol. 42, no. 1, pp. 3–48, Jan. 2001.

[5] C. Fullwood and G. Doherty-Sneddon, "Effect of gazing at the camera during a video link on recall," *Appl. Ergonom.*, vol. 37, no. 2, pp. 167–175, Mar. 2006.

[6] G. O. Deák, "When and where do infants follow gaze?" in *Proc. Joint IEEE Int. Conf. Develop. Learn. Epigenetic Robot.*, Aug. 2015, pp. 182–187.

[7] K. Kompatsiari, F. Ciardo, D. De Tommaso, and A. Wykowska, "Measuring engagement elicited by eye contact in human-robot interaction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nov. 2019, pp. 6979–6985.

[8] D. Miyauchi, "Bidirectional eye contact for human-robot communication," *IEICE Trans. Inf. Syst.*, vol. E88-D, no. 11, pp. 2509–2516, Nov. 2005.

[9] M. V. Cranach, *The Role of Orienting Behavior in Human Interaction, Behavior and Environment*. New York, NY, USA: Plenum Press, 1971, pp. 217–237.

[10] Y. Yoshikawa, K. Shinozawa, H. Ishiguro, N. Hagita, and T. Miyamoto, "The effects of responsive eye movement and blinking behavior in a communication robot," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2006, pp. 4564–4569.

[11] I. Brinck, L. Heco, K. Sikström, V. Wandsleb, B. Johansson, and C. Balkenius, "Humans perform social movements in response to social robot movements: Motor intention in human-robot interaction," in *Proc. Joint IEEE 10th Int. Conf. Develop. Learn. Epigenetic Robot. (ICDL-EpiRob)*, Oct. 2020, pp. 1–6.

[12] M. M. Hoque, Y. Kobayashi, and Y. Kuno, "A proactive approach of robotic framework for making eye contact with humans," *Adv. Hum.-Comput. Interact.*, vol. 2014, Jul. 2014, Art. no. 694046.

[13] A. Saran, E. S. Short, A. Thomaz, and S. Niekum, "Enhancing robot learning with human social cues," in *Proc. 14th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2019, pp. 745–747.

[14] S. Ali, F. Mehmood, Y. Ayaz, M. J. Khan, H. Sadia, and R. Nawaz, "Comparing the effectiveness of different reinforcement stimuli in a robotic therapy for children with ASD," *IEEE Access*, vol. 8, pp. 13128–13137, 2020.

[15] H.-L. Cao, R. E. Simut, N. Desmet, A. De Beir, G. Van De Perre, B. Vanderborght, and J. Vanderfaeillie, "Robot-assisted joint attention: A comparative study between children with autism spectrum disorder and typically developing children in interaction with NAO," *IEEE Access*, vol. 8, pp. 223325–223334, 2020.

[16] N. Bergstrom, T. Kanda, T. Miyashita, H. Ishiguro, and N. Hagita, "Modeling of natural human-robot encounters," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2008, pp. 2623–2629.

[17] B. Schauerte, J. Richarz, and G. A. Fink, "Saliency-based identification and recognition of pointed-at objects," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2010, pp. 4638–4643.

[18] K. Yamazaki, M. Kawashima, Y. Kuno, N. Akiya, M. Burdelski, A. Yamazaki, and H. Kuzuoka, "Prior-to-request and request behaviors within elderly day care: Implications for developing service robots for use in multiparty settings," in *Proc. ECSCW*. London, U.K.: Springer, 2007, pp. 61–78.

[19] H. Kuzuoka, Y. Suzuki, J. Yamashita, and K. Yamazaki, "Reconfiguring spatial formation arrangement by robot body orientation," in *Proc. 5th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2010, pp. 285–292.

[20] R. Gockley, J. Forlizzi, and R. Simmons, "Natural person-following behavior for social robots," in *Proc. ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*. New York, NY, USA: ACM, 2007, pp. 17–24.

