FORM PTO-1449 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTY DOCKET NO.	SERIAL NO. 12/474,453
INFORMATION DISCLOSURE CITATION IN AN APPLICATION (Use several sheets if necessary)	APPLICANT Gregory N. Snook et al.	
	FILING DATE May 29, 2009	GROUP 2713

U.S. PATENT DOCUMENTS

Examiner Initials		DOCUMENT NUMBER	DATE	NAME	CLASS	SUBCLASS	FILING DATE If appropriate
		5,486,001	1-1996	Baker			

FOREIGN PATENT DOCUMENTS

		DOCUMENT NUMBER	DATE	COUNTRY	CLASS	SUBCLASS	Translation YES	NO
		19259/92	2-1995	Australia				
		WO92/21412	12-1992	WIPO				

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, etc.)

		Fletcher, Andrew, EyeToy: "Kinetic Combat- Review," http://www.gamecritics.com/node/3775, 2007, printed from the Internet on Jan. 24, 2011, 4 pages.
		Davis, James et al., "Virtual PAT: A Virtual Personal Aerobics Trainer," MIT Media Lab, 1998, 6 pages.
		"Virtual PAT: A Virtual Personal Aerobics Trainer- Microsoft Academic Search," http://academic.research.microsoft.com/Paper/350476.aspx, printed from the Internet on January 10, 2011, 2 pages.
		Wang, Yong et al., "USING HUMAN BODY GESTURES AS INPUTS FOR GAMING VIA DEPTH ANALYSIS," IEEE International Conference on Multimedia & Expo, 2008, pp. 993-996.

EXAMINER	DATE CONSIDERED
/Truc Chuong/	01/13/2013

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP 609; Draw line through citation if not in conformance and not considered. Include a copy of this form with next communication to applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.C./

ITW

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN THE APPLICATION OF:

APPLICANT : **GREGORY N. SNOOK et al.**

APPL. NO. : **12/474,453** **ART UNIT: ~~2713~~ 2173**

FILED : **MAY 29, 2009** **EXAMINER: VU, KIEU**

FOR : **GESTURE COACH**

COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

TRANSMITTAL

Sir:

Transmitted herewith are:

1. Third party submission (1 page)
2. PTO 1449, listing 3 patent documents and 4 other documents
3. Publications (68 pages)
4. Proof of Service on Applicant
5. Check for $180 (third party submission)

Any required fees may be charged to Deposit Account No. 12-1662 of Richard C. Litman.

Respectfully submitted,

Charles K. Friedman
Registration No. 39,195

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN THE APPLICATION OF:

APPLICANT	:	GREGORY N. SNOOK et al.	
APPL. NO.	:	12/474,453	ART UNIT: 2713
FILED	:	MAY 29, 2009	EXAMINER: VU, KIEU
PUB. NO.	:	2010/0306712	PUB.DATE:Dec.2, 2010
FOR	:	GESTURE COACH	

COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

THIRD PARTY SUBMISSION IN PUBLISHED APPLICATION

UNDER 37 CFR 1.99

Sir:

An information disclosure statement (FORM PTO-1449) containing citations for one U.S. patent document, two foreign patent documents, and four other documents is enclosed herewith.

This submission has been served on the Applicant and proof of service is attached. A check in the amount of $180.00 as required by 37 CFR 1.17(p) is attached.

Respectfully Submitted,

Charles K. Friedman
Reg. No. 39,195

02/02/2011 AWONDAF1 00000031 12474453
01 FC:1806 180.00 OP

CERTIFICATE OF SERVICE

The undersigned hereby certifies that a true copy of

THIRD PARTY SUBMISSION IN PUBLISHED APPLICATION UNDER 37 CFR 1.99
PTO 1449
1 US PATENT DOCUMENT
2 FOREIGN PATENT DOCUMENTS
4 PUBLICATIONS

was served upon Counsel by First Class Mail on the 31st day of January, 2011:

Peter Trahms-Neudorfer
WOODCOCK WASHBURN LLP (MICROSOFT CORPORATION)
CIRA CENTRE, 12TH FLOOR
2929 ARCH STREET
PHILADELPHIA PA 19104-2891

Charles K. Friedman
Registration No. 39,195

FORM PTO-1449 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTY DOCKET NO.	SERIAL NO. 12/474,453
INFORMATION DISCLOSURE CITATION IN AN APPLICATION (Use several sheets if necessary)	APPLICANT Gregory N. Snook et al.	
	FILING DATE May 29, 2009	GROUP 2713

U.S. PATENT DOCUMENTS

Examiner Initials		DOCUMENT NUMBER	DATE	NAME	CLASS	SUBCLASS	FILING DATE If appropriate
		5,486,001	1-1996	Baker			

FOREIGN PATENT DOCUMENTS

		DOCUMENT NUMBER	DATE	COUNTRY	CLASS	SUBCLASS	Translation YES	NO
		19259/92	2-1995	Australia				
		WO92/21412	12-1992	WIPO				

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, etc.)

		Fletcher, Andrew, EyeToy: "Kinetic Combat- Review," http://www.gamecritics.com/node/3775, 2007, printed from the Internet on Jan. 24, 2011, 4 pages.
		Davis, James et al., "Virtual PAT: A Virtual Personal Aerobics Trainer," MIT Media Lab, 1998, 6 pages.
		"Virtual PAT: A Virtual Personal Aerobics Trainer- Microsoft Academic Search," http://academic.research.microsoft.com/Paper/350476.aspx, printed from the Internet on January 10, 2011, 2 pages.
		Wang, Yong et al., "USING HUMAN BODY GESTURES AS INPUTS FOR GAMING VIA DEPTH ANALYSIS," IEEE International Conference on Multimedia & Expo, 2008, pp. 993-996.

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP 609; Draw line through citation if not in conformance and not considered. Include a copy of this form with next communication to applicant.

AU9219259

(12) PATENT ABRIDGMENT (11) Document No. AU-B-19259/92
(19) AUSTRALIAN PATENT OFFICE (10) Acceptance No. 656781

(54) Title
PERSONALIZED INSTRUCTIONAL AID

International Patent Classification(s)
(51)[5] A63B 069/36 G06F 015/62

(21) Application No. : 19259/92 (22) Application Date : 01.06.92

(87) PCT Publication Number : WO92/21412

(30) Priority Data

(31) Number	(32) Date	(33) Country
PK6405	30.05.91	AU AUSTRALIA

(43) Publication Date : 08.01.93

(44) Publication Date of Accepted Application : 16.02.95

(71) Applicant(s)
RICHARD JOHN BAKER

(72) Inventor(s)
RICHARD JOHN BAKER

(74) Attorney or Agent
JOHN R. G. GARDNER , PO Box 443, MUDGEERABA QLD 4213

(56) Prior Art Documents
US 4942391
US 4852184
US 4137566

(57)

An instructional aid (10) for enabling and assisting a person to emulate a predetermined movement such as a golf swing including an image/data capturing device (11) for scanning capturing and storing the image/data signals of the person's golf swing or movement, and a transceiver (14) for transmitting the image/data signals over a communications network (15) to a remote computer (17). The computer (16) contains a database (17) which stores in digital form image signals and bio-mechanical data of a number of preselected swings or movements. The computer (16) analyses the person's image/data signals against image/data of a preselected swing and generates in digital form further visual images/data signals which can be transformed so as to assist or enable a person to be able to emulate the preferred swing. The instructional aid may also be used for analysing or reporting on particular geographical regions or objects and provides a specialized presentation for development of that region or object from information stored in the computer database.

OPI DATE 08/01/93 APPLN ID 19259/92
AOJP DATE 25/02/93 PCT NUMBER PCT/AU92/00253

AU9219259

TY (PCT)

(51) International Patent Classification 5 :
A63B 69/36, G06F 15/62

A1

(11) International Publication Number: WO 92/21412

(43) International Publication Date: 10 December 1992 (10.12.92)

(21) International Application Number: PCT/AU92/00253

(22) International Filing Date: 1 June 1992 (01.06.92)

(30) Priority data:
PK 6405 30 May 1991 (30.05.91) AU

(71)(72) Applicant and Inventor: BAKER, Richard, John [AU/AU]; 10 Dundee Court, Robina Woods, QLD 4226 (AU).

(74) Agent: GARDNER, John, R., G.; P.O. Box 443, Mudgeeraba, QLD 4213 (AU).

(81) Designated States: AT, AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH, CH (European patent), CI (OAPI patent), CM (OAPI patent), CS, DE, DE (European patent), DK, DK (European patent), ES, ES (European patent), FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), GN (OAPI patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU, LU (European patent), MC (European patent), MG, ML (OAPI patent), MN, MR (OAPI patent), MW, NL, NL (European patent), NO, PL, RO, RU, SD, SE, SE (European patent), SN (OAPI patent), TD (OAPI patent), TG (OAPI patent), US.

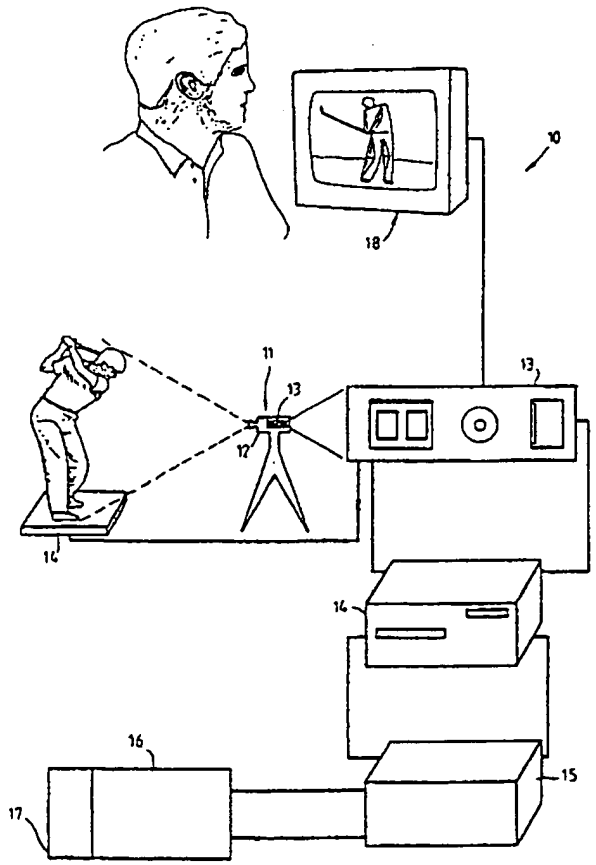
Published
With international search report.

656781

(54) Title: PERSONALIZED INSTRUCTIONAL AID

(57) Abstract

An instructional aid (10) for enabling and assisting a person to emulate a predetermined movement such as a golf swing including an image/data capturing device (11) for scanning capturing and storing the image/data signals of the person's golf swing or movement, and a transceiver (14) for transmitting the image/data signals over a communications network (15) to a remote computer (17). The computer (16) contains a database (17) which stores in digital form image signals and bio-mechanical data of a number of preselected swings or movements. The computer (16) analyses the person's image/data signals against image/data of a preselected swing and generates in digital form further visual images/data signals which can be transformed so as to assist or enable a person to be able to emulate the preferred swing. The instructional aid may also be used for analysing or reporting on particular geographical regions or objects and provides a specialized presentation for development of that region or object from information stored in the computer database.



PERSONALIZED VIDEO AID

Technical Field

This invention relates to a method and means for providing an expert knowledge base in a video teaching aid which in one aspect may be used as an audio visual teaching aid and in particular as an audio visual teaching aid for teaching of movements and techniques, such as used in sports for the game of golf. The present invention, however, is applicable to other movements, sports pastimes or techniques.

In a further aspect the present invention relates to a method and means for providing an expert visual or audio visual assessment of a particular proposal relating to geographical locations. One such application is to generate a visual display of a particular tract of land for a golf course layout made up from a number of pre stored hole designs or computer generated hole designs which have taken into account all aspects of the final layouts design on the overall development, with the inclusion of audio or audio visual data pertaining to structural methods for success.

The present invention has further application in providing an assessment on a particular object by for example comparing properties of that object with those of a preferred object or objects.

Background Art

As golf has become an increasingly popular pastime over the years, many players wishing to improve their own personal techniques, have pursued professional assistance. Often players, be they professionals or amateurs, may wish to emulate successful movements of top players or coaches who are exposed in the mass-media. At the present time, these players have a number of choices available to them in attempting to try to emulate such movements, which may include general practice, private coaching, viewing current audio visual instruction tapes, recording their own personal movements and analyzing same, or studying photographs/slides etc. of their swing in action.

In all the above cases, the onus is placed upon

the individual or coach, to perceive those areas of movement that differ between the current technique of the player and that of the preferred technique they are trying to achieve. Even for a specialized person working in their own chosen sport or profession, such differences in movement can be so minute or hard to detect, with say a person's weight shift or pressure in their grip or arms, that they can go undetected and so handicap that persons potential rate of improvement. The reason for this particularly in the above instructional visual teaching cases is that the viewer watches information flowing to him in one direction only regarding the current or preferred technique and thus the viewer cannot perceive all aspects of the movement because of no corresponding physical forces or certain performance data to relate specifically with because such visual presentations do not show their current technique along with the preferred technique with the addition of bio mechanical information. Also in the use of current general prerecorded instructional audio visual presentations they direct their information towards a mass audience as a whole without any allowance for a persons particular technique or requirement of problem solving.

In other situations where conceptual instructional advice is required, the quality of advice can vary, for example when a person is considering whether or how to best construct a particular tract of land he owns or is considering acquiring, say for the development or remodelling of a golf course, normally expensive preliminary surveying, design and feasibility works are carried out, to which he is reliant on all aspects of the development being expertly covered, to which this is not always the case. Similarly, if alternative ventures in particular areas are proposed, no audio visual means are presently available which provides a customized or personalized completely regenerated, expertly compiled computer analysis in the visual or audio visual form of potential works, described within.

Summary of the Invention

The present invention aims to provide in one aspect an

expertly analyzed instructional aid employing visual or audio/visual techniques which allows for a flow of information in both directions between say selected coaches and pupils so as to help that pupil/viewer improve his or her personal movements and techniques more quickly and without as much guesswork on his or her part.

The present invention also provides a method and means for teaching or instructing such movements or techniques in precise ways, that enables a person/mechanism to more clearly emulate another movement. Whilst the present invention is particularly suited in one application to sports coaching it may also be applied to other areas where an emulation of a precise movement/procedure is required, as will become more apparent from the following operating description. The present invention may additionally be applied to providing expertly compiled opinion in the form of a visual or audio/visual overview of a particular geographical location altered in accordance with selected criteria.

