System and Method of Facial and Body Gesture-Based Interaction

FIELD OF INVENTION

The present invention relates to a facial and body gesture-based interactive system. In particular, the present invention provides an interactive system and method that automatically recognises faces and body gestures to allow for a continuous narrative that evolves with each facial expression or body gesture made by a user.

BACKGROUND ART

A central core of a human being is having the ability to communicate and express emotions which is needed in performing daily activities. Humans express emotions through speech and action. A high percentage of communication in humans is non-verbal. Hence, facial expression and body gestures play an important role in human interaction.

In recent years, analysis of facial and body gestures by recognition has become a topic of interest. The automation of the process of facial recognition is an area that has seen large progress in the past few years. Facial and body gesture recognition is an application that provides benefit in various fields such as healthcare, entertainment, advertisement communication, education, robotics and computing.

Detection and analysis of facial expression coupled with body gestures to provide an interactive experience continue to be a complex and challenging field of computer science and design. This is due to the complexity of the face itself as detection of the human face involves multiple conditions such as face gestures, poses and lighting conditions. Facial and body gesture recognition is a computer-based technology that uses mathematical algorithms to analyse faces and bodies in images, videos or in a direct interaction with a person.

Conventionally, advertisements display still images or a pre-recorded video to draw customers to an advertised product. In earlier times, brochures and leaflets are a common sight in shopping malls and retail outlets to market a certain product.

A drawback in using conventional advertisements is that there is lack of an interactive experience according to the user's emotion, facial expression and body gestures. Providing an interactive experience to a user or a customer is a powerful marketing tool that has potential in drawing customers or users to the product that is being advertised.

Numerous interactive systems have been proposed that involve facial and body gesture recognition.

One example of an interactive system is disclosed in China Patent Publication No. CN 101393599 B (hereinafter referred to as CN 599 B) entitled 'Game role control method based on human face expression' having a filing date of 19 September 2007, Applicant: Chinese ACAD Inst Automation. CN 599 B discloses a method for controlling game characters based on analyzing the human face. CN 599 B discloses methods including processing an image, detecting the human face and positioning key points, analyzing the positioning results to obtain human face gesture and expression information and mapping information to game control commands to achieve real-time control function of the faces of the game characters. The method of CN 599 B enables game players to perform an interactive operation that allows for a more interactive gaming experience.

Another example of an interactive system is disclosed in United States Patent Publication No. US 2014/0365333 A1 (hereinafter referred to as US 333 A1) entitled 'Retail Store Customer Natural-Gesture Interaction with Animated 3D Images Using Sensor Array' having a filing date of 7 June 2013, Applicant: Matthew Hurewitz. US 333 A1 discloses a system for selling products in a physical retail store that comprises a virtual interactive product display. The interactive product display comprises sensors that provide gesture recognition, face recognition and gesture analysis. This feature of US 333 A1 allows for retail store customers to view and interact with 3D images of products for sale. US 333 A1 further discloses a feature that allows the user to self-identify in order to save pre-selected products to view at the interactive display.

A further example of an interactive system is disclosed in International Patent Publication No. WO 2018/140397 A1 (hereinafter referred to as WO 397 A1) entitled 'System for Interactive Image Based Game' having a filing date of 23 January 2018, Applicant: Odile Aimee Furment et al. WO 397 A1 discloses a method for real time processing of images providing a real time preview of face images for interaction with a game environment.

The image sensors present in the invention of WO 397 A1 process images providing modified images of a certain size and shape that detect and track faces for a real time preview of face images overlaid on a game environment. The facial gestures allow for execution of game commands allowing control of the real time preview of the displayed face images for interaction with the game environment. WO 397 A1 also discloses a feature to allow second face images for interaction with the game environment.

As outlined above, various interactive systems have been developed which involve facial and gesture recognition to create an interactive experience between a user and the digital environment.

It is an object of the present invention to provide a facial and body gesture recognition system which facilitates complex interactions with a user.

SUMMARY OF INVENTION

The present invention relates to a system and method of facial and body gesture-based interaction. In particular, the present invention provides an interactive system and method that recognises faces and body gestures to allow for a continuous narrative that evolves with each facial expression or body gesture made by a user.