[21] M. A. Yousuf, Y. Kobayashi, Y. Kuno, A. Yamazaki, and K. Yamazaki, "How to move towards visitors: A model for museum guide robots to initiate conversation," in *Proc. IEEE Int. Work. Rob. Hum. Int. Commun. (RO-MAN)*, Aug. 2013, pp. 587–592.

[22] S. Andrist, M. Gleicher, and B. Mutlu, "Looking coordinated: Bidirectional gaze mechanisms for collaborative interaction with virtual characters," in *Proc. CHI Conf. Hum. Factors Comput. Syst.* New York, NY, USA: ACM, May 2017, pp. 2571–2582.

[23] A. M. Sabelli and T. Kanda, "Robovie as a mascot: A qualitative study for long-term presence of robots in a shopping mall," *Int. J. Social Robot.*, vol. 8, no. 2, pp. 211–221, Apr. 2016.

[24] T. Edward, *Hall. The Hidden Dimension: Man's Use of Space in Public and Private*. London, U.K.: Bodley Head, 1966.

[25] E. Goffman, *Behavior in Public Relation: Notes on the Social Organization of Gatherings*. New York, NY, USA: Macmillan, 1963.

[26] K. Adam, "Spatial organization in social encounters: The F-formation system," in *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge, U.K.: Cambridge Univ. Press, 1990, pp. 209–238.

[27] C. Breazeal, *Designing Sociable Robots*. Cambridge, MA, USA: MIT Press, 2004.

[28] H. Kozima and H. Yano, "In search of otogenetic prerequisites for embodied social intelligent," in *Proc. Workshop Emergence Develop. Embodied Cognition, Int. Conf. Cognit. Sci. (IACS)*, 2001, pp. 30–34.

[29] B. M. Scassellati, "Foundations for a theory of mind for a humanoid robot," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, May 2001.

[30] H. Kozima and H. Yano, "A robot that learns to communicate with human caregivers," in *Proc. Int. Workshop Epigenetic Robot.*, C. Balkenius, Ed. Lund, Sweden: Lund University Cognitive Studies, Sep. 2001.

[31] M. Fujita and H. Kitano, "Development of an autonomous quadruped robot for robot entertainment," *Auto. Robots*, vol. 5, no. 1, pp. 7–18, 1998.

[32] T. Tashima, S. Saito, T. Kudo, M. Osumi, and T. Shibata, "Interactive pet robot with an emotion model," *Adv. Robot.*, vol. 13, no. 1, pp. 225–226, Jan. 1999.

[33] J. Pineau, M. Montemerlo, M. Pollack, N. Roy, and S. Thrun, "Towards robotic assistants in nursing homes: Challenges and results," *Robot. Auto. Syst.*, vol. 42, nos. 3–4, pp. 271–281, Mar. 2003.

[34] K. Severinson-Eklundh, A. Green, and H. Hüttenrauch, "Social and collaborative aspects of interaction with a service robot," *Robot. Auto. Syst.*, vol. 42, nos. 3–4, pp. 223–234, Mar. 2003.

[35] A. Maejima, T. Kuratate, B. Pierce, S. Morishima, and G. Cheng, "Automatic face replacement for a humanoid robot with 3D face shape display," in *Proc. 12th IEEE-RAS Int. Conf. Humanoid Robots (Humanoids)*, Nov. 2012, pp. 469–474.

[36] A. Kshirsagar, M. Lim, S. Christian, and G. Hoffman, "Robot gaze behaviors in human-to-robot handovers," *IEEE Robot. Autom. Lett.*, vol. 5, no. 4, pp. 6552–6558, Oct. 2020.

[37] R. Gockley, A. Bruce, J. Forlizzi, M. Michalowski, A. Mundell, S. Rosenthal, B. Sellner, R. Simmons, K. Snipes, A. C. Schultz, and J. Wang, "Designing robots for long-term social interaction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Aug. 2005, pp. 1338–1343.