With the above, and other objects in view, the present invention provides in a first preferred aspect, a method for providing an instructional aid for assisting a person to emulate a preferred movement including the steps of capturing and storing visual image signals representative of a particular movement, comparing said captured image signals with stored image signals representative of a selected preferred movement, and regenerating further visual image signals based on said image signals of said particular movement adjusted to emulate said stored image signals of said selected movement.

In a second aspect, the present invention provides apparatus for providing an instructional aid for assisting a person personally to emulate a preferred movement or technique, said apparatus including means for capturing visual image signals of a particular movement or technique, means for storing said captured visual image signals, means for storing visual image signals of a plurality of preferred movements or techniques, means for comparing said captured visual image signals with the stored image signals of a selected said

preferred movement or technique, and means for regenerating further visual image signals of said particular movement or technique adjusted in accordance with said selected preferred movement or technique to assist a person to emulate said selected preferred movement or technique. The term "image signals" as used herein includes signals representative of such images such as digital information which can be converted in to visual images.

Preferably the visual images of the particular movement are captured and stored in digital form using known video recording techniques such as with a video camera which produces a video cassette containing the images of said particular movement. Alternatively other means may be provided for capturing and storing the visual images as described further below.

The visual images of the preferred movements are preferably stored in digital form in a database of a computer so that a selected preferred movement can be readily accessed. The computer is most preferably located at a remote location so as to be accessible from a number of different locations.

Preferably means are provided to transmit the captured images in digital form for receipt by the computer. The transmitting means preferably transmits the images over the normal telephone network, however, other communications links may be employed.

The comparison in the computer can be carried out using either software techniques or hardware techniques and various sampling procedures. The computer is also preferably software or hardware programmed to take into account the size of the player and other physical characteristics such as length of arm so that a compensation factor may be included whereby the movements such as swings of the player and the preprogrammed stored movements or swing in the computer can be truly compared. This comparison can take place at different times or positions in say a swing to produce signals representative of the difference between the swing. This comparison signal may comprise a measurement of distance,

angle, speed of swing or other parameters. Such signals may then be employed in regenerating the new visual image which approximates the captured image altered to emulate the preferred image.

Whilst the invention described within has particular applications to audio visual teachings in sport, so as to help and assist a person emulate a particular movement or technique, it may also be applied to many other situations and fields where an emulation or overview of a precise movement or technique is required. Thus, the invention may be applied to various applications in the arts fields, for example the teaching of dance steps, the computing field where precise movements of a person using a keyboard, mouse or the like are required to be studied to help an end-user progress in learning software, or within the medical field where precise movements are required to be studied, an expert advice given such as with human body movements.

Additionally the present invention may be applied to fields where expert audio visual systems appraisals and emulations of a development structure may be required such as in the property field for precise calculations and designs of future development proposals, or where expert audio visual systems appraisal of a particular object/s is required by comparison with a preferred or ideal or object/s.

The present invention thus provides in a further aspect a method of providing an instructional aid relating to a development proposal comprising the steps of capturing and storing visual image and location data signals of a proposed development location, analyzing said captured signals against the signals of a selected or preferred known development or means for forming a new design development and regenerating further visual image/data signals based on said signals of said proposed development adjusted in accordance with said preferred development or said newly formed design development so as to provide a recommendation for said preferred development proposal.

In a further aspect the present invention provides apparatus for providing an instructional aid based on a visual image of a particular movement or location, said apparatus including means for visually scanning said movement or location and for producing visual image signals representative of said movement or location, means for storing said visual image signals, means for transmitting said visual image signals to a remote location, means for comparing said transmitted visual image signals with corresponding signals of a selected prestored image or images, means for regenerating further image signals comprising said image signals of said particular movement or location adjusted in accordance with said prestored signals of said image or images, and means for retransmitting said regenerated image signals and converting said image signals into a format which enables viewing thereof.

In yet a further aspect, the present invention provides apparatus for providing a personalised instructional aid in a development presentation based upon a visual image signal of a particular location, said apparatus including means for visually scanning said location and for producing image and informational data signals representative of said location, means for storing said visual/data signals, means for transmitting said visual/data signals to a remote location, means for comparing said visual/data signals with the signals of a selected prestored location and the structural procedure relating thereto visual/data signals, means for regenerating further visual data/signals based on said signals of said particular location adjusted in accordance with said prestored procedure signals to assist in said development, means for adding dialogue and information data to said regenerated signals and means for transmitting said regenerated image/data signals and converting said signals into a format which enables viewing thereof.

In another aspect, the present invention provides apparatus for providing a personalised audio visual aid based upon optical photometric characteristics of a selected

object, said apparatus including means for visually scanning, sensing and capturing visual/data signals representative of said object, means for storing said visual/data signals, means for transmitting said visual/data signals to a remote location, means for comparing said captured visual/data signals with prestored image signals of a selected object/s, means for regenerating further visual data/signals comprising said signals of said captured object adjusted in accordance with said prestored preferred object/s signals, means for transmitting said regenerated image/data signals and converting said image signals into a format which enables viewing thereof.

Preferably the characteristics of the object are sensed using imaging and non-imaging sensor systems. Preferably also means are provided for adding dialogue or informational data to the regenerated format.

In yet a further aspect the present invention provides a method for providing a personalized audio visual teaching aid for assisting a person to emulate, or consider advice on a preferred activity or subject of interest, said method including the steps of capturing and storing visual image, audio and/or data signals representative of a particular activity or subject of interest, comparing said captured signals with stored image, and audio and/or data signals representative of a selected preferred activity or subject of interest, and generating audio visual image signals based on said captured signals analysed against the signals of said selected preferred activity or subject of interest.

In yet a further aspect the present invention provides apparatus for providing an audio visual teaching aid for assisting a person personally to emulate or consider advice on a preferred activity or subject of interest, said apparatus including means for capturing visual image, and audio and/or data signals of a particular activity or subject of interest, means for storing said captured signals, means for storing signals of a plurality of preferred activities or subjects of interest, means for

comparing said captured signals with said stored signals of a selected preferred activity or subject of interest, and computer means for regenerating further audio visual and data signals of said particular activity or subject of interest adjusted in accordance with said selected preferred activity or subject of interest whereby to assist a person to emulate or consider said selected preferred activity or subject of interest.

Brief Description of the Drawings

Reference will now be made to the accompanying drawings which illustrate a preferred embodiment of the invention and wherein:-

Fig. 1 illustrates in block diagram the instructional aid according to the present invention applied to the game of golf; and

Fig. 2 is a flow chart showing the procedure involved in the invention.

Detailed Description of the Preferred Embodiment

Fig. 1 illustrates the apparatus for use in the present invention described below in reference to emulation of a golf swing but useable in any situation where emulation of personalised audio visual interactive advice relating to a particular movement is required. The apparatus 10 includes a visual image capturing device 11 which is adapted to capture movement of the person undertaking a golf swing for subsequent analysis and correction as described below with reference to Fig. 1 and the flow chart of Fig.2.

To capture the initial detailed visual images relating to physical characteristics including movement techniques, employed during the golf swing various scanning or electronic collecting procedures are used. For this purpose the image capturing device 11 preferably incorporates or comprises a portable electronic camera such as a video camera 12 or the like having an associated recording or signal storage means 13 as is known in the art to form and store signals representative of the characteristics of movement etc. being viewed.

Where recording of movements takes place, bio-

mechanical information/data may be transmitted to the recording or storage means 13 by an electronic means, hard wiring or other communication means, either at the same time, as the captured visual signals relating to scanned images of the person being viewed or at a different time. In the embodiment illustrated, bio-mechanical information may be captured using a weight sensing pad or pads 13' which will sense weight transference during the swing.

The means used to collect and transmit such bio-mechanical information/data, is at all times non restrictive in its overall performance, so as to allow the pupil being analyzed to achieve normal free flowing movement without restriction. One such means may take the form of mechanical device/s affixed, or attached, to the person or golf club moved by the person so as to show data such as pressures transferred through the arms or hands, proper releasing techniques, velocities achieved etc. Alternatively or additionally, sensing devices maybe positioned alongside, below or above the person whose image is being captured. Such sensing devices as above may comprise weight or load sensitive pads, monitors electronic light beaming device or the like. Bio-mechanical, signals may also be obtained by various scanning, editing and digital techniques, used within a remote computer database as described further below, once the signal has been received by the remote database, to determine also the many characteristics of the person being studied.

The image capturing device 11 is preferably of portable construction, and of high recording quality, so as to scan, receive and store images and information/ data, in broadcast quality. These visual or information signals are converted by sensors or the like, into electrical charge patterns etc, which are read out or scanned to produce, digital electronic signals representative thereof, from electrical or light signals,

produced by an electronic pick up tube, charge-coupled device (CCD), or the like means within the camera framework, or by the information signalling devices referred to above.

For initially capturing the necessary signals for analysis, more than one image capturing device 11 may be used, so as to channel information regarding various visual details, information/data or the like, to one recording or storage medium. In this regard various switching techniques may also be used to switch between selected image capturing devices 11.

The signal storage or recording means may take the form of Magnetic Tape/Disk, Laser Based or Optical Video Disk, Magneto-Optic Disk, Magnetic-Laser-Optical Disk or other storage medium. Alternatively digitised signals representative of the scanned images or bio mechanical and information data may be also transferred or stored onto a computer memory chip or the like, held within the camera framework.

The apparatus 10 also includes an audio/visual signal transceiver 14 adapted to receive and transmit audio/visual signals and other information signals in the manner described below.

The image capturing device 11 referred to above in one form is a unit discrete from the audio/visual signal transceiver 14 and adapted to be connected thereto or the two maybe one and the same, so that the image capturing device 11 also incorporates a unit for transmission and reception of signals to and from remote databases.

The audio/visual signal transceiver 14 may take the form of a portable signal unit or a personal computer (PC), which can convert, compress or suitably accept, a down loading of captured signals, stored within the storage means 13 of the portable camera unit and then reform, record or store such signals in a suitable digital state for later transmission over a communications network 15. The communication network 15 may comprise the general switched telephone network, satellite links, television quality cable links or the like over which the signals representative of the persons movement together with bio-mechanical information where stored are directed as

electrical or light signals to a remote computer 16 for storage in a computer database 17 for later analysis against selected preferred signals also stored in a database of the computer or a further computer. As is conventional a modem or other similar signal conversion device may be associated with the computer 16 to receive the transmitter signals and convert those signals into a suitable form for reading by the computer 16. The modem or other similar signal conversion device may also be employed for transmission of signals from the computer 16 as described further below.

Initial capturing of audio, visual and data signal information, may also be undertaken using the audio visual signal transceiver 14 (if not a personal computer) to which a video camera 12 may be directly connected so as to enable signals captured by the camera 12 to be recorded directly onto suitable storage medium held within the transceiver 14. The transceiver 14 is then able to transmit these captured signals to remote locations, when directly connected to the communications network described above. Thus a single unit may perform all functions of recording, transmitting and receiving audio, visual and data signal information. The transceiver 14 for this purpose may include dual signal storage facilities, to achieve the above functions, however the transceiver 14 may also be constructed so as to enable all functions of recording and playback in each mode within a single signal storage device.

This transceiver 14 is software and hardware programmed and may offer some of the following features:

1) The senders details for later personalizing effects to the final video presentation i.e. name and spoken language and means for selecting the person with whom the golf swing of the person is to be compared;

2) The transmitting parties/agents details for account keeping purposes and security for the system i.e. digitised security code;

3) Circuitry for increased transmission speed;

4) Circuitry and displays for showing details regarding

signalling times to reach database and regeneration times once retransmitted back to pupil/agent remote location, for account keeping purposes;

5). Circuitry and displays for showing format of transmitted or requested signals in i.e. PAL, NTSC or SECAM etc;

6). Circuitry for Time, Date, Auto Dial - Redial, Stop, Start, Receive, Send and possibly an Advertising channel etc; and

7). Circuitry and displays for showing to whom the regenerated signals belong and to enable the retransmitted regenerated signal to be added to the already prerecorded signal, described below, if desired.

As stated above, the computer 16 is arranged normally at a remote location from the audio/visual transceiver 14 and includes a database 17 which stores selected prerecorded video signals in digital form which in the case of golf swings may include a selection of swings showing the movement of different golf professionals which are required to be emulated. As the computer 16 and image capturing device 11 are not physically affixed or attached together, the computer 16 and database 17 may be of a sufficient size and capacity to enable the desired functions to be achieved at one and the same location.

The computer 16 may be at many remote locations and holds within its database the following prerecorded information in digital form, which provide a basis for forming the regenerated audio visual signals.

1). A digitised audio visual library of selected coaches, advisers professional or mechanisms (such as a golf simulation machine) movements, techniques, steps, objects or procedures, that have been previously analysed, recorded and reformed into suitable digitised signals, for later analysis and regeneration in the database against the remote signals received.

2). A digitised bio mechanical library of selected coaches or mechanisms movements measured individually or as a group, for later display with the regenerated audio/visual

signals.

The computer 16 incorporates a means for receiving and editing the transmitted visual data or bio-mechanical and information data signals, relating to the movement of the golfer and arranges same in such a way as to be able to match the received visual data signals with the previously established and selected database signals of their preferred applications and generates a new video signal to enable the viewer to see personally, in regenerated form, those differences that exist in (a) a teaching mode, between the two movements, that of the pupil and the selected professional or coach and (b) in an analysis mode, those ideas that might be possible for their preferred swings (or applications in development fields where the method and apparatus of the invention is used in this application.) For this purpose, the computer is software and/or hardware programmed to compare the respective swings of the person being analyzed and the selected preferred swing with adjustments made such as for size of person, speed of swing and other parameters to allow a correct comparison. The differences between the swings are measured by sampling or other techniques and this difference added or subtracted from a persons recorded swing and new visual images generated on this basis which emulates the preferred movements.

The computer 16 may also include a means for adding dialogue to the regenerated audio visual signal recordings, so that the viewer can receive personalized instruction or advice from afar, similar to having private instruction or personal consultation.

The computer 16 may further include a means for converting received signals, if needed, in one broadcasting mode, into those of another countries i.e. PAL-NTSC-SECAM etc. To accomplish this within the computer, a Standards Converter or the like may be used, using Optical Scan or Digital Standards features or the like.

The computer 16 may additionally include a means for determining costing on each and every received and

regenerated signal that passes through the system, based upon the following information:-

(a) Signal transmission time to and from the computer.

(b) The amount of analysis and regeneration time required, for each and every presentation, such as with the showing of audio visual faults, bio mechanical/characterized data graphs, dialogue and broadcasting changes etc.