One aspect of the present invention provides a system of facial and body gesture-based interaction for recognising faces and body gestures to allow for a continuous narrative that evolves with each user facial expression or body gesture. The system of facial and body gesture based interaction comprises a middleware platform that links facial vectors or body data for playing a series of branching videos, a display screen for projecting the series of branching videos from the middleware platform, a microcomputer having an algorithm that works in communication with the middleware platform by supplying electronic triggers to the middleware platform, a hardware having an inbuilt artificial intelligence mechanism for detecting facial and body vectors to trigger signals to the middleware platform, a computer vision hardware camera for receiving images and video signals of the user's facial expression or body gesture and further analysing and running the images and video signals through a series of algorithms for input to the middleware platform and a cloud database for storing information of vectors on body pose estimation as reference data to match with a specific branching video once images and video signals of the user are received. The middleware platform that links facial vectors or body data for playing a series of branching videos comprises a video server where branching videos, audios or graphics are arranged to be displayed in sequence when electronic triggers are received by the middleware platform and a content management module for executing machine readable instructions to structure digital assets based on a plurality of artificial intelligence vectors that involve face and body detection.

Yet another aspect of the present invention provides that the plurality of artificial intelligence vectors of the content management module further cover at least 106 facial points that involve facial expressions, attributes, actions and recognition.

Another aspect of the present invention provides that the plurality of artificial intelligence vectors of the content management module further cover at least 14 body points that involve body pose detections, distances, body structure attributes and human attire.

Yet another aspect of the present invention provides that the digital assets are videos, images or motion graphics.

Another aspect of the present invention provides that the middleware platform that links facial vectors for playing a series of branching videos further comprises that the facial vectors involve smiling in order to play a series of branching videos where the user is able to obtain a reward, play or vote.

Yet another aspect of the present invention provides that the middleware platform that links facial vectors for playing a series of branching videos further comprises that the middleware platform works in communication with a gantry system to allow for opening and closing of a gate upon verification of facial vectors of the system.

Another aspect of the present invention provides that the hardware having an inbuilt artificial intelligence mechanism for detecting facial and body vectors to trigger signals to the middleware platform further comprises machine readable instructions selected from SMRT, UOB or WINK.

Another aspect of the present invention provides a system of facial and body gesture-based interaction for recognising faces and gestures to allow for a continuous narrative that evolves with each user facial expression or body gesture, the system comprising a middleware platform that links facial vectors or body data for interacting with a 3D avatar or other such intelligent objects; a display screen for projecting the 3D avatar from the middleware platform; a microcomputer having an algorithm that works in communication with the middleware platform by supplying electronic triggers to the middleware platform; a hardware having an inbuilt artificial intelligence mechanism for detecting facial and body vectors to trigger signals to the middleware platform; a computer vision hardware camera for receiving images and video signals of the user's facial expression or body gesture and further analysing and running the images and video signals through a series of algorithms for input to the middleware platform; and a cloud database (109) for storing information of vectors on body pose estimation as reference data to match with a specific 3D avatar once images and video signals of the user are received. The middleware platform that links facial vectors or body data for interacting with a 3D avatar or other such intelligent objects comprises a video server where the 3D avatar or intelligent objects are arranged to interact with the user when electronic triggers are

received by the middleware platform and a content management module for executing machine readable instructions to structure digital assets based on a plurality of artificial intelligence vectors that involve face and body detection.

A further aspect of the present invention provides a system for automated facial and body gesture-based interaction (100). The system comprising a middleware platform (102); an imaging device for capturing images of a user's facial expressions and/or body gestures, and processing those images; a processor including an AI mechanism to detect facial and body vectors from the imaging device output to trigger signals to the middleware platform; the system being further adapted to communicate with a cloud database (109) which stores information of vectors on body pose estimation as reference data; wherein the middleware platform determines from the inputs a meaning for at least some facial expressions and/or gestures, and causes an output to be generated responsive to that meaning.

Yet another aspect of the present invention provides a method for recognising faces and gestures to allow for a continuous narrative that evolves with each user facial expression or body gesture made by a user in a system of facial and body gesture-based interaction. The method comprises steps of initiating an interactive system flow where data from a camera frame is filtered and analysed, receiving an image byte array data as detected by the camera frame, converting the image byte array data to a YUV data format, analysing the YUV data format. The method further comprises detecting if a human is present, if a human is detected, analysing of facial and body attributes are determined by tracking semantic key points, if there is no human detected, routing of interactive system flow to initiating the interactive system flow by receiving a camera frame. The method further comprises matching a plurality of facial and body attributes to a branching video, if there is a match, playing a specific branching video, if there is no match, routing of interactive system flow to initiating the interactive system flow by receiving a camera frame. Further, the method comprises playing a further specific sequence of branching video, if there is further interaction, playing a sequence of branching video before ending the session, if there is no further interaction, routing of interactive system flow to initiating the interactive system flow by receiving a camera frame.