[38] K. Hayashi, D. Sakamoto, T. Kanda, M. Shiomi, S. Koizumi, H. Ishiguro, T. Ogasawara, and N. Hagita, "Humanoid robots as a passive-social medium-a field experiment at a train station," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interact. (HRI)*, Mar. 2007, pp. 137–144.

[39] S. Li, B. Wrede, and G. Sagerer, "A dialog system for comparative user studies on robot verbal behavior," in *Proc. 15th IEEE Int. Symp. Robot Hum. Interact. Commun. (ROMAN)*, Sep. 2006, pp. 129–134.

[40] M. P. Michalowski, S. Sabanovic, and R. Simmons, "A spatial model of engagement for a social robot," in *Proc. 9th IEEE Int. Workshop Adv. Motion Control*, Mar. 2006, pp. 762–767.

[41] M. Yamamoto and T. Watanabe, "Timing control effects of utterance to communicative actions on embodied interaction with a robot," in *Proc. 13th IEEE Int. Workshop Robot Hum. Interact. Commun. (RO-MAN)*, 2004, pp. 467–472.

[42] K. Dautenhahn, M. Walters, S. Woods, K. L. Koay, C. L. Nehaniv, A. Sisbot, R. Alami, and T. Siméon, "How may i serve you? A robot companion approaching a seated person in a helping context," in *Proc. 1st ACM SIGCHI/SIGART Conf. Hum.-Robot Interact.* New York, NY, USA: Association for Computing Machinery, 2006, pp. 172–179.

[43] M. Buss, D. Carton, B. Gonsior, K. Kuehnlenz, C. Landsiedel, N. Mitsou, R. de Nijs, J. Zlotowski, S. Sosnowski, E. Strasser, and M. Tscheligi, "Towards proactive human-robot interaction in human environments," in *Proc. 2nd Int. Conf. Cognit. Infocommun. (CogInfoCom)*, 2011, pp. 1–6.

[44] A. Cesta, G. Cortellessa, M. V. Giuliani, F. Pecora, M. Scopelliti, and L. Tiberio, "Psychological implications of domestic assistive technology for the elderly," *Psychol. J.*, vol. 5, no. 3, pp. 229–252, 2007.

[45] M. M. Hoque, T. Onuki, E. Tsuburaya, Y. Kobayashi, Y. Kuno, T. Sato, and S. Kodama, "An empirical framework to control human attention by robot," in *Proc. Asian Conf. Comput. Vis.* Berlin, Germany: Springer, 2010, pp. 430–439.

[46] C. Shi, M. Shimada, T. Kanda, H. Ishiguro, and N. Hagita, "Spatial formation model for initiating conversation," in *Proc. Robot., Sci. Syst.*, Los Angeles, CA, USA, Jun. 2011, pp. 305–313.

[47] C. Sirithunge, A. G. B. P. Jayasekara, and D. P. Chandima, "Proactive robots with the perception of nonverbal human behavior: A review," *IEEE Access*, vol. 7, pp. 77308–77327, 2019.

[48] S. Satake, T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, and N. Hagita, "How to approach humans?: Strategies for social robots to initiate interaction," in *Proc. 4th ACM/IEEE Int. Conf. Human Robot Interact. (HRI)*, Mar. 2009, pp. 109–116.

[49] S. Satake, T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, and N. Hagita, "A robot that approaches pedestrians," *IEEE Trans. Robot.*, vol. 29, no. 2, pp. 508–524, Apr. 2013.

[50] N. Mitsunaga, T. Miyashita, H. Ishiguro, K. Kogure, and N. Hagita, "Robovie-IV: A communication robot interacting with people daily in an office," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2006, pp. 5066–5072.

[51] A. Kendon, *Features of the Structural Analysis of Human Communicational Behavior: Aspects of Nonverbal Communication*. Holland, The Netherlands: Swets Zeitlinger, 1980.