The computer 16 is software and hardware programmed, to accept digitised signals transmitted to it, from remote locations, associated with the audio visual transceiver 14, so as to personally relate in the final regeneration and editing phases to the viewer, this being performed by, using various digital editing techniques and procedures as described below.

The audio visual presentation generated by the computer 16 can take the form of split-screening (generating new visual images of the preferred technique displayed along side the present technique), superimposing (where the preferred technique is superimposed over the present technique) or other display. Graphics and charts etc, for bio-mechanical/ information displays, so that the viewer can clearly see those changes that are required or need to be performed, to develop or pursue the preferred movement, procedure or proposal are also displayed. These superimposing effects or the like, may take the form of stick figures, detailed drawings etc, to normal human or visual appearances, so as to show the viewer any variances between his stroke and the preferred stroke. For ease of editing the subject material within this computer database, a time code or the like, maybe also burnt into the original recorded signal, so as to make final editing much more efficient within the remote database. This procedure may also be linked with an edit controller or the like, within the computer 16, which locks the two signals electronically together during edits.

To personalize the regenerated signal, dialogue is also added when required, so that the selected instructor or adviser may relate to the viewer/pupil more personally, than with other audio visual presentations. Part of this personalizing effect is initially accomplished in the original transmission phase, by the name, spoken language, broadcasting mode or location

statistics etc of the final viewer being captured, within this database, to which an addition of dialogue can also be added when required, in the regeneration signal phase, showing those changes they need to perform or consider, to obtain the preferred movement, thus totally personalizing the new regenerated signal to the viewer.

For this purpose such dialogue, may be initially captured from selected instructors and advisers and stored in the data base and may only take the form of a brief statement, so as to make the viewer fully aware of those changes that are required, or need their consideration, to perfect the movement, procedure or development etc. In a further aspect where a range of model swings are being analysed, the computer may also form the dialogue from previously learned words or phrases.

The means for adding such dialogue may take the form of a voice energy monitor, speech synthesiser or software programmed computer which can analyse incoming information from the viewer pupil and scan its memory bank of stored or learned words or phrases to find suitable terms, to bring attention to these problem areas or matters needing their attention.

Dialogue may also be added relating to recommendations for the viewers consideration in regards to, suitability of equipment being used to perform the desired movements and techniques.

To increase acceptance as a true teaching aid, scientific data, which in many applications could only be assumed before, such as with personalized coaching, may be presented to a viewer, when needed, in the form of graphs, charts, or the like, set either above, below, alongside or within the total regenerated video signal recording. This information is gathered from their received signals, being matched to, their preferred established signals contained within the database 17 so that the person/pupil can clearly see, those bio mechanical differences that exist between, their current procedures/techniques and that of their selected coaches/professionals procedures/techniques.

Once the new audio visual signal is regenerated the signal is then retransmitted back by the computer 16 and via a modem or the like and the communications network 15 to the remote senders/agents transceiver 14. This procedure maybe performed, from within the computer 16 itself, using various digital, compression and sampling techniques, or it can be again sent back by a further audio/visual transceiver associated with the computer 16.

On receipt at the remote pupils/agents receiving location, the regenerated signal may be in a form suitable for direct saving on suitable saleable storage means such as video tape, disk, chip or the like which enables home viewing such as by a video cassette or disk player or the signal could be shown directly through a monitor 18. If not in a suitable form for direct saving when it arrives at agents/pupils receiving location, the signal may be reformed for storage as above on storage means which permits home viewing by the use of either the receiving audio visual transceiver, or by other suitable conventional means, offering various slowing down or speeding up signal procedures. If desired the regenerated audio/visual signal may be stored on the original storage means eg. video tape containing the players current movements or techniques or stored on a new storage means.

In the embodiment described and shown, the image capturing device 11 stores the captured images on a video cassette 19 of known form in the storage means 13 and the regenerated images are stored on the same cassette 19 or alternatively an additional cassette. The regenerated image in this embodiment is in the form of a stick figure 20 which is superimposed on the true visual image 21 of the players swing or movement and which indicates to the player the angles and positions which he must adopt to emulate the swing.

To add further to this advanced audio visual teaching aids value, as a complete teaching system, compared to other audio visual presentations, previously recorded instructional material, of a non-personalized nature, may

also be added, to the final home storage presentation means such as a video cassette. Thus this regenerated transmitted signal of the pupil's audio visual presentation and data pertaining to movements, maybe then simply added to the prerecorded material already held at the senders/agents remote location. In this format the single recording may then show (a) an initial complete recorded procedure of instruction, as currently

presented with audio visual presentations (b) an addition of their current movements, procedures or images etc, taken before regeneration and (c) the final regenerated personalized audio visual changes, that are required, or affect, the desired changes needed by the viewer.

A typical presentation with a preferred format and structure in the final video tape or other viewable means could be as follows :

Pupil: Mr Rick Baker

Selected coach : Mr Bobby Jones

Applications for: Golf

Dialogue in: English

1. Preferred audio visual start/introduction:

"Hi, Rick, I am Bobby Jones, it is my pleasure to have studied your golf swing and these are the points I feel you should be concentrating on : "

2. Preferred audio visual presentation in progress:

During the presentation whilst the pupil, Mr Baker is viewing the subject matter, the viewing monitor is showing segments of the previously recorded swing, taken at the recording facility, with superimposed preferred movements of Mr Jones against Mr. Baker's current movements and pointing out with the addition of dialogue, those areas that he should be concentrating on, so as to improve his current technique.

Bio mechanical data in the form of graphs or charts etc. are also added during the presentation, so as that he and his local coach may also quite clearly see scientific data pertaining to those performed movements. Dialogue may also be added regarding suitability of equipment being used at this stage.

3. Towards the conclusion of this audio visual presentation, the coach could conclude by saying:

"Thank you Rick for allowing me to help you improve your game, but please continue to work on these points with the help of your local professional, who is also there to assist you in overcoming these problems. Good golfing,

Bobby Jones."

The above embodiment has been described with reference to a golf swing movement, however, as previously stated the invention may be readily applied where other movements are required to be emulated. Such movement may comprise movements of persons or movements of objects or mechanisms of which emulation is required.

In a further form the invention may be applied to the property field to enable professional assessments and presentations of designs for future development proposals or the like. In this form visual images and data relating to signals of a particular area of ground are captured using remote sensing techniques. For this purpose a camera and remote sensing device may be fixed to the underside of an aircraft so as to capture and record geometrical or statistical information of the earth or objects featured beneath. Geographical data achieved through using this remote sensing as above may be computer coded or digitised using video cameras, array scanners, laser beam, microwave or other means.

Information regarding the geographical area being scanned is then transmitted as before using a transceiver to a remote computer containing a database of prestored expert information on attributes required of a particular area of land. Thus where the best standard of golf course is required to be laid out upon a specific parcel of land the data base may include a library of selected hole layouts from different courses. The computer which is software programmed may then select preferred hole layouts to match to the area of land of which geographical features has been scanned and recorded. Alternatively, the computer may itself design the appropriate hole layouts to best fit the area of land and proposed development. The computer database may additionally include feasibility study means for determining and displaying measurements, costings and other procedures required to achieve the end result. The computer may further be programmed to take into account residential housing around the course and the location of same, so as to maximize view of the course etc.

As before, dialogue may be added to the regenerated audio visual presentation as well as statistical graphs or other information relating to the development being studied.

The present invention may also be applied to enable a particular object to be analysed against prestored data relating to a currently known object. In this embodiment and as before an object is scanned so that its characteristics may be determined. Scanning is carried out to obtain the optical photometric characteristics of the object using imaging and non-imaging sensor systems. As before the data obtained is stored and then transmitted to a remote location for analysis in a computer with the corresponding data of a preferred object whose data information is captured in a data base. After comparison and adding, where appropriate, suitable dialogue from say an expert in that field, a regenerated signal is formed and transmitted back to the sender so that the sender has an expert audio visual assessment of the differences between the respective objects. A particular application in this embodiment is for component parts as used in an assembly. The data information obtained relating to that part may be compared to a perfect component part whose data is stored at the remote location so that the differences may be assessed. This embodiment may also be employed for detecting for example faults in a particular component part.

It will be apparent that the latter embodiment may be applied to any suitable targets of which analysis is required.

The present invention thus provides an interactive method and means for expertly teaching, advising or instructing procedures, strategies, assessments or perceptions in audio visual form to enable a viewer/end-user to more closely emulate calculate or develop techniques or make decisions suitable to their various applications.

Whilst the above has been given by way of illustrative embodiment of the invention, all such modifications and variations thereto as would be apparent to

persons skilled in the art are deemed to fall within the broad scope and ambit of the invention as herein defined by the appended claims.

CLAIMS

1. A method for providing an instructional aid for assisting a person to emulate a preferred movement including the steps of capturing and storing visual image signals representative of a particular movement, comparing said captured image signals with stored image signals representative of a selected preferred movement, and regenerating further visual image signals based on said image signals of said particular movement adjusted to emulate said stored image signals of said selected movement.

2. A method according to Claim 1 wherein said image signals of said particular movement are captured and stored using a video camera.

3. A method according to Claim 1 wherein said visual image signals of said preferred movements are stored in a database of a computer and wherein a selected preferred movement is accessed from said data base for comparison with said particular movement.

4. A method according to Claim 3 wherein said image signals are captured and stored at a first location and wherein said computer is located at a second location remote from said first location and wherein means are provided for transmitting said captured images from said first location to said computer at said second location.

5. A method according to Claim 4 wherein said transmitting means comprise independent signal transmitting means or a personal computer.

6. A method according to Claim 4 wherein said transmitting means transmits said image signals over a communication network to said computer.

7. A method according to Claim 4 wherein said regenerated signals are transmitted from said remote second

location to said first location.

8. A method according to Claim 7 wherein said regenerated image signals are stored in image presentation means which permits viewing thereof by said person.

9. A method according to Claim 8 and wherein dialogue relating to said regenerated image signals is added for presentation with said regenerated image signals.

10. A method according to claim 9 wherein said dialogue is selected from words or phrases stored in a database.

11. A method according to claim 7 wherein bio mechanical data relating to said regenerated images is displayed within image presentation means.

12. A method of providing an instructional aid relating to a development proposal comprising the steps of capturing and storing visual image and location data signals of a proposed development location, analyzing said captured signals against the signals of a selected or preferred known development or means for forming a new design development and regenerating further visual image/data signals based on said signals of said proposed development adjusted in accordance with said preferred development or said new design development so as to provide a recommendation for said preferred development.

13. A method according to Claim 12 wherein said visual image/data signals of said proposed location are captured and stored using a video camera and remote sensing devices.

14. A method according to Claim 12 wherein said signals of said selected or preferred known development or new design development are stored in a database of a computer.

15. A method according to Claim 14 wherein said computer is arranged at a location remote from said proposed location and wherein said captured image signals are transmitted to said remote location for said analysis by said computer.

16. A method according to Claim 15 and wherein said regenerated image/data signals are transmitted from said computer to said proposed location and are provided in a format to enable when transferred or added to home viewing means for viewing thereof.

17. A method according to Claim 16 and including means for adding dialogue to said format so as to provide an audio visual presentation.

18. A method according to Claim 17 wherein said dialogue is selected from words or phrases previously stored in a database.

19. Apparatus for providing an instructional aid for assisting a person personally to emulate a preferred movement or technique, said apparatus including means for capturing visual image signals of a particular movement or technique, means for storing said captured visual image signals, means for storing visual image signals of a plurality of preferred movements or techniques, means for comparing said captured visual image signals with the stored image signals of a selected said preferred movement or technique, and means for regenerating further visual image signals of said particular movement or technique adjusted in accordance with said selected preferred movement or technique to assist a person to emulate said selected preferred movement or technique.

20. Apparatus according to Claim 19 wherein said means for storing visual image signals of a plurality of preferred movements or techniques comprises a computer database and

wherein said comparing and regenerating means comprises a computer associated with said computer database.

21. Apparatus according to Claim 20 and including transceiver means for transmitting to said computer said visual image signals of said particular movement or technique and for receiving from said computer said regenerated visual images signals.

22. Apparatus according to Claim 20 and including means for converting said regenerated image signals into a format which allows viewing thereof.

23. Apparatus according to Claim 22 and including means for adding to said format dialogue relating to said regenerated images.

24. Apparatus according to Claim 23 and including means for scanning and selected stored words or phrases for addition to said format as said dialogue.

25. Apparatus according to Claim 22 and including means for adding to said format bio-mechanical data relating to said regenerated images.

26. Apparatus for providing an instructional aid based on a visual image of a particular movement or location, said apparatus including means for visually scanning said movement or location and for producing visual image signals representative of said movement or location, means for storing said visual image signals, means for transmitting said visual image signals to a remote location, means for comparing said transmitted visual image signals with corresponding signals of a selected prestored image or images, means for regenerating further image signals comprising said image signals of said particular movement or location adjusted in accordance with said prestored signals of said image or images, and means for retransmitting said

regenerated image signals and converting said image signals into a format which enables viewing thereof.

27. Apparatus for providing a personalised instructional aid in a development presentation based upon a visual image signal of a particular location, said apparatus including means for visually scanning said location and for producing image and informational data signals representative of said location, means for storing said visual/data signals, means for transmitting said visual/data signals to a remote location, means for comparing said visual/data signals with the visual/data signals of a selected prestored location and the procedure relating to the development thereof, means for regenerating further visual data/signals based on said signals of said particular location adjusted in accordance with said prestored procedure signals to assist in said development, means for adding dialogue and information data to said regenerated signals and means for transmitting said regenerated image/data signals and converting said signals into a format which enables viewing thereof.

28. Apparatus for providing a personalised audio visual aid based upon optical photometric characteristics of a selected object, said apparatus including means for visually scanning, sensing and capturing visual/data signals representative of said object, means for storing said visual/data signals, means for transmitting said visual/data signals to a remote location, means for comparing said visual/ data signals with corresponding prestored signals of a selected object, means for regenerating further visual data/signals comprising said signals of said particular object adjusted in accordance with said prestored preferred object signals, means for transmitting said regenerated image/data signals and converting said image signals into a format which enables viewing thereof.

29. Apparatus according to Claim 28 and including

means for adding dialogue or informational data to said regenerated format.

30. Apparatus according to Claim 28 wherein said object scanning means comprise imaging or non-imaging sensor systems.

31. A method for providing a personalized audio visual teaching aid for assisting a person to emulate, or consider advice on a preferred activity or subject of interest, said method including the steps of capturing and storing visual image, audio and/or data signals representative of a particular activity or subject of interest, comparing said captured signals with stored image, and audio and/or data signals representative of a selected preferred activity or subject of interest, and generating audio visual image signals based on said captured signals analysed against the signals of said selected preferred activity or subject of interest.