Yet another aspect of the present invention provides that analysing of facial and body attributes are determined by tracking semantic key points which further comprise tracking semantic key points by estimating and analyzing body parts such as a bottom torso, left hip, left knee, left foot, right hip, right knee, right foot, centre torso, upper torso, neck base, centre head, right shoulder, right elbow, right hand, left shoulder, left elbow and left hand and further running a set of data on the semantic key points through a series of algorithms and further match the set of data to stored parameters of within the interactive system.

The present invention consists of features and a combination of parts hereinafter fully described and illustrated in the accompanying drawings, it being understood that various changes in the details may be made without departing from the scope of the invention or sacrificing any of the advantages of the present invention.

BRIEF DESCRIPTION OF ACCOMPANYING DRAWINGS

To further clarify various aspects of some embodiments of the present invention, a more particular description of the invention will be rendered by references to specific embodiments thereof, which are illustrated in the appended drawings. It is appreciated that these drawings depict only typical embodiments of the invention and are therefore not to be considered limiting of its scope. The invention will be described and explained with additional specificity and detail through the accompanying drawings in which:

Figure 1.0a illustrates an architecture of the general overview of a facial and body gesture interactive system of the present invention.

Figure 1.1a illustrates an architecture of the general overview of a facial and body gesture interactive system for interacting with a 3D avatar of the present invention.

Figure 1.0b illustrates a plurality of modules of the middleware platform of the facial and gesture interactive system of the present invention.

Figure 2.0 is a flowchart illustrating a method for recognising faces and body gestures of the present invention.

Figure 3.0a illustrates key point location of a human body.

Figure 3.0b illustrates a formula for determining key point location of the human body.

Figure 3.0c illustrates examples of preference data for human body pose.

Figure 3.0d is a flowchart illustrating body pose detection.

Figure 3.0e illustrates 106 facial points that involve facial expressions, attributes, actions and recognition.

Figure 4.0 illustrates an example of gesture detection for a campaign theme ‘Love it, Like It & It’s Okay’ using key points.

Figure 5.0 is a flowchart illustrating an example of a branching sequence of the present invention.

Figure 6.0 is an example of a physical setup for a published campaign application.

DETAILED DESCRIPTION OF THE DRAWINGS

The present invention relates to a system and method of facial and body gesture-based interaction. In particular, the present invention provides an interactive system and method that recognises faces and body gestures to allow for a continuous narrative that evolves with each facial expression or body gesture made by a user. Hereinafter, this specification will describe the present invention according to the preferred embodiments. It is to be understood that limiting the description to the preferred embodiments of the invention is merely to facilitate discussion of the present invention and it is envisioned without departing from the scope of the appended claims.

The present invention provides a facial and body gesture interactive system that creates an interactive experience between a user and the digital environment. In one application, the facial and gesture interactive system of the present invention provides for an interactive way of advertising in the advertising industry. The present invention redefines advertising and engages users or customers for marketing and branding campaigns. The facial and body gesture recognition interactive system of the present invention is provided by a middleware platform for playing a series of branching videos. These branching videos will lead to a continuous narrative that evolves with each facial expression or body gesture when a customer or user interacts with the system.

The interactive system of the present invention, in another implementation, allows a user to play games, redeem vouchers and vote by various facial expressions or body gestures. An example of a facial expression is smiling which can be further analysed by the strength of the smile. An interactive video from the database of the present invention will be matched according to the strength of a user's smile and the relevant video will then be played. For gestures in an interactive system, angles of a body part are further analysed such as angles to a hand gesture to trigger playing of a relevant branching video.

The interactive system of the present invention is a principal application that can be used in various interactive models. Various examples of the interactive models are described as follows.

Example 1

Interactive model where a user smiles for rewards

This interactive model enables the user to earn points by smiling at a display screen. By detecting facial attributes such as a smile, branching videos will be played as an interactive content. At the end of the interactive content, a QR code is displayed for a user to scan which will provide the user with details on various loyalty programmes.