[52] R. A. Brooks, C. Breazeal, M. Marjanovi, B. Scassellati, and M. M. Williamson, *The Cog Project: Building a Humanoid Robot*, vol. 1562. Berlin, Germany: Springer, Mar. 1998, pp. 52–87.

[53] T. Kanda, H. Ishiguro, M. Imai, and T. Ono, "Development and evaluation of interactive humanoid robots," *Proc. IEEE*, vol. 92, no. 11, pp. 1839–1850, Nov. 2004.

[54] Y. Matsusaka, T. Tojo, S. Kubota, K. Furukawa, D. Tamiya, K. Hayata, Y. Nakano, and T. Kobayashi, "Multi-person conversation via multi-modal interface—A robot who communicate with multi-user," in *Proc. Eur. Conf. Speech Commun. Technol.* Shanghai, China: ISCA Archive, Sep. 1999, pp. 1723–1726.

[55] B. Mutlu, T. Kanda, J. Forlizzi, J. Hodgins, and H. Ishiguro, "Conversational gaze mechanisms for humanlike robots," *ACM Trans. Interact. Intell. Syst.*, vol. 1, no. 2, pp. 1–33, Jan. 2012.

[56] T. Yonezawa, H. Yamazoe, A. Utsumi, and S. Abe, "Gaze-communicative behavior of stuffed-toy robot with joint attention and eye contact based on ambient gaze-tracking," in *Proc. 9th Int. Conf. Multimodal Interfaces (ICMI)*. New York, NY, USA: ACM, 2007, pp. 140–145.

[57] M. M. Hoque, T. Onuki, Y. Kobayashi, and Y. Kuno, "Effect of robot's gaze behaviors for attracting and controlling human attention," *Adv. Robot.*, vol. 27, no. 11, pp. 813–829, Aug. 2013.

[58] M. M. Hoque, D. Das, Y. Kobayashi, and Y. Kuno, "Design an intelligent robotic head to interacting with humans," in *Proc. 15th Int. Conf. Comput. Inf. Technol. (ICCIT)*, Dec. 2012, pp. 539–545.

[59] M. M. Hoque, Q. D. Hossian, K. Deb, D. Das, Y. Kobayashi, and Y. Kuno, "An empirical robotic framework for initiating interaction with the target human in multiparty settings," *J. Circuits, Syst. Comput.*, vol. 24, no. 2, Feb. 2015, Art. no. 1540003.

[60] C.-M. Huang and A. Thomaz, "Joint attention in human-robot interaction," in *Proc. AAAI Fall Symp., Dialog Robots*, Jan. 2010, pp. 32–37.

[61] C.-M. Huang and A. L. Thomaz, "Effects of responding to, initiating and ensuring joint attention in human-robot interaction," in *Proc. RO-MAN*, Jul. 2011, pp. 65–71.

[62] G. Mehlmann, M. Häring, K. Janowski, T. Baur, P. Gebhard, and E. André, "Exploring a model of gaze for grounding in multimodal HRI," in *Proc. 16th Int. Conf. Multimodal Interact.*, Nov. 2014, pp. 247–254.

[63] A. S. Phyo, H. Fukuda, A. Lam, Y. Kobayashi, and Y. Kuno, "A human-robot interaction system based on calling hand gestures," in *Proc. Int. Conf. Intell. Comput.* Cham, Switzerland: Springer, 2019, pp. 43–52.

[64] D. Maurer, "Infants' perception of facedness," in *Social Perception in Infants*, vol. 8, T. M. Field and N. A. Fox, Eds. Norwood, NJ, USA: Ablex, May 1985, pp. 73–100.

[65] M. M. Hoque, D. Das, T. Onuki, Y. Kobayashi, and Y. Kuno, "Robotic system controlling target human's attention," in *Proc. Intell. Comput. Theories Appl. (ICIC)* (Lecture Notes in Computer Science), vol. 7390. Berlin, Germany: Springer, 2012.