32. A method according to Claim 31 wherein said visual image, audio and/or data signals relating to said particular activity or subject of interest are captured and stored by using a portable video camera and imaging or non-imaging remote sensing devices.

33. A method according to Claim 31 wherein said visual image, audio and/or data signals of said preferred activity or subject of interest are captured and stored in a database of a computer and wherein signals representative of said selected preferred activity or subject of interest are accessed from said data base for comparison with said signals of said particular activity or subject of interest.

34. A method according to Claim 33 wherein said visual image, and audio and/or data signals of said particular activity or subject of interest are captured and stored at a first location and wherein said computer is located at a

second location remote from said first location and wherein means are provided for transmitting said captured visual image, and audio and/or data signals from said first location to said computer at said second location.

35. A method according to Claim 34 wherein said transmitting means comprises independent signal transmitting means or a personal computer.

36. A method according to Claim 34 or 35 wherein said transmitting means transmits said visual image, and audio and/or data signals over communication network means to said computer.

37. A method according to any one of Claims 34 to 36 wherein said generated signals are transmitted from said remote second location to said first location by communication network means.

38. A method according to Claim 37 wherein said generated image and, audio and/or data signals are stored in image presentation means, or displayed on a monitor for permitting viewing thereof by said person.

39. A method according to any one of Claims 31 to 38 wherein dialogue is added for presentation with said generated audio visual image signals.

40. A method according to Claim 39 wherein said dialogue is selected from words and phrases stored within a database.

41. Apparatus for providing an audio visual teaching aid for assisting a person personally to emulate or consider advice on a preferred activity or subject of interest, said apparatus including means for capturing visual image, and audio and/or data signals of a particular activity or subject of interest, means for storing said captured

signals, means for storing signals of a plurality of preferred activities or subjects of interest, means for comparing said captured signals with said stored signals of a selected preferred activity or subject of interest, and computer means for regenerating further audio visual and data signals of said particular activity or subject of interest adjusted in accordance with said selected preferred activity or subject of interest whereby to assist a person to emulate or consider said selected preferred activity or subject of interest.

42. Apparatus according to Claim 41 wherein said means for storing visual image, audio and/or data signals of said particular activity or subject of interest captures, stores and compresses said signals at a first location and wherein said signals of said preferred activities or subjects of interest are stored within a database associated with said computer means at a second location remote from said first location, said computer means being adapted to compare and regenerate said further audio visual signals for transmission to said first location.

43. Apparatus according to Claim 42 and including transceiver means for transmitting said visual image, audio and/or data signals of said particular activity or subject of interest over communications networks means to said computer means and for receiving from said computer means said regenerated audio visual, and/or data signals adjusted in accordance with said selected preferred activity or subject of interest.

44. Apparatus according to Claim 43 wherein said transceiver means comprises a portable transceiver means or a personal computer.

45. Apparatus according to any one of Claims 41 to 44 and including means for converting and presenting said regenerated visual image, and audio and/or data signals into

a format which allows viewing thereof through a monitor.

46. Apparatus according to any one of Claims 41 to 45 wherein said computer means adds to said regenerated signals dialogue relating to said signals, said dialogue being selected from words or phrases stored in a database.

47. A method for providing an instructional aid substantially as hereinfore described with reference to the accompanying drawings.

48. A method for providing a personalized audio visual teaching aid substantially as hereinbefore described with reference to the accompanying drawings.

49. Apparatus for providing an instructional aid substantially as hereinbefore described with reference to the accompanying drawings.

50. Apparatus for providing an audio visual teaching aid substantially as hereinbefore described with reference to the accompanying drawings.

DATED this nineteenth day of October 1994

RICHARD JOHN BAKER
By My Patent Attorney

JOHN R. G. GARDNER

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : A63B 69/36, G06F 15/62

A1

(11) International Publication Number: WO 92/21412

(43) International Publication Date: 10 December 1992 (10.12.92)

(21) International Application Number: PCT/AU92/00253

(22) International Filing Date: 1 June 1992 (01.06.92)

(30) Priority data:
PK 6405 30 May 1991 (30.05.91) AU

(71)(72) Applicant and Inventor: BAKER, Richard, John [AU/AU]; 10 Dundee Court, Robina Woods, QLD 4226 (AU).

(74) Agent: GARDNER, John, R., G.; P.O. Box 443, Mudgeeraba, QLD 4213 (AU).

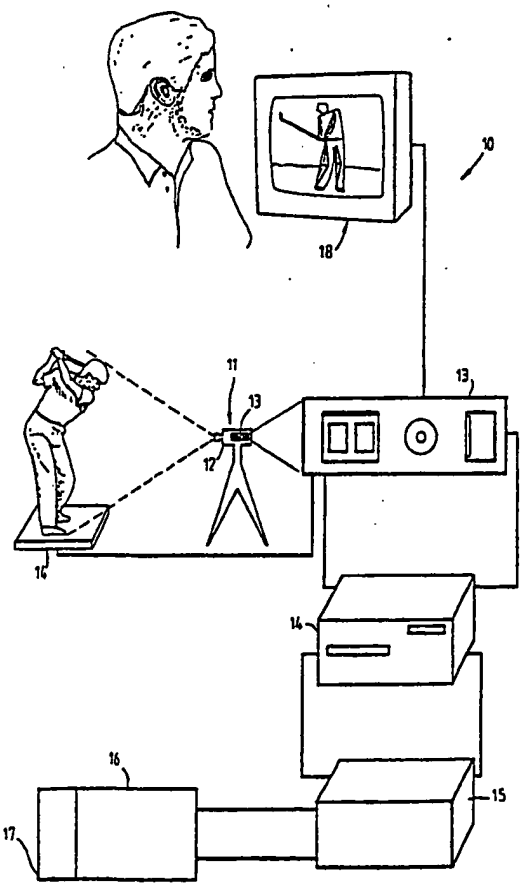
(81) Designated States: AT, AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH, CH (European patent), CI (OAPI patent), CM (OAPI patent), CS, DE, DE (European patent), DK, DK (European patent), ES, ES (European patent), FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), GN (OAPI patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU, LU (European patent), MC (European patent), MG, ML (OAPI patent), MN, MR (OAPI patent), MW, NL, NL (European patent), NO, PL, RO, RU, SD, SE, SE (European patent), SN (OAPI patent), TD (OAPI patent), TG (OAPI patent), US.

Published
With international search report.

(54) Title: PERSONALIZED INSTRUCTIONAL AID

(57) Abstract

An instructional aid (10) for enabling and assisting a person to emulate a predetermined movement such as a golf swing including an image/data capturing device (11) for scanning capturing and storing the image/data signals of the person's golf swing or movement, and a transceiver (14) for transmitting the image/data signals over a communications network (15) to a remote computer (17). The computer (16) contains a database (17) which stores in digital form image signals and bio-mechanical data of a number of preselected swings or movements. The computer (16) analyses the person's image/data signals against image/data of a preselected swing and generates in digital form further visual images/data signals which can be transformed so as to assist or enable a person to be able to emulate the preferred swing. The instructional aid may also be used for analysing or reporting on particular geographical regions or objects and provides a specialized presentation for development of that region or object from information stored in the computer database.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CS	Czechoslovakia
DE	Germany
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
MC	Monaco
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
NL	Netherlands
NO	Norway
PL	Poland
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SN	Senegal
SU	Soviet Union
TD	Chad
TG	Togo
US	United States of America

PERSONALIZED INSTRUCTIONAL AID

Technical Field

THIS INVENTION relates to a method and means for providing an expert knowledge base in an instructional aid which in one aspect may be used as a teaching aid and in particular as a teaching aid for teaching of movements and techniques, such as used in sports for the game of golf. The present invention, however, is applicable to other movements, sports, pastimes or techniques.

In a further aspect the present invention relates to a method and means for providing an expert visual or audio visual assessment of a particular proposal relating to geographical locations. One such application is to generate a visual display of a particular tract of land for a golf course layout made up from a number of pre stored hole designs or computer generated hole designs which have taken into account all aspects of the final layouts design on the overall development, with the inclusion of audio or audio visual data pertaining to structural methods for success.

The present invention has further application in providing an assessment on a particular object by for example comparing properties of that object with those of a preferred object.

Background Art

As golf has become an increasingly popular pastime over the years, many players wishing to improve their own personal techniques, have pursued professional assistance. Often players, be they professionals or amateurs, may wish to emulate successful movements of top players or coaches who are exposed in the mass-media. At the present time, these players have a number of choices available to them in attempting to try to emulate such movements, which may include general practice, private coaching, viewing current audio visual instruction tapes, recording their own personal movements and analyzing same, or studying photographs/slides etc. of their swing in action.

In all the above cases, the onus is placed upon the

individual or coach, to perceive those areas of movement that differ between the current technique of the player and that of the preferred technique they are trying to achieve. Even for a specialized person working in their own chosen sport or profession, such differences in movement can be so minute or hard to detect, with say a person's weight shift or pressure in their grip or arms, that they can go undetected and so handicap that persons potential rate of improvement. The reason for this particularly in the above instructional visual teaching cases is that the viewer watches information flowing to him in one direction only regarding the current or preferred technique and thus the view cannot perceive all aspects of the movement because of no corresponding physical forces or certain performance data to relate specifically with because such visual presentations do not show their current technique along with the preferred technique with the addition of bio mechanical information. Also in the use of current general prerecorded instructional audio visual presentations they direct their information towards a mass audience as a whole without any allowance for a persons particular technique or requirement of problem solving.

In other situations where conceptual instructional advice is required, the quality of advice can vary, for example, when a person is considering whether or how to best construct a particular tract of land he owns or is considering acquiring, say for the development or remodelling of a golf course, normally expensive preliminary surveying, design and feasibility works are carried out, to which he is reliant on all aspects of the development being expertly covered, to which this is not always the case. Similarly, if alternative ventures in particular areas are proposed, no audio visual means are presently available which provides a customized or personalized completely regenerated, expertly compiled computer analysis in the visual or audio visual form of potential works, described within.

Summary of the Invention

The present invention aims to provide in one aspect an

expertly analyzed instructional aid employing visual or audio/visual techniques which allows for a flow of information in both directions between say selected coaches and pupils so as to help that pupil/viewer improve his or her personal movements and techniques more quickly and without as much guesswork on his or her part.

The present invention also provides a method and means for teaching or instructing such movements or techniques in precise ways, that enables a person/mechanism to more clearly emulate another movement. Whilst the present invention is particularly suited in one application to sports coaching it may also be applied to other areas where an emulation of a precise movement/procedure is required, as will become more apparent from the following operating description. The present invention may additionally be applied to providing expertly compiled opinion in the form of a visual or audio/visual overview of a particular geographical location altered in accordance with selected criteria.

With the above, and other objects in view, the present invention provides in a first preferred aspect, a method for providing an instructional aid for assisting a person to emulate a preferred movement including the steps of capturing and storing visual image signals representative of a particular movement, comparing said captured image signals with stored image signals representative of a selected preferred movement, and regenerating further visual image signals based on said image signals of said particular movement adjusted to emulate said stored image signals of said selected movement.

In a second aspect, the present invention provides apparatus for providing an instructional aid for assisting a person personally to emulate a preferred movement or technique, said apparatus including means for capturing visual image signals of a particular movement or technique, means for storing said captured visual image signals, means for storing visual image signals of a plurality of preferred movements or techniques, means for comparing said captured visual image signals with the stored image signals of a selected said

preferred movement or technique, and means for regenerating further visual image signals of said particular movement or technique adjusted in accordance with said selected preferred movement or technique to assist a person to emulate said selected preferred movement or technique. The term "image signals" as used herein includes signals representative of such images such as digital information which can be converted in to visual images.

Preferably the visual images of the particular movement are captured and stored in digital form using known video recording techniques such as with a video camera which produces a video cassette containing the images of said particular movement. Alternatively other means may be provided for capturing and storing the visual images as described further below.

The visual images of the preferred movements are preferably stored in digital form in a database of a computer so that a selected preferred movement can be readily accessed. The computer is most preferably located at a remote location so as to be accessible from a number of different locations.

Preferably means are provided to transmit the captured images in digital form for receipt by the computer. The transmitting means preferably transmits the images over the normal telephone network, however, other communications links may be employed.

The comparison in the computer can be carried out using either software techniques or hardware techniques and various sampling procedures. The computer is also preferably software or hardware programmed to take into account the size of the player and other physical characteristics such as length of arm so that a compensation factor may be included whereby the movements such as swings of the player and the preprogrammed stored movements or swing in the computer can be truly compared. This comparison can take place at different times or positions in say a swing to produce signals representative of the difference between the swing. This comparison signal may comprise a measurement of distance,

angle, speed of swing or other parameters. Such signals may then be employed in regenerating the new visual image which approximates the captured image altered to emulate the preferred image.

Whilst the invention described within has particular applications to teachings in sport, so as to help and assist a person emulate a particular movement or technique, it may also be applied to many other situations and fields where an emulation or overview of a precise movement is required. Thus, the invention may be applied to various applications in the arts fields, for example the teaching of dance steps or within the medical field where precise movements are required to be studied, an expert advice given such as with human body movements.

Additionally the present invention may be applied to fields where expert systems appraisals and emulations of a development structure may be required such as in the property field for precise calculations and designs of future development proposals or where expert appraisal of a particular object is required by comparison with a preferred or ideal object.

The present invention thus provides in a further aspect a method of providing an instructional aid relating to a development proposal comprising the steps of capturing and storing visual image and location data signals of a proposed development location, analyzing said captured signals against the signals of a selected or preferred known development or means for forming a new design development and regenerating further visual image/data signals based on said signals of said proposed development adjusted in accordance with said preferred development or said newly formed design development so as to provide a recommendation for said preferred development.

In a further aspect the present invention provides apparatus for providing an instructional aid based on a visual image of a particular movement or location, said apparatus including means for visually scanning said movement or location and for producing visual image signals

representative of said movement or location, means for storing said visual image signals, means for transmitting said visual image signals to a remote location, means for comparing said transmitted visual image signals with corresponding signals of a selected prestored image or images, means for regenerating further image signals comprising said image signals of said particular movement or location adjusted in accordance with said prestored signals of said image or images, and means for retransmitting said regenerated image signals and converting said image signals into a format which enables viewing thereof.