Example 2

Interactive model where a user smiles to play

This interactive model uses facial and gesture recognition to engage a user to play games. An example is a maze arcade where a user controls an avatar of the maze arcade with body gestures. Another example is attractiveness scores can be displayed using facial recognition when a user interacts with the system.

Example 3

Interactive model where a user smiles to vote

This interactive model uses facial and gesture recognition to allow a user to vote. A user is able to vote for their favourite image or video on the display screen by smiling. The number of smiles collected can be tracked. This provides useful and measurable data to advertisers in real-time.

Example 4

Interactive model for opening and closing of gates based on facial recognition

This interactive model uses facial recognition to manage entry of people applicable in offices, shopping centres or other buildings. Some features included in the interactive model are temperature screening and mask detection for initiating opening and closing of the gate.

With the interactive system of the present invention, demographic reports and daily reports are generated. These data are important to advertisers where with the data provided, advertisers will be able to analyse their users and modify their products according to users' preference. As data is collected from facial and gestures displayed by a user, data is continuously being stored in the system which will be analysed for generation of reports allowing a user to review campaign results in real-time. A productions system is also available for advertisers to upload and manage content

remotely. Data analytics of the system of the present invention covers total traffic count of users, total views, duration of view, unique identification (ID) gender, age range, mask wearing and emotions.

Reference is first made to Figure 1.0a. Figure 1.0a illustrates an architecture of a general overview of a facial and gesture interactive system (100) of the present invention. The facial and gesture interactive system (100) recognises faces and body gestures to allow for a continuous narrative that evolves with each user facial expression or body gesture. The system comprises a middleware platform (102), a display screen (105), a microcomputer (106), a hardware having an inbuilt artificial intelligence mechanism (107), a computer vision hardware camera (108) and a cloud database (109).

The middleware platform (102) links facial vectors or body data for playing a series of branching videos. The facial vectors involve smiling in order to play a series of branching videos where the user is able to obtain a reward, play or vote. Besides that, the middleware platform works in communication with a gantry system (110) to allow for opening and closing of a gate upon verification of facial vectors of the system. The display screen (105) projects the series of branching videos from the middleware platform (102). The microcomputer (106) which is a controller board has an algorithm that works in communication with the middleware platform (102) by supplying electronic triggers to the middleware platform (102). The microcomputer is preferably an RK 3399 microcomputer. However, the microcomputer is not limited to the RK 3399 processor and can include other processors that are low in power but high in performance type processor. The hardware having an inbuilt artificial intelligence mechanism (107) detects facial and body vectors to trigger signals to the middleware platform. The computer vision hardware camera (108) receives images and video signals of the user's facial expression or body gesture and further analyses and runs the images and video signals through a series of algorithms for input to the middleware platform. The cloud database (109) of the present invention stores information of vectors on body pose estimation as reference data to match with a specific branching video once images and video signals of the user is received.

In describing further on the middleware platform (102), the middleware platform (102) comprises a video server (104) and a content management module (103). The video server is where branching videos, audios or graphics are arranged to be displayed in

sequence when electronic triggers are received by the middleware platform (102) whereas the content management module (103) executes machine readable instructions to structure digital assets based on a plurality of artificial intelligence vectors that involve face and body detection.

In describing another embodiment of the present invention with respect to Figure 1.0a, the present invention provides a system for automated facial and body gesture-based interaction (100). The system comprising a middleware platform (102); an imaging device for capturing images of a user's facial expressions and/or body gestures, and processing those images; a processor including an AI mechanism to detect facial and body vectors from the imaging device output to trigger signals to the middleware platform. The system being further adapted to communicate with a cloud database (109) which stores information of vectors on body pose estimation as reference data; wherein the middleware platform determines from the inputs a meaning for at least some facial expressions and/or gestures, and causes an output to be generated responsive to that meaning.