[66] L. M. Pönkänen and J. K. Hietanen, "Eye contact with neutral and smiling faces: Effects on autonomic responses and frontal EEG asymmetry," *Frontiers Hum. Neurosci.*, vol. 6, no. 122, pp. 201–213, May 2012.

[67] T. Kanda, H. Ishiguro, T. Ono, M. Imai, and R. Nakatsui, "Development and evaluation of an interactive humanoid robot robovie," in *Proc. IEEE Int. Conf. Robot. Autom.*, vol. 2, Feb. 2002, pp. 1848–1855.

[68] A. Fridlund, "The new ethology of human facial expressions," in *The Psychology of Facial Expression*, J. Russell and J. Fernández-Dols, Eds. Cambridge, U.K.: Cambridge Univ. Press, Mar. 1997, pp. 103–130.

[69] J. E. Hanna and S. E. Brennan, "Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation," *J. Memory Lang.*, vol. 57, no. 4, pp. 596–615, Nov. 2007.

[70] H. H. Clark, *Using Language*. Cambridge, U.K.: Cambridge Univ. Press, 1996.

[71] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Dec. 2001, p. 1.

[72] W. K. Pratt, *Digital Image Processing*, 3rd ed. Hoboken, NJ, USA: Wiley, 2001, ch. 8, pp. 61–86.

[73] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.

[74] E. R. Davies, "A modified Hough scheme for general circle location," *Pattern Recognit. Lett.*, vol. 7, no. 1, pp. 37–43, Jan. 1988.

[75] O. Déniz, M. Castrillon, J. Lorenzo, L. Anton, and G. Bueno, "Smile detection for user interfaces," in *Proc. Int. Symp. Vis. Comput.* Berlin, Germany: Springer, Dec. 2008, pp. 602–611.

[76] C. L. Sidner, C. Lee, L.-P. Morency, and C. Forlines, "The effect of head-nod recognition in human-robot conversation," in *Proc. 1st ACM SIGCHI/SIGART Conf. Hum.-Robot Interact. (HRI)*, New York, NY, USA, Jan. 2006, pp. 290–296.

[77] T. Sano, A. Yuguchi, G. A. G. Ricardez, J. Takamatsu, A. Nakazawa, and T. Ogasawara, "Evaluating imitation of human eye contact and blinking behavior using an Android for human-like communication," in *Proc. 28th IEEE Int. Conf. Robot Hum. Interact. Commun. (RO-MAN)*, Oct. 2019, pp. 1–6.

[78] M. Argyle and J. Dean, "Eye-contact, distance and affiliation," *Sociometry*, vol. 28, pp. 289–304, Sep. 1966.

[79] J. Mennesson, A. Dahmane, T. Danisman, and I. M. Bilasco, "Head yaw estimation using frontal face detector," in *Proc. 11th Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.* Setúbal, Portugal: SciTePress, Jan. 2016, pp. 517–524.

SHAYLA SHARMIN (Graduate Student Member, IEEE) received the B.Sc. degree in engineering and the M.Sc. degree in computer science and engineering from the Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh, in 2014 and 2020, respectively. She is currently working as an Assistant Professor with the Department of Computer Science and Engineering, CUET. Her research interests include image processing, human robot/computer interaction, computer vision, and artificial intelligence. She is a member of IEB, Bangladesh.