In yet a further aspect, the present invention provides apparatus for providing a personalised instructional aid in a development presentation based upon a visual image signal of a particular location, said apparatus including means for visually scanning said location and for producing image and informational data signals representative of said location, means for storing said visual/data signals, means for transmitting said visual/data signals to a remote location, means for comparing said visual/data signals with the signals of a selected prestored location and the structural procedure relating thereto visual/data signals, means for regenerating further visual data/signals based on said signals of said particular location adjusted in accordance with said prestored procedure signals to assist in said development, means for adding dialogue and information data to said regenerated signals and means for transmitting said regenerated image/data signals and converting said signals into a format which enables viewing thereof.

In another aspect, the present invention provides apparatus for providing a personalised instructional aid based upon optical photometric characteristics of a selected target, said apparatus including means for visually scanning, sensing and capturings visual/data signals representative of said target, means for storing said visual/data signals, means for transmitting said visual/data signals to a remote location, means for comparing said visual/ data signals with corresponding prestored signals of a selected target, means for

regenerating further visual data/signals comprising said signals of said particular target adjusted in accordance with said prestored preferred target signals, means for transmitting said regenerated image/data signals and converting said image signals into a format which enables viewing thereof.

Preferably the characteristics of the target which for example may be an object are sensed using imaging and non-imaging sensor systems. Preferably also means are provided for adding dialogue or informational data to be regenerated format.

Brief Description of the Drawings

Reference will now be made to the accompanying drawings which illustrate a preferred embodiment of the invention and wherein:-

Fig. 1 illustrates in block diagram the instructional aid according to the present invention applied to the game of golf; and

Fig. 2 is a flow chart showing the procedure involved in the invention.

Detailed Description of the Preferred Embodiment

Fig. 1 illustrates the apparatus for use in the present invention described below in reference to emulation of a golf swing but usable in any situation where emulation of movement is required. The apparatus 10 includes a visual image capturing device 11 which is adapted to capture movement of the person undertaking a golf swing for subsequent analysis and correction as described below with reference to Fig. 1 and the flow chart of Fig.2.

To capture the initial detailed visual images relating to physical characteristics including movement techniques, employed during the golf swing various scanning or electronic collecting procedures are used. For this purpose the image capturing device 11 preferably incorporates or comprises a portable electronic camera such as a video camera 12 or the like having an associated recording or signal storage means 13 as is known in the art to form and store signals representative of the characteristics of movement etc. being viewed.

Where recording of movements tasks place, bio

mechanical information/ data, may also be transmitted or collected upon the same storage means 13 contained within the framework of the portable video camera 12, so as to enhance further final specialized teaching or information features. Such bio-mechanical information/data may be transmitted to the recording or storage means 13 by an electronic means, hard wiring or other communication means, either at the same time, as the captured visual signals relating to scanned images of the person being viewed or at a different time. In the embodiment illustrated, bio-mechanical information may be captured using a weight sensing pad or pads 14 which will sense weight transference during the swing.

The means used to collect and transmit such bio-mechanical information/data, is at all times non- restrictive in its overall performance, so as to allow the pupil being analyzed to achieve normal free flowing movement without restriction. One such means may take the form of mechanical device/s affixed, or attached, to the person or golf club moved by the person so as to show data such as pressures transferred through the arms or hands, proper releasing techniques, velocities achieved etc. Alternatively or additionally, sensing devices maybe positioned alongside, below or above the person whose image is being captured. Such sensing devices as above may comprise weight or load sensitive pads, monitors electronic light beaming device or the like. Bio-mechanical, signals may also be obtained by various scanning, editing and digital techniques, used within a remote computer database as described further below, once the signal has been received by the remote database, to determine also the many characteristics of the person being studied.

The image capturing device 11 is preferably of portable construction, and of high recording quality, so as to scan, receive and store images and information/ data, in broadcast quality. These visual or information signals are converted by sensors or the like, into electrical charge patterns etc, which are read out or scanned to produce, digital electronic signals representative thereof, from electrical or light signals,

produced by an electronic pick up tube, charge-coupled device (CCD), or the like means within the camera framework, or by the information signalling devices referred to above.

For initially capturing the necessary signals for analysis, more than one image capturing device 11 may be used, so as to channel information regarding various visual details, information/data or the like, to one recording or storage medium. In this regard various switching techniques may also be used to switch between selected image capturing devices 11.

The signal storage or recording means may take the form of Magnetic Tape/Disk, Laser Based or Optical Video Disk, Magneto-Optic Disk, Magnetic-Laser-Optical Disk or other storage medium. Alternatively digitised signals representative of the scanned images or bio mechanical and information data may be also transferred or stored onto a computer memory chip or the like, held within the camera framework.

The apparatus 10 also includes an audio/visual signal transceiver 14 adapted to receive and transmit audio/visual signals and other information signals in the manner described below.

The image capturing device 11 referred to above in one form is a unit discrete from the audio/visual signal transceiver 14 and adapted to be connected thereto or the two maybe one and the same, so that the image capturing device 11 also incorporates a unit for transmission and reception of signals to and from remote databases.

The audio/visual signal transceiver 14 may take the form of a portable signal unit or a personal computer (PC), which can convert, compress or suitably accept, a down loading of captured signals, stored within the storage means 13 of the portable camera unit and then reform, record or store such signals in a suitable digital state for later transmission over a communications network 15. The communication network 15 may comprise the general switched telephone network, satellite links, television quality cable links or the like over which the signals representative of the persons movement together with bio-mechanical information where stored are directed as

electrical or light signals to a remote computer 16 for storage in a computer database 17 for later analysis against selected preferred signals also stored in a database of the computer or a further computer. As is conventional a modem or other similar signal conversion device may be associated with the computer 16 to receive the transmitter signals and convert those signals into a suitable form for reading by the computer 16. The modem or other similar signal conversion device may also be employed for transmission of signals from the computer 16 as described further below.

Initial capturing of audio, visual and data signal information, may also be undertaken using the audio visual signal transceiver 14 (if not a personal computer) to which a video camera 12 may be directly connected so as to enable signals captured by the camera 12 to be recorded directly onto suitable storage medium held within the transceiver 14. The transceiver 14 is then able to transmit these captured signals to remote locations, when directly connected to the communications network described above. Thus a single unit may perform all functions of recording, transmitting and receiving audio, visual and data signal information. The transceiver 14 for this purpose may include dual signal storage facilities, to achieve the above functions, however the transceiver 14 may also be constructed so as to enable all functions of recording and playback in each mode within a single signal storage device.

This transceiver 14 is software and hardware programmed and may offer some of the following features:

1) The senders details for later personalizing effects to the final video presentation i.e. name and spoken language and means for selecting the person with whom the golf swing of the person is to be compared;

2) The transmitting parties/agents details for account keeping purposes and security for the system i.e. digitised security code;

3) Circuitry for increased transmission speed;

4) Circuitry and displays for showing details regarding

signalling times to reach database and regeneration times once retransmitted back to pupil/agent remote location, for account keeping purposes;

5) Circuitry and displays for showing format of transmitted or requested signals in i.e. PAL, NTSC or SECAM etc;

6) Circuitry for Time, Date, Auto Dial - Redial, Stop, Start, Receive, Send and possibly an Advertising channel etc; and

7) Circuitry and displays for showing to whom the regenerated signals belong and to enable the retransmitted regenerated signal to be added to the already prerecorded signal, described below, if desired.

As stated above, the computer 16 is arranged normally at a remote location from the audio/visual transceiver 14 and includes a database 17 which stores selected prerecorded video signals in digital form which in the case of golf swings may include a selection of swings showing the movement of different golf professionals which are required to be emulated. As the computer 16 and image capturing device 11 are not physically affixed or attached together, the computer 16 and database 17 may be of a sufficient size and capacity to enable the desired functions to be achieved at one and the same location.

The computer 16 may be at many remote locations and holds within its database the following prerecorded information in digital form, which provide a basis for forming the regenerated audio visual signals.

1) A digitised audio visual library of selected coaches, advisers professional or mechanisms (such as a golf simulation machine) movements, techniques, steps or procedures, that have been previously analysed, recorded and reformed into suitable digitised signals, for later analysis and regeneration in the database against the remote signals received.

2) A digitised bio mechanical library of selected coaches or mechanisms movements measured individually or as a group, for later display with the regenerated audio/visual signals.

The computer 16 incorporates a means for receiving and

editing the transmitted visual data or bio-mechanical and information data signals , relating to the movement of the golfer and arranges same in such a way as to be able to match the received visual signals with the previously established and selected database signals of their preferred applications and generate a new video signal to enable the viewer to see personally, in regenerated form, those differences that exist in (a) a teaching mode, between the two movements, that of the pupil and the selected professional or coach and (b) in an analysis mode, those ideas that might be possible for their preferred swings (or applications in development fields where the method and apparatus of the invention is used in this application.) For this purpose, the computer is software and/or hardware programmed to compare the respective swings of the person being analyzed and the selected preferred swing with adjustments made such as for size of person, speed of swing and other parameters to allow a correct comparison. The differences between the swings are measured by sampling or other techniques and this difference added or subtracted from a persons recorded swing and new visual images generated on this basis which emulates the preferred movements.

The computer 16 may also include a means for adding dialogue to the regenerated audio visual signal recordings, so that the viewer can receive personalized instruction or advice from afar, similar to having private instruction or personal consultation.

The computer 16 may further include a means for converting received signals, if needed, in one broadcasting mode, into those of another countries i.e. PAL-NTSC-SECAM etc. To accomplish this within the computer, a Standards Converter or the like may be used, using Optical Scan or Digital Standards features or the like.

The computer 16 may additionally include a means for determining costing on each and every received and regenerated signal that passes through the system, based upon the following information:-

(a) Signal transmission time to and from the computer.

(b) The amount of analysis and regeneration time required, for each and every presentation, such as with the showing of audio visual faults, bio mechanical/characterized data graphs, dialogue and broadcasting changes etc.

The computer 16 is software and hardware programmed, to accept digitised signals transmitted to it, from remote locations, associated with the audio visual transceiver 14, so as to personally relate in the final regeneration and editing phases to the viewer, this being performed by, using various digital editing techniques and procedures as described below.

The audio visual presentation generated by the computer 16 can take the form of split-screening (generating new visual images of the preferred technique displayed along side the present technique), superimposing (where the preferred technique is superimposed over the present technique) or other display. Graphics and charts etc, for bio-mechanical/ information displays, so that the viewer can clearly see those changes that are required or need to be performed, to develop or pursue the preferred movement, procedure or proposal are also displayed. These superimposing effects or the like, may take the form of stick figures, detailed drawings etc, to normal human or visual appearances, so as to show the viewer any variances between his stroke and the preferred stroke. For ease of editing the subject material within this computer database, a time code or the like, maybe also burnt into the original recorded signal, so as to make final editing much more efficient within the remote database. This procedure may also be linked with an edit controller or the like, within the computer 16, which locks the two signals electronically together during edits.

To personalize the regenerated signal, dialogue is also added when required, so that the selected instructor or adviser may relate to the viewer/pupil more personally, than with other audio visual presentations. Part of this personalizing effect is initially accomplished in the original transmission phase, by the name, spoken language, broadcasting mode or location

statistics etc of the final viewer being captured, within this database, to which an addition of dialogue can also be added when required, in the regeneration signal phase, showing those changes they need to perform or consider, to obtain the preferred movement, thus totally personalizing the new regenerated signal to the viewer.

For this purpose such dialogue, may only take the form of a brief statement, so as to make the viewer fully aware of those changes that are required, or need their consideration, to perfect the movement, procedure or development etc.

The means for adding such dialogue may take the form of a voice energy monitor or the like, contained within the computer, which on analysis scans a memory bank of learned words or phrases to find suitable terms, to bring attention to these problem areas or matters needing their attention.

Dialogue may also be added relating to recommendations for the viewers consideration in regards to, suitability of equipment being used to perform the desired movements and techniques.

To increase acceptance as a true teaching aid, scientific data, which in many applications could only be assumed before, such as with personalized coaching, may be presented to a viewer, when needed, in the form of graphs, charts, or the like, set either above, below, alongside or within the total regenerated video signal recording. This information is gathered from their received signals, being matched to, their preferred established signals contained within the database 17 so that the person/pupil can clearly see, those bio mechanical differences that exist between, their current procedures/techniques and that of their selected coaches/professionals procedures/techniques.

Once the new audio visual signal is regenerated the signal is then retransmitted back by the computer 16 and via a modem or the like and the communications network 15 to the remote senders/agents transceiver 14. This procedure maybe performed, from within the computer 16 itself, using various digital, compression and sampling techniques, or it can be

again sent back by a further audio/visual transceiver associated with the computer 16.

On receipt at the remote pupils/agents receiving location, the regenerated signal may already be in suitable form, for direct saving on suitable storage means such as video tape, disk or the like which enables home viewing such as by means for a video cassette player and playback through a monitor 18. If not in a suitable form for direct saving, the signal may be reformed for storage as above on storage means which permits home viewing other than through a computer monitor by the use of either the receiving audio visual transceiver, or by other suitable conventional means, offering various slowing down or speeding up signal procedures. If desired the regenerated audio/visual signal may be stored on the original storage means eg. video tape containing the players current movements or techniques or stored on a new storage means.

In the embodiment described and shown, the image capturing device 11 stores the captured images on a video cassette 19 of known form and the regenerated images are stored on the same cassette 19 or alternatively an additional cassette. The regenerated image in this embodiment is in the form of a stick figure 20 which is superimposed on the true visual image 21 of the players swing or movement and which indicates to the player the angles and positions which he must adopt to emulate the swing.

To add further to this advanced audio visual teaching aids value, as a complete instructional system, compared to other audio visual presentations, previously recorded instructional material, of a non-personalized nature, may also be added, to the final home storage presentation means such as a video cassette. Thus this regenerated transmitted signal of the pupil's audio visual presentation and data pertaining to movements, maybe then simply added to the prerecorded material already held at the senders/agents remote location. In this format the single recording may then show (a) an initial complete recorded procedure of instruction, as currently

presented with audio visual presentations (b) an addition of their current movements, procedures or images etc, taken before regeneration and (c) the final regenerated personalized audio visual changes, that are required, or affect, the desired changes needed by the viewer.

A typical presentation with a preferred format and structure in the final video tape or other viewable means could be as follows :

Pupil: Mr Rick Baker

Selected coach : Mr Bobby Jones

Applications for: Golf

Dialogue in: English

1. Preferred audio visual start/introduction:

"Hi, Rick, I am Bobby Jones, it is my pleasure to have studied your golf swing and these are the points I feel you should be concentrating on : "

2. Preferred audio visual presentation in progress:

During the presentation whilst the pupil, Mr Baker is viewing the subject matter, the viewing monitor is showing segments of the previously recorded swing, taken at the recording facility, with superimposed preferred movements of Mr Jones against Mr. Baker's current movements and pointing out with the addition of dialogue, those areas that he should be concentrating on, so as to improve his current technique.