In describing a further embodiment of the present invention, reference is made to Figure 1.1a. Figure 1.1a describes a system of facial and body gesture based interaction (1100) for recognising faces and gestures to allow for a continuous narrative that evolves with each user facial expression or body gesture comprising a middleware platform (1102) that links facial vectors or body data for interacting with a 3D avatar or other such intelligent objects (1101). The system comprises a display screen (1105) for projecting the 3D avatar from the middleware platform, a microcomputer (1106) having an algorithm that works in communication with the middleware platform by supplying electronic triggers to the middleware platform, a hardware having an inbuilt artificial intelligence mechanism (1107) for detecting facial and body vectors to trigger signals to the middleware platform, a computer vision hardware camera (1108) for receiving images and video signals of the user's facial expression or body gesture and further analysing and running the images and video signals through a series of algorithms for input to the middleware platform and a cloud database (1109) for storing information of vectors on body pose estimation as reference data to match with a specific 3D avatar once images and video signals of the user are received. The middleware platform (1102) that links facial vectors or body data for interacting with a 3D avatar or other such intelligent objects comprises a video server where the 3D avatar or intelligent objects are

arranged to interact with the user when electronic triggers are received by the middleware platform and a content management module (1103) for executing machine readable instructions to structure digital assets based on a plurality of artificial intelligence vectors that involve face and body detection.

Figure 1.0b illustrates a plurality of modules of the middleware platform of the facial and gesture interactive system. There are three main modules in the middleware platform which are the branching sequence, artificial intelligence and the digital assets.

A branching sequence is defined as a flow of advertisement which initiates the interactive system to play a specific video or sequence of videos in a particular order based on logic, conditions and algorithms that are specific to the campaigns. The input of a user, particularly facial expressions and gesture of the user will determine the flow of branching sequence. The working principle of the branching sequence is further described in Figure 5.0. Figure 5.0 is a flowchart illustrating an example of a branching sequence of the present invention. When the interactive system receives data on a human being detected, the data will be analysed and further categorised in order to play the most suited branching sequence video according to a user's facial expression or body gestures in real time. An example of categorisation is to detect gender of the user. Depending on the gender, a specific branching video sequence will be played. A sequence of branching videos will then be played progressively as more information based on the user's profile such as the age range and facial expression such as smiling is introduced into the system in real time. The sequence of videos that form a continuous narrative will be played as a user interacts with the interactive system of the present invention.

An artificial intelligence mechanism (107) captures and processes image feed from the cameras. The artificial intelligence mechanism is selected from SMRT, UOB or WINK. This mechanism of the interactive system will cause data to run against the interactive system's stored parameters which will be used in determining the branching sequence that is to be selected. The artificial intelligence mechanism enables data to be generated into reports. The artificial intelligence mechanism produces artificial intelligence vectors from data received by face and body detection. A content management module executes machine readable instruction based on the artificial intelligence vectors which translate data received on face and body detection. There are at least 106 facial points that involve facial expressions, attributes, actions and recognition as illustrated in Figure 3.0e

and at least 14 body points that involves body pose detections, distances, body structure attributes and human attire attributes as illustrated in Figure 3.0a.

Digital assets are significant modules that make up the middleware platform. These digital assets come in videos, images and motion graphic forms.

Reference is further made to Figure 2.0. Figure 2.0 is a flowchart illustrating a method for recognising faces and gestures to allow for a continuous narrative that evolves with each user facial expression or body gesture made by a user in a system of facial and body gesture-based interaction. The method (200) starts (202) by initiating an interactive system flow where data from a camera frame (204) is filtered and analysed. An image byte array data as detected by the camera frame is then received (206) and the image byte array data is converted to a YUV data format (208). The interactive system then analyses the data in YUV data format (210). When a user or customer appears in front of the interactive system, the interactive system will be prompted to detect if a human is present (212). If a human is detected, the interactive system analysing of a plurality of facial and body attributes are determined by tracking semantic key points (214). If there is no human detected, the interactive system flow will be routed to initiating the interactive system flow by receiving a camera frame (204). When a human is detected and the facial and body attributes are analysed, the plurality of facial and body attributes will be matched to a branching video (216). If there is a match, a specific branching video will be played (218). If there is no match, the interactive system flow will be routed to initiating the interactive system flow by receiving a camera frame (204). From playing a specific branching video, a sequence of branching videos will be triggered where the interactive system will further play specific sequence of branching video (220). If there is further interaction, the sequence of branching video will be played before ending the session (222). Once completed the branching video ends, the interactive system flow will end (224). If there is no further interaction, the interactive system flow will be routed to initiating the interactive system flow by receiving a camera frame (204).