MOHAMMED MOSHIUL HOQUE (Senior Member, IEEE) received the Ph.D. degree from the Department of Information and Computer Sciences, Saitama University, Japan, in 2012. He was worked as the Head of the Department of CSE, CUET, and the Director of Student's Welfare, CUET. He is currently a Distinguish Professor with the Department of Computer Science and Engineering (CSE), CUET. He is also working as the Director of Sheikh Kamal IT Business Incubator with CUET. He is the Founding Director of CUET NLP Lab and a Coordinator of CUET Fab Lab. He published more than 140 publications in several international journals, and conferences. His research work awarded in several international conferences, such as HIS 2011, EICT 2013, IEMIS 2018, ICO 2019, ICIoTCT 2020, and CONSTRAINT-AAAI 2021. One of his work received second prize in a national level completion at Solve-a-Thon Category. His research interests include human robot/computer interaction, artificial intelligence, machine learning, and natural language processing. He is a Senior Member of the IEEE Computer Society, the IEEE Robotics and Automation Society, the IEEE Women in Engineering, the IEEE Signal Processing Society, USA, and the Fellow of Institute of Engineers, Bangladesh. He actively engaged in several activities of IEEE as an IEEE volunteer and served several positions of IEEE Bangladesh Section (BDS) and Society Chapters as an Excom. He served as the IEEE BDS Vice-Chair (technical), in 2019 and 2020, the Conference Coordinator, in 2018, the Award Coordinator, in 2017, the Award Coordinator and Vice-chair (Activity) of IEEE CS Bangladesh Chapter, and the Educational Activity Coordinator of IEEE RAS Bangladesh Chapter, from 2017 to 2020. From 2019 to 2021, he served as the Advisor of the IEEE CUET SB, RAS/SPS CUET Chapter and WIE Affinity Group. He served as the TPC Chair, a TPC Co-Chair, a Publication Chair, and a TPC members in several international conferences.

S. M. RIAZUL ISLAM (Member, IEEE) was as an Assistant Professor with the University of Dhaka, Bangladesh, and a Lecturer with the Department of Electrical and Electronic Engineering, from 2005 to 2014. He worked as a Chief Engineer with the Department of Solution Laboratory for Advanced Research, Samsung Research and Development Institute, Bangladesh, in 2014. From 2014 to 2017, he worked as a Postdoctoral Fellow with the Wireless Communications Research Center, Inha University, South Korea. From 2016 to 2017, he was also affiliated with Memorial University, Canada. Since March 2017, he has been working as an Assistant Professor with the Department of Computer Science and Engineering, Sejong University, South Korea. His research interests include wireless communications, Internet of Things, and applied artificial intelligence.

MD. FAZLUL KADER (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in computer science and engineering from the Chittagong University of Engineering and Technology, Chittagong, Bangladesh, in November 2005 and January 2014, respectively, and the Ph.D. degree from the Kumoh National Institute of Technology, Gumi, South Korea, in February 2018. Since 2007, he has been a Faculty Member with the Department of Electrical and Electronic Engineering, University of Chittagong, Chittagong, where he is currently an Associate Professor. His major research interests include 5G, cognitive radio networks, cooperative communications, MIMO, computer networks, NOMA, spatial modulation, and so on. He is an Associate Editor of IEEE Access. He regularly serves as a Reviewer and a TPC member in many reputed journals and conferences.

IQBAL H. SARKER (Member, IEEE) received the Ph.D. degree from the Department of Computer Science and Software Engineering, Swinburne University of Technology, Melbourne, VIC, Australia, in 2018. He is currently working as a Faculty Member with the Department of Computer Science and Engineering, Chittagong University of Engineering and Technology, Bangladesh. He has published a number of peer-reviewed journals and conferences in top venues, such as the *Journal of Network and Computer Applications* (Elsevier, USA), *Internet of Things* (Elsevier), the *Journal of Big Data* (Springer Nature, U.K.), *Mobile Network and Applications* (Springer, Netherlands), *Sensors* (MDPI, Switzerland), *AI & Society* (Springer, Germany), *Computer Journal*, (Oxford University, U.K.), and conferences, such as the IEEE DSAA-Canada, IEEE Percom-Greece, ACM Ubicomp-USA, and Germany; ACM Mobiquitous-Australia, and Springer PAKDD–Australia. His research interests include data science, AI and machine learning, cybersecurity data analytics, cyber threat intelligence, context-aware smart computing, and IoT analytics. He is a member of ACM, and one of the research founders of International AIQT Foundation, Switzerland.

● ● ●