Bio mechanical data in the form of graphs or charts etc. are also added during the presentation, so as that he and his local coach may also quite clearly see scientific data pertaining to those performed movements. Dialogue may also be added regarding suitability of equipment being used at this stage.

3. Towards the conclusion of this audio visual presentation, the coach could conclude by saying:

"Thank you Rick for allowing me to help you improve your game, but please continue to work on these points with the help of your local professional, who is also there to assist you in overcoming these problems. Good golfing,

Bobby Jones."

The above embodiment has been described with reference to a golf swing movement, however, as previously stated the invention may readily applied where other movements are required to be emulated. Such movement may comprise movements of persons or movements of objects or mechanisms of which emulation is required.

In a further form the invention may be applied to the property field to enable professional assessments and presentations of designs for future development proposals or the like. In this form visual images and data relating to signals of a particular area of ground are captured using remote sensing techniques. For this purpose a camera and remote sensing device may be fixed to the underside of an aircraft so as to capture and record geometrical or statistical information of the earth or objects featured beneath. Geographical data achieved through using this remote sensing as above may be computer coded or digitised using video cameras, array scanners, laser beam, microwave or other means.

Information regarding the geographical area being scanned is then transmitted as before using a transceiver to a remote computer containing a database of prestored expert information on attributes required of a particular area of land. Thus where the best standard of golf course is required to be laid out upon a specific parcel of land the data base may include a library of selected hole layouts from different courses. The computer which is software programmed may then select preferred hole layouts to match to the area of land of which geographical features has been scanned and recorded. Alternatively, the computer may itself design the appropriate hole layouts to best fit the area of land and proposed development. The computer database may additionally include feasibility study means for determining and displaying measurements, costings and other procedures required to achieve the end result. The computer may further be programmed to take into account residential housing around the course and the location of same, so as to maximize view of the course etc.

As before, dialogue may be added to the regenerated

audio visual presentation as well as statistical graphs or other information relating to the development being studied.

The present invention may also be applied to enable a particular object to be analysed against prestored data relating to a currently known object. In this embodiment and as before an object is scanned so that its characteristics may be determined. Scanning is carried out to obtain the optical photometric characteristics of the object using imaging and non-imaging sensor systems. As before the data obtained is stored and then transmitted to a remote location for analysis in a computer with the corresponding data of a preferred object whose data information is captured in a data base. After comparison and adding, where appropriate, suitable dialogue from say an expert in that field, a regenerated signal is formed and transmitted back to the sender so that the sender has an expert assessment of the differences between the respective objects. A particular application in this embodiment is for component parts as used in an assembly. The data information obtained relating to that part may be compared to a perfect component part whose data is stored at the remote location so that the differences may be assessed. This embodiment may also be employed for detecting for example faults in a particular component part.

It will be apparent that the latter embodiment may be applied to any suitable targets of which analysis is required.

The present invention thus provides a method and means for expertly teaching or instructing procedures, strategies, assessments or perceptions in audio visual form to enable a viewer to more closely emulate calculate or develop techniques suitable to their various applications.

Whilst the above has been given by way of illustrative embodiment of the invention, all such modifications and variations thereto as would be apparent to persons skilled in the art are deemed to fall within the broad scope and ambit of the invention as herein defined by the appended claims.

CLAIMS

1. A method for providing an instructional aid for assisting a person to emulate a preferred movement including the steps of capturing and storing visual image signals representative of a particular movement, comparing said captured image signals with stored image signals representative of a selected preferred movement, and regenerating further visual image signals based on said image signals of said particular movement adjusted to emulate said stored image signals of said selected movement.

2. A method according to Claim 1 wherein said image signals of said particular movement are captured and stored using a video camera.

3. A method according to Claim 1 wherein said visual image signals of said preferred movements are stored in a database of a computer and wherein a selected preferred movement is accessed from said data base for comparison with said particular movement.

4. A method according to Claim 3 wherein said image signals are captured and stored at a first location and wherein said computer is located at a second location remote from said first location and wherein means are provided for transmitting said captured images from said first location to said computer at said second location.

5. A method according to Claim 4 wherein said transmitting means comprise independent signal transmitting means or a personal computer.

6. A method according to Claim 4 wherein said transmitting means transmits said image signals over a communication network to said computer.

7. A method according to Claim 4 wherein said regenerated signals are transmitted from said remote second location to said first location.

8. A method according to Claim 7 wherein said regenerated image signals are stored in image presentation means which permits viewing thereof by said person.

9. A method according to Claim 8 and wherein dialogue relating to said regenerated image signals is provided within said image presentation means.

10. A method according to claim 7 wherein bio-mechanical data relating to said regenerated images is displayed within image presentation means.

11. A method of providing an instructional aid relating to a development proposal comprising the steps of capturing and storing visual image and location data signals of a proposed development location, analyzing said captured signals against the signals of a selected or preferred known development or means for forming a new design development and regenerating further visual image/data signals based on said signals of said proposed development adjusted in accordance with said preferred development or said newly formed design development so as to provide a recommendation for said preferred development.

12. A method according to Claim 11 wherein said visual image/data signals of said proposed location are captured and stored using a video camera and remote sensing devices.

13. A method according to Claim 10 wherein said signals of said known developments are stored in a database of a computer.

14. A method according to Claim 11 wherein said computer is arranged at a location remote from said proposed location and wherein said captured image signals are transmitted to said

remote location for analysis by said computer.

15. A method according to Claim 14 and wherein said regenerated image/data signals are transmitted from said computer to said proposed location and are provided in a format to enable when transferred or added to home viewing means viewing thereof.

16. A method according to Claim 15 and including means for adding dialogue to said format so as to provide an audio visual presentation.

17. Apparatus for providing an instructional aid for assisting a person personally to emulate a preferred movement or technique, said apparatus including means for capturing visual image signals of a particular movement or technique, means for storing said captured visual image signals, means for storing visual image signals of a plurality of preferred movements or techniques, means for comparing said captured visual image signals with the stored image signals of a selected said preferred movement or technique, and means for regenerating further visual image signals of said particular movement or technique adjusted in accordance with said selected preferred movement or technique to assist a person to emulate said selected preferred movement or technique.

18. Apparatus according to Claim 17 wherein said means for storing visual image signals of a plurality of preferred movements or techniques comprises a computer database and wherein said comparing and regenerating means comprises a computer associated with said computer database.

19. Apparatus according to Claim 18 and including transceiver means for transmitting to said computer said visual image signals of said particular movement or technique and for receiving from said computer said regenerated visual images signals.

20. Apparatus according to Claim 18 and including means for converting said regenerated image signals into a format which allows viewing thereof.

21. Apparatus according to Claim 20 and including means for adding to said format dialogue relating to said regenerated images.

22. Apparatus according to Claim 20 and including means for adding to said format bio-mechanical data relating to said regenerated images.

23. Apparatus for providing an instructional aid based on a visual image of a particular movement or location, said apparatus including means for visually scanning said movement or location and for producing visual image signals representative of said movement or location, means for storing said visual image signals, means for transmitting said visual image signals to a remote location, means for comparing said transmitted visual image signals with corresponding signals of a selected prestored image or images, means for regenerating further image signals comprising said image signals of said particular movement or location adjusted in accordance with said prestored signals of said image or images, and means for retransmitting said regenerated image signals and converting said image signals into a format which enables viewing thereof.

24. Apparatus for providing a personalised instructional aid in a development presentation based upon a visual image signal of a particular location, said apparatus including means for visually scanning said location and for producing image and informational data signals representative of said location, means for storing said visual/data signals, means for transmitting said visual/data signals to a remote location, means for comparing said visual/data signals with the signals of a selected prestored location and the structural procedure relating thereto visual/data signals, means for regenerating

further visual data/signals based on said signals of said particular location adjusted in accordance with said prestored procedure signals to assist in said development, means for adding dialogue and information data to said regenerated signals and means for transmitting said regenerated image/data signals and converting said signals into a format which enables viewing thereof.

25. Apparatus for providing a personalised instructional aid based upon optical photometric characteristics of a selected target, said apparatus including means for visually scanning, sensing and capturings visual/data signals representative of said target, means for storing said visual/data signals, means for transmitting said visual/data signals to a remote location, means for comparing said visual/data signals with corresponding prestored signals of a selected target, means for regenerating further visual data/signals comprising said signals of said particular target adjusted in accordance with said prestored preferred target signals, means for transmitting said regenerated image/data signals and converting said image signals into a format which enables viewing thereof.

26. Apparatus according to Claim 25 and including means for adding dialogue or informational data to said regenerated format.

27. Apparatus according to Claim 25 wherein said target scanning means comprise imaging or non-imaging sensor systems.

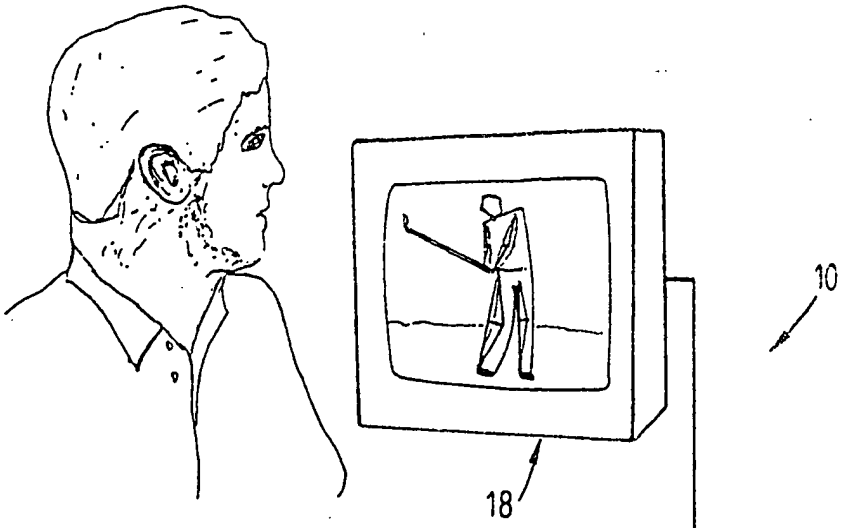
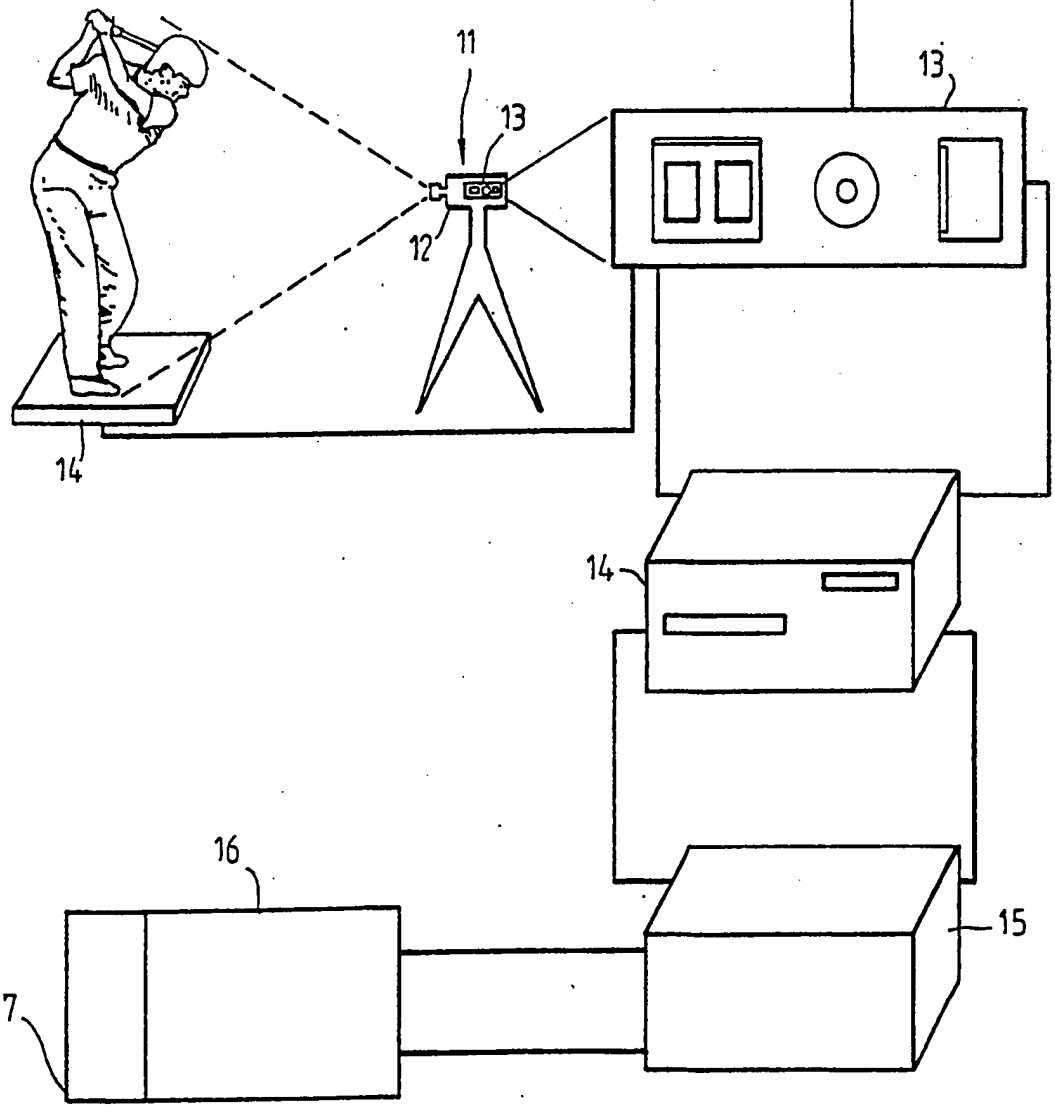