In describing further on key point location of the human body and analysing of facial and body attributes determined by tracking semantic key points, reference is made to Figure 3.0a. Figure 3.0a illustrates key point location of a human body. Examples of key point locations that are identified by the interactive system of the present invention are the bottom torso, left hip, left knee, left foot, right hip, right knee, right foot, centre torso, upper torso, neck base, centre head, right shoulder, right elbow, right hand, left shoulder,

left elbow and left hand. Tracking of semantic key points takes place when an AI algorithm tracks movements of the user based on the key point location. The data on semantic key points will be run through a series of algorithms and further matched to stored parameters within the interactive system. Digital videos having motion in pictures will be converted into numbers which become data for the interactive system. Data that has been received in the system will be stored.

A formula as shown in Figure 3.0b is used to determine key point location of the human body. As shown on Figure 3.0b, combination of different key point location of the human body will form a pose. These poses will then be translated to detection scores based on the formula. Data on the detection scores of poses will be received and analysed by the system before matching cross matching the body poses against stored parameters in the interactive system. Examples of data stored for human body pose are shown in Figure 3.0c.

A flowchart illustrating body pose detection is shown in Figure 3.0d. This is a general flow of body pose detection. The body pose detection flow starts by determining a list of preferred body poses. If a body pose is detected, the preferred body pose pointers are then extracted. The angles of detected body pose points are then determined. If no preferred body pose is detected, the system will tabulate preferred body pose scores and the highest score will be determined. From the highest preferred body pose score, the body pose can be detected and the system will end.

Another example of a campaign produced using the present invention is a gesture detection campaign known as ‘Love it, Like It & It’s Okay’ as shown in Figure 4.0. The examples show three separate categories based on a range of angles. Depending on the angle of a body gesture of a user, data will be collected in the system on preference of a user or customer based on the categories ‘Love it’ which is defined as a first stae of approval, ‘Like It’ which is the second stage of approval and ‘It’s Okay’ which is a third stage of approval. The campaigns produced using the interactive system of the present invention is connected to a physical setup to be displayed on various physical devices as shown in Figure 6.0. A physical setup displaying the campaigns comprise a camera, 4G antenna, a TV screen for displaying the videos, an antenna & camera cable, HDMI cables, a 12v power adapter, power cable and a microcontroller set which comprises a micro-controller board and 4G LTE module.

Throughout this specification, unless the context requires otherwise, the word “comprise”, or variations such as “comprises” or “comprising”, will be understood to imply the inclusion of a stated step or element or integer or group of steps or elements or integers, but not the exclusion of any other step or element or integer or group of steps, elements or integers. Thus, in the context of this specification, the term “comprising” is used in an inclusive sense and thus should be understood as meaning “including principally, but not necessarily solely”.

CLAIMS

1. A system of facial and body gesture-based interaction (100) for recognising faces and gestures to allow for a continuous narrative that evolves with each user facial expression or body gesture, the system comprising:

a middleware platform (102) that links facial vectors or body data for playing a series of branching videos;

a display screen (105) for projecting the series of branching videos from the middleware platform;

a microcomputer (106) having an algorithm that works in communication with the middleware platform by supplying electronic triggers to the middleware platform;

a hardware having an inbuilt artificial intelligence mechanism (107) for detecting facial and body vectors to trigger signals to the middleware platform;

a computer vision hardware camera (108) for receiving images and video signals of the user's facial expression or body gesture and further analysing and running the images and video signals through a series of algorithms for input to the middleware platform; and

a cloud database (109) for storing information of vectors on body pose estimation as reference data to match with a specific branching video once images and video signals of the user are received;

characterized in that

the middleware platform (102) that links facial vectors or body data for playing a series of branching videos comprises:

a video server (104) where branching videos, audios or graphics are arranged to be displayed in sequence when electronic triggers are received by the middleware platform; and

a content management module (103) for executing machine readable instructions to structure digital assets based on a plurality of artificial intelligence vectors that involve face and body detection.

2. The system of facial and body gesture-based interaction (100) of claim 1, wherein the plurality of artificial intelligence vectors of the content management module (103) further cover at least 106 facial points that involve facial expressions, attributes, actions and recognition.

3. The system of facial and body gesture-based interaction (100) of claim 1, wherein the plurality of artificial intelligence vectors of the content management module (103) further cover at least 14 body points that involve body pose detections, distances, body structure attributes and human attire.