FIG.1



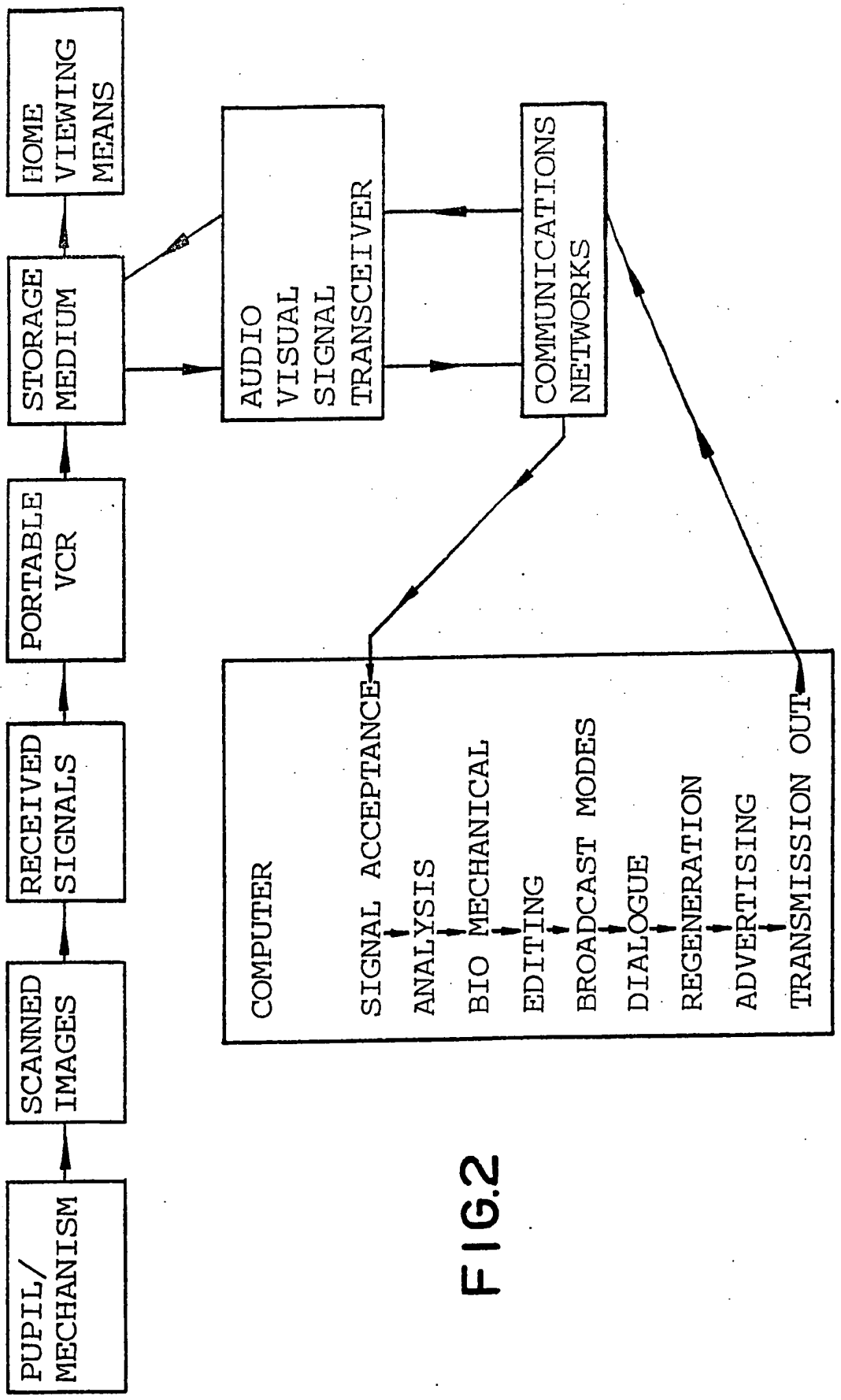

PUPIL/ MECHANISM
SCANNED IMAGES
RECEIVED SIGNALS
PORTABLE VCR
STORAGE MEDIUM
HOME VIEWING MEANS
AUDIO VISUAL SIGNAL TRANSCEIVER
COMMUNICATIONS NETWORKS
COMPUTER
SIGNAL ACCEPTANCE
ANALYSIS
BIO MECHANICAL
EDITING
BROADCAST MODES
DIALOGUE
REGENERATION
ADVERTISING
TRANSMISSION OUT

FIG.2

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.[5] A63B 69/36, G06F 15/62

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC A63B 69/36, G06F 15/62

US Cl. 434/252

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

AU : IPC as above

Electronic data base consulted during the international search (name of data base, and where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate of the relevant passages	Relevant to Claim No.
A	Patents Abstracts of Japan, C-823, page 79, JP,A, 03-26281 (OYO KEISOKU KENKYUSHO K.K.) 4 February 1991 (04.02.91)	(1-10,17-23)
A	Patents Abstracts of Japan, C-818, page 87, JP,A, 03-12180 (MIZUNO CORP) 21 January 1991 (21.01.91)	(1-10,17-23)
A	US,A, 4942391 (KIKUTA) 17 July 1990 (17.07.90) See the abstract	(1-27)
A	US,A, 4852184 (TAMURA et al) 25 July 1989 (25.07.89) See the abstract; column 5, lines 37-47	(11-16,23,24)

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 September 1992 (11.09.92)	18 Sept 1992 (18.09.92)

Name and mailing address of the ISA/	Authorized officer
AUSTRALIAN PATENT OFFICE PO BOX 200 WODEN ACT 2606 AUSTRALIA	R FINZI for J.W. THOMSON
Facsimile No. 06 2853929	Telephone No. (06) 2832214

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate of the relevant passages	Relevant to Claim No.
A	Derwent Abstract Accession no. 89-056750/08, Class WO2, JP,A, 01-007860 (TAKARA KK) 11 January 1989 (11.01.89)	(23-25)
A	AU,A, 15932/88 (ASC DEVELOPMENT CORPORATION) 6 October 1988 (06.10.88) See the whole specification	(1-27)
A	US,A, 4137566 (HAAS et al) 30 January 1979 (30.01.79) See column 4, lines 12-25; Figure 1	(1-10,17-23,25)
A	US,A, 3408750 (McCOLLOUGH et al) 5 November 1968 (05.11.68) See column 1, lines 25-60; column 2, lines 14-18	(1-10,17-23)

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report	Patent Family Member					
US 4942391	JP	1059568				
US 4852184	DE	3736932	FR	2606244	JP	63113676
	KR	9100591				
US 4137566	CA	1119723	GB	2005548		
AU 15932/88	JP	2502790	US	4828500	WO	8807735
						END OF ANNEX

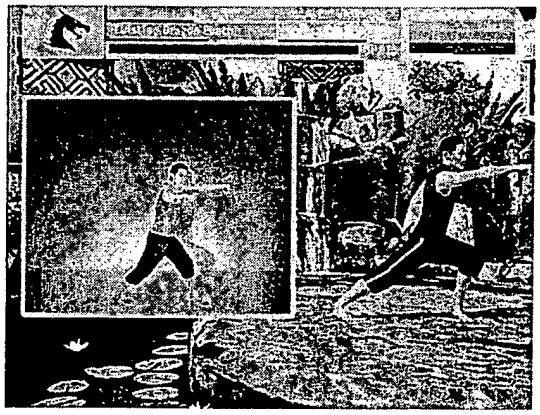
EyeToy: Kinetic Combat – Review

By **Andrew Fletcher** on January 3, 2007 - 5:31pm.

Last year, I called Sony's fitness title *EyeToy: Kinetic* a quiet revolution in game design, and this martial arts based sequel certainly serves as a further curiosity tweak to the faintly odd notion of sports games—that is, *real* sports games—flourishing in future console generations. On this evidence, that revolution will boil down to design specifics as much as it will persuasive concepts.

Bravely, *EyeToy: Kinetic Combat* veers off from the 'all-round' fitness theme of *Kinetic* and attempts to teach players the basics of Hung Gar Kung Fu. Contrary to what other lazy reviewers (and I suppose I mean that literally) have said, the fitness aspect is really now a fringe benefit rather than the focus. Sony's London Studio should be applauded for testing themselves with such a peculiar and unprecedented goal as teaching a specific martial art, but sadly *Kinetic Combat* is not the inspirational software it could have been.

The core teaching is done through tutorials in which the player must copy the onscreen trainer's movements, performed on one side of the screen as a game graphic and on the other as a semi-transparent shadow over the player's own EyeToy image. The idea is that one matches up movements to the trainer's ghostly outline, which is shaped according to the height and body build descriptions players input at the start of the game. It sounds simple enough, but sadly the game's main flaws can be extrapolated from that one key screen.



First of all, the screen is horribly cluttered. The time and effort the gamer spends paying attention to the instructor's movements and comments, trying to fit inside and in time with the shadow, and checking the *five* different performance feedbacks (red/green motion evaluating blocks; red/green/amber rep. blocks; current grade letter; current performance evaluation; pointless line graph) leaves him or her even less chance to concentrate on the power and technique that is supposedly at the core of this martial art. The new instructor doesn't rescue the presentation either, being painfully dull in both

Similar Game Reviews/Articles
- Eyetoy: Kinetic - Review
- EyeToy: Play - Review
- EyeToy: Play - Second Opinion
- Eye Toy: Groove - Review
- Eyetoy: Kinetic - Consumer Guide

Recent Video Game Reviews
- Dead Rising 2: Case West Second Opinion
- Metal Gear Solid: Peace Walker Review
- ilomilo Review
- Splatterhouse Review
- Dead Rising 2: Case West Review
- Enslaved: Odyssey to the West Second Opinion
- Raskulls Review
- 3D Dot Game Heroes Second Opinion
- 999: Nine Hours, Nine Persons, Nine Doors Review
- Majin and the Forsaken Kingdom Review

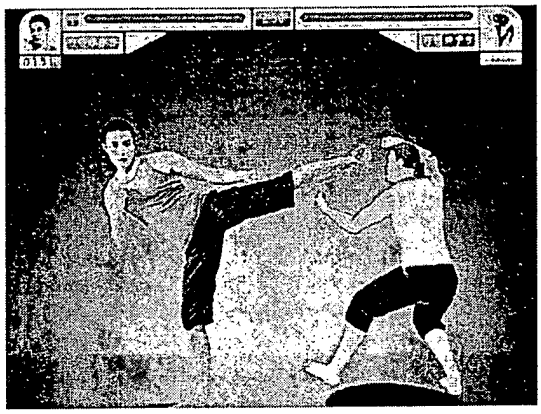
more

Recent Blog Posts
- Consoleation: December 2010 console gaming sales data
- Want to know what's wrong with the PlayStation 3?
- Consoleation: Street un-smart
- The Horror Geek presents: Official launch trailer for Dead Space 2
- A matter of pacing: Red Dead Redemption

appearance and (largely unhelpful) dialogue, which affects the game more than you'd think.

The restricting decision to judge players on their ability to match the onscreen trainer's movements in real time (like a kind of unfair, rhythm-action 'Simon Says...') seems to misjudge the *Kinetic* audience. Whereas the more flexible and forgiving first game praised players for their achievements, here they are immediately told that they're falling short of a perfect standard. Yet that evaluation is itself undermined since the onscreen shadow against which the player is tested is almost useless to anyone with less than four or five metres between themselves and the TV. It doesn't help that the standard Hung Gar stance is particularly low, and dipping to try and fit the trainer's shadow is uncomfortable and potentially unsafe.

Many will quickly tire of being told that their movements are inaccurate just because of their spatial imposition and the software's inability to compensate. I have a perfectly reasonable room size, yet *Combat* remained inflexible in allowing me to fit my onscreen tutor's shadow. Besides, even though the motion sensing technology is sound in concept (disregarding its execution), it quickly becomes dull performing the endless reps that *Combat* asks without any sense of freedom or fun. Perhaps that's a personal opinion about the appeal of martial arts training, but at the end of the day *Combat* has in no way endeared it to me. Being tutored on technique by an impossibly wooden game character is simply more tedious and less game-like than the high energy exercise of *Kinetic*.



Outside of direct tutorship, however, there are training games and sparring fights, which recall the last title's exhausting freeform workouts and are all the more entertaining for it. Their effectiveness varies and the games will ask players to perform moves that they haven't even learnt yet, but they still feel beneficial for only suggesting moves, and most of all just for being enjoyable.

It is crucial that *Kinetic* games make me *want* to rearrange my furniture and work up a sweat, but whereas the original was at pains to present a workout experience as fun, comfortable and considerate as possible, *Combat* can make me feel like I'm hooked up to performance-testing lab machines.

Kinetic, in a flawed and even haphazard kind of way, taught us that with the right kind of guidance and encouragement, a pupil can be helped to learn at their own rate and successfully so. *Kinetic Combat* teaches us that dry didacticism and overwhelming, in-your-face feedback merely stifle the process. If games are going to start educating us (as the recent *Brain Training*, *TalkMan* and *Kinetic* games suggest they would like to), then they're going to have to be more fun, personable and accessible than other methods of learning, not less so.

Best Game Review Sites

- Game Rankings
- Metacritic
- Zero Punctuation
- IGN
- GameSpot
- GameFAQs
- GameTrailers
- Kotaku
- G4 TV
- GamesRadar
- 1up.com
- GameSpy
- Joystiq
- Giant Bomb
- Destructoid

- Video: Have Princess Peach and Zelda fallen for their captors?
- Amnesia woes, Case West fumbles and Preacher rocks
- PlayStation 2sDay: Tony Hawk's Pro Skater 3
- Organization XIII: How bosses should be made
- GameCritics.com Podcast Episode 47 Transcript

more

Recent Podcasts

- GameCritics.com Podcast Episode 47 Transcript
- GameCritics.com Podcast Episode 47: Third Annual Holiday Awards Spectacular
- GameCritics.com Podcast Episode 46 Transcript
- GameCritics.com Podcast Episode 46: World of Warcraft: Cataclysm, Epic Dungeon, 2011 Gaming Resolutions
- GameCritics.com Podcast Episode 45 Transcript

more

Recent comments

- PUT A TV out on it and I am
  5 hours 11 min ago
- still not sold on it.
  8 hours 30 min ago
- Not impressed. Graphics
  11 hours 41 min ago
- I am not at all interested
  12 hours 30 min ago
- EDF worst 360 game ever?
  13 hours 34 min ago
- Second choice
  17 hours 15 min ago
- Picked the 2nd choice. I
  1 day 12 hours ago
- I liked Silent Hill but this
  1 day 22 hours ago
- but
  2 days 6 hours ago
- Blame update process
  3 days 4 hours ago

Forum Discussions

- Uncharted Waters Online
- Needing Approval to Post
- Ep 47: CONTEST!
- DS vs PSP
- PS3 Mass Effect 2
- Pan's Labyrinth - Film
- Uncharted 3: Drake's Deception Release Date, Trailers, Screenshots, Gameplay Previews
- Smartphones as a gaming platform
- Scott Pilgrim vs The World
- DC Universe Online anyone?

As for Hung Gar Kung Fu: try taking some evening classes, and quit those instead. [Rating: 5.0 out of 10]

Disclaimer: This title is currently only available in EU territories with no US release currently slated.

‹ Previous Index EyeToy: Kinetic Combat – Consumer Guide ›

» printer-friendly version | add new comment | ShareThis

Category Tags
Platform(s): **PS2**
Developer(s): **Sony London**
Publisher: **Sony**
Series: **EyeToy**
Genre(s): **Weird Music/Dance**
Articles: **Game Reviews**

Code of Conduct
Comments are subject to approval/deletion based on the following criteria:
1) Treat all users with respect.
2) Post with an open-mind.
3) Do not insult and/or harass users.
4) Do not incite flame wars.
5) Do not troll and/or feed the trolls.
6) No excessive whining and/or complaining.

Please **report any offensive posts** here.

For more video game discussion with the our online community, become a member of our **forum.**

Our Game Review Philosophy and Ratings Explanations.

About Us | Privacy Policy | Review Game | Contact Us | Twitter | Facebook | RSS

more

Game Tips Forum

- Dead Space 2 Tips Achievements FAQ Help Tricks Hints Cheat Codes Strategy Guide
- Two Worlds 2 Tips Achievements FAQ Help Tricks Hints Cheat Codes Strategy Guide
- Lord of Arcana Tips Trophies FAQ Help Tricks Hints Cheat Codes Strategy Guide
- Exorcist Tips Achievements Trophies FAQ Help Tricks Hints Cheat Codes Strategy Guide
- Paws and Claws: Marine Rescue Tips FAQ Help Tricks Hints Cheat Codes Strategy Guide
- Dino Strike Tips Achievements FAQ Help Tricks Hints Cheat Codes Strategy Guide
- Breach Tips Achievements Trophies FAQ Help Tricks Hints Cheat Codes Strategy Guide
- World of Tanks Tips Achievements FAQ Help Tricks Hints Cheat Codes Strategy Guide
- Jam City Rollergirls Tips FAQ Help Tricks Hints Cheat Codes Strategy Guide
- Rocks n' Rockets Tips Achievements FAQ Help Tricks Hints Cheat Codes Strategy Guide

more

User Submissions

- Please rate this review: Enslaved: Odyssey to the West
- Rate this 'article': Video Game Maps (i)
- Zippy reviews NEIR
- Please rate this review: Call of Duty: Black Ops
- Please Rate This Review: Singularity 2nd Opinion

more

Tweets by Staff Members
GameCritics.com

trentfingland So at work there's a resident who puts cheese in her coffee. Is this a thing that I'm unfamiliar with? It seems too deliberate to be random.
40 minutes ago · reply

horrorgeek Will Orson Welles' Lost Film Finally Be Seen? **http://aol.it/fVCPrr**

Virtual PAT: A Virtual Personal Aerobics Trainer

James W. Davis Aaron F. Bobick
MIT Media Lab
20 Ames Street, Cambridge, MA 02139
jdavis@media.mit.edu, bobick@media.mit.edu

Abstract

A prototype system for implementing a virtual Personal Aerobics Trainer (PAT) is presented. Unlike workout video tapes or TV exercise shows, this system allows the user to create and personalize an aerobics session to meet the user's needs and desires. Various media technology and computer vision algorithms are used to enhance the interaction of a virtual instructor by enabling it to watch and talk to the user. Throughout the paper, we report the system components and discuss the advantages, problems, and extensions of the design.

Figure 1. Virtual PAT. A virtual Personal Aerobics Trainer. Photo credit: Webb Chappell. Copyright: Webb Chappell 1998.

1. Introduction

In this paper we discuss the design and implementation of a prototype virtual Personal Aerobics Trainer (PAT). The underlying motivation for building such a system is that many forms of media that *pretend* to be interactive are in fact deaf, dumb, and blind. For example, most of the aerobics workout videos that one can buy or rent present an instructor that blindly expels verbal re-enforcements (e.g. "very good!") whether or not a person is doing the moves (or is even in the room!). There would be a substantial improvement if the TV just knew whether or not a person was moving in front of it. A feeling of awareness would then be associated with the system. And because of the repetitiveness of watching the same exercise videos, this "programmable" system heightens the interest of the user by allowing the design of specialized workouts (e.g. exercising only the upper body).

The system here creates a personalized aerobics session for the user and displays the resulting instruction on a large TV screen. Here the user can select which moves, their duration, the music, and which instructor are desired for the workout. Depending on the mood of the user, the choice of instructor could make a large difference in the workout. For instance, if the user were tired and required strong motivation during the workout, a brash Army Drill Sargent as the instructor might be an appropriate choice. Along those lines, the prototype system here makes available an Army Drill Sargent character as the instructor. The session is scripted by the user, and is then automatically generated and started when the user enters the area in front of the TV screen (See Figure 1).

In addition to the video demonstrations, the user periodically receives audio feedback from the virtual instructor on how he/she is currently doing. We place a video camera (or multiple cameras) in the room and use real-time computer vision techniques to recognize the aerobic movements of the user. Using the output of the vision system, the virtual instructor then responds accordingly (e.g "good job!", if the vision system recognizes that the user is performing the aerobic move). This vision technology is different from many other sensing technologies in that the user need not wear

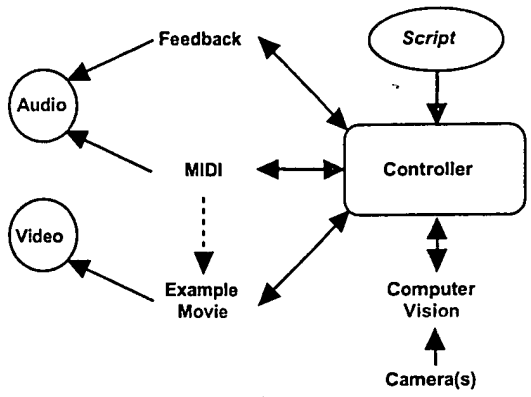


Figure 2. Component diagram showing the interconnection of the media, control, and computer vision modules within the PAT system.

any special clothing/devices or be tethered to machines with bundles of wires. This enables the experience to be more natural and desirable [2, 9]. We now explain the design of the system.

2. System design

The PAT system is a modular design of media and vision components connected to a controller (See Figure 2). All software[1] was written in C++ and run across SGI R10000 O2 computer systems (though we believe the components could be placed within a lower-end hardware setup). For the current system, the output video (a series of movie clips showing the instructor performing the moves) is sent to a large TV screen, the feedback is routed through the audio channel, and the music is stored/played in the form of MIDI music files.

Because the current version of the system uses real video clips it would be tedious to record all possible feedbacks during all possible moves. Therefore, the audio is decoupled from the video (i.e. the lips of the instructor do not move, as if speaking the lines). The separating of audio feedback from the visual example permits the simple combination of various spoken feedback phrases and multiple aerobic examples. It would be possible to integrate the spoken feedback and the visually displayed by using a computer graphics model of the instructor, rather than using movie clips. The correct combined state of the instructor (e.g. doing jumping jacks while complementing the user) would be controlled and rendered at run-time. It might even be quite fun to have virtual cartoonish-like characters as the instructors. Each character could have their own "attitude" and behavior [1], which would possibly increase the entertainment value of the system. But in the current system, we chose to use stored movie clips for simplicity. Also, to date we have only recorded one instructor though the system is designed to have multiple instructors from which the user can choose.

When personalizing (or creating) the aerobics workout session, the user need not worry about specific timing information of the moves, feedback, or music to coordinate the different media events or structure of the session. These components each have underlying temporal information (e.g. movie length, tempo of music) which is communicated between different modules. The user need only manipulate a simple script of the workout session.

We now give a brief explanation of each of the components in the system. The components residing on different machines have a direct communication line to a main control program using the Parallel Virtual Machine (PVM) software designed for heterogeneous network computing.

2.1. Scripting

Since most instructional systems employ some underlying notion of event ordering, we can exploit this to allow the user to create and structure a personalized session. The system was designed so that each session is guided from a simple script which controls the flow of the session (similar to [6]). Included in the script are the names for the workout moves, the time allotted for each move, the choice of music (song titles) for the workout, and the instructor (by name) to run the session. This allows the user to easily choose their own tailored workout. Placing the available options and examples in a GUI would make the script creation process quite trivial, but for now it is created in a text editor. The controller loads this script and initiates the program. The program is immediately available upon instantiation of the system with a script (it is not generated or compiled off-line).

2.2. Controller

A simple state-based controller was developed to run the workout script and act as a central node connecting the various modules. The controller consists of seven states: Pause, Startup, Introduction, Workout, Close, Shutdown, and PreClose. The system begins in the Pause state, where it resides until a person enters the space, determined by querying the vision module. The entering person event triggers Startup state which

[1] All media components were developed using SGI's Digital Media utilities/libraries.

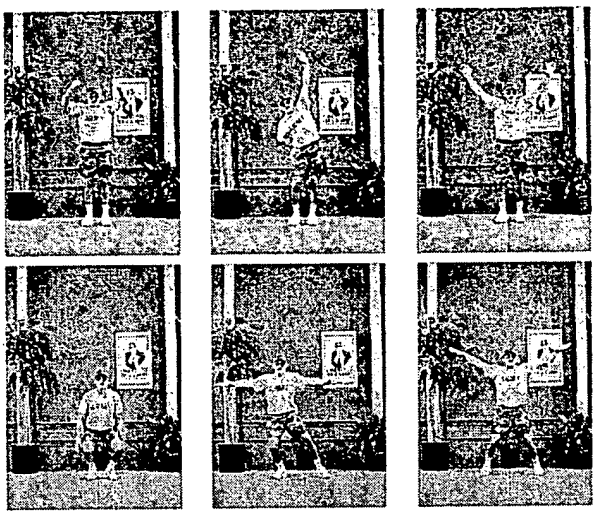


Figure 3. Video output of the virtual instructor.

opens the video window and performs some system preparation. Next is the Introduction state, where the virtual instructor appears and welcomes the user. After the brief introduction, the system loops in the Workout state (a loop for each move in the session) until all moves are completed. Feedback comments from the instructor are given here (comment categories are determined by the recognition results from the vision module). Then a Close state gives the final goodbye comments from the instructor, followed by the Shutdown state where the display is turned off and then system cleanup is initiated. There is an additional PreClose state which is entered if the user prematurely leaves the space (recognized by the vision module). Here, the instructor realizes the user is no longer present, and then starts a shutdown (or pause) of the system (the program will not continue if no one is there to participate).

2.3. Instruction by example

In the TV display is a window showing the instructor performing the aerobic moves. Figure 3 shows a few key frames from some stretches and aerobic exercises. This is provided so the user can follow along, watching how to perform the moves, as done with a real instructor.

Currently, a set of movie clips, showing a full view of the instructor, is used. In each clip (for each move), the instructor performs a single cycle of the move. This clip is looped (or swinged) for the duration of the time devoted for that move. Given the beat track of the music, the speed of the current movie clip can be altered to be in synchronization with the music currently being played. (It can be quite annoying for the video/user to be moving out of synchronization with the music if there is a strong beat.) We now discuss a music format that can be used to synchronize the example movies with the music.

2.4. MIDI music and synchronization

Music is used in aerobics primarily to regulate the speed and consistency of the moves as well as to maintain the energy of the exerciser. We currently use MIDI music files/songs as our source of music: a MIDI file is a symbolic description of the music similar to the sheet music or score for a song. By connecting a synthesizer and speakers to a computer (some computers provide an internal software synthesizer), the song can be played by sending the MIDI file from the computer to the synthesizer which then generates the music. One advantage of using MIDI is that it is easy to alter the music by simply inserting, deleting, or changing individual MIDI events. Another important feature of MIDI is that there is a tempo associated with each song.

To suit our purposes, we would like to synchronize our instructor's movements from the movie clips to the beat of the music. Using the MIDI format, we can add a silent beat track — a metronome click — using the tempo of the song. The beats act as a trigger to which the instructor movies can be synchronized. The movie cycle duration is set to be equal to the time required for a fixed number of beats, and the movie is looped/swinged after a certain beat count.

The beat track is added as an additional track to the MIDI file and is set at a non-perceivable volume. This way, the computer can watch for the beat events while not actually playing audible notes. We note that there exists methods for accomplishing the above beat track synchronization using other musical media forms (e.g. a CD); such a method is found in [8].

At times, there can be a bit of a jump when looping the movie clips in synchronization with the beat track of the music. The actual time *measured* between beats in a song and also the response time of the movie adjustment can sometimes vary slightly. This causes the looping to become minutely jagged at the transition for the next loop. More precise control and reading of the beat track can fix this problem. Another synchronization problem can arise if there is substantial video buffering (or lag) from the video generator to the display unit, causing the virtual instructor to appear as dragging behind the beat. Due to the above problems, we built two systems where one version synchronizes the video and music and another version where the two

are unsynchronized (the music was carefully selected in this case).

The MIDI songs for the workout are all queued up at the beginning of the session so one song plays immediately after another. At the end of the session, the music is stopped.

2.5. Feedback on performance

Each time a new aerobic move begins, the feedback module is activated and a brief statement about the new move is given. For some moves, the comment may give the name of the movement (e.g. "It's time for some jumping jacks") or for other moves explain their purpose (e.g. "This move is going to work the shoulders"). For the feedback to the user the system contains many *positive* comments (e.g. "good job!", "fantastic!") and many *negative* feedback comments (e.g. "get moving!", "concentrate!"). Whenever the controller decides it is time for a feedback comment[2], it randomly picks a comment from the appropriate category. This way, one does not always here the same comment many times in a row or hear the same ordering of comments. There is an opportunity here to record very expressive comments for the system, which increases the entertainment value of the system as well as its usefulness.

3. Recognizing aerobic movements with computer vision

In this section, we give a brief explanation of the methods used for recognizing the movements of the user. We use real-time computer vision techniques to "watch" the user and determine if he/she is currently performing the same move as the instructor. A background-subtracted image of the room yielding only a silhouette of the person [4, 2, 9] is used as the input to the vision system.

3.1. Robust silhouette extraction technique

The silhouetting method used in this system is based on the optical blocking (or eclipsing) of specialized non-visible light (e.g. infrared light) rather than the color differences between the person and background. In our approach, the person is always able to wear any color clothing, and the end effect is the same (unlike the color-based methods). The method also permits the display of video graphics behind the user on the background. Figure 4 shows the environment configuration. This system is similar in nature to the infrared designs of [7, 5].

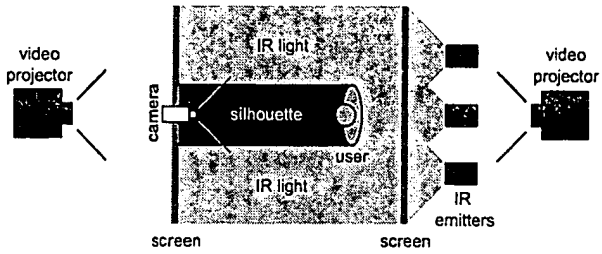


Figure 