4. The system of facial and body gesture-based interaction (100) of claim 1, wherein the digital assets are videos, images or motion graphics.

5. The system of facial and body gesture-based interaction (100) of claim 1, wherein the middleware platform (102) that links facial vectors for playing a series of branching videos further comprises that the facial vectors involve smiling in order to play a series of branching videos where the user is able to obtain a reward, play or vote.

6. The system of facial and body gesture-based interaction (100) of claim 1, wherein the middleware platform (102) that links facial vectors for playing a series of branching videos further comprise that the middleware platform (102) works in communication with a gantry system (110) to allow for opening and closing of a gate upon verification of facial vectors of the system.

7. The system of facial and body gesture-based interaction (100) of claim 1, wherein the hardware having an inbuilt artificial intelligence mechanism (107) for detecting facial and body vectors to trigger signals to the middleware platform (102) further comprises machine readable instructions selected from SMRT, UOB or WINK.

8. A system of facial and body gesture-based interaction (1100) for recognising faces and gestures to allow for a continuous narrative that evolves with each user facial expression or body gesture, the system comprising:

a middleware platform (1102) that links facial vectors or body data for interacting with a 3D avatar or other such intelligent objects;

a display screen (1105) for projecting the 3D avatar from the middleware platform;

a microcomputer (1106) having an algorithm that works in communication with the middleware platform by supplying electronic triggers to the middleware platform;

a hardware having an inbuilt artificial intelligence mechanism (1107) for detecting facial and body vectors to trigger signals to the middleware platform;

a computer vision hardware camera (1108) for receiving images and video signals of the user's facial expression or body gesture and further analysing and running the images and video signals through a series of algorithms for input to the middleware platform; and

a cloud database (1109) for storing information of vectors on body pose estimation as reference data to match with a specific 3D avatar once images and video signals of the user are received;

characterized in that

the middleware platform (1102) that links facial vectors or body data for interacting with a 3D avatar or other such intelligent objects comprises:

a video server where the 3D avatar or intelligent objects are arranged to interact with the user when electronic triggers are received by the middleware platform; and

a content management module (1103) for executing machine readable instructions to structure digital assets based on a plurality of artificial intelligence vectors that involve face and body detection.

9. A system for automated facial and body gesture-based interaction (100), the system comprising:

a middleware platform (102);

an imaging device for capturing images of a user's facial expressions and/or body gestures, and processing those images;

a processor including an AI mechanism to detect facial and body vectors from the imaging device output to trigger signals to the middleware platform;

the system being further adapted to communicate with a cloud database (109) which stores information of vectors on body pose estimation as reference data; wherein the middleware platform determines from the inputs a meaning for at least some facial expressions and/or gestures, and causes an output to be generated responsive to that meaning.

10. A method (200) of facial and body gesture-based interaction for recognising faces and gestures to allow for a continuous narrative that evolves with each user facial expression or body gesture made by a user in a system of facial and body gesture-based interaction, the method comprising steps of:

initiating an interactive system flow where data from a camera frame (204) is filtered and analysed;

receiving an image byte array data as detected by the camera frame (206);

converting the image byte array data to a YUV data format (208);

analysing the YUV data format (210);

detecting if a human is present (212);

if a human is detected;

analysing of facial and body attributes are determined by tracking semantic key points (214),

if there is no human detected;

routing of interactive system flow to initiating the interactive system flow by receiving a camera frame (204);

matching a plurality of facial and body attributes to a branching video (216);

if there is a match;

playing a specific branching video (218),

if there is no match;

routing of interactive system flow to initiating the interactive system flow by receiving a camera frame (204); and

playing a further specific sequence of branching video (220);

if there is further interaction;

playing a sequence of branching video before ending the session (222);

if there is no further interaction;

routing of interactive system flow to initiating the interactive system flow by receiving a camera frame (204).

11. The method of claim 10, wherein analysing of facial and body attributes are determined by tracking semantic key points (214), said method further comprises:

tracking semantic key points by estimating and analyzing body parts such as a bottom torso, left hip, left knee, left foot, right hip, right knee, right foot, center torso, upper torso, neck base, center head, right shoulder, right elbow, right hand, left shoulder, left elbow and left hand and further running a set of data on the semantic key points through a series of algorithms and further match the set of data to stored parameters within the interactive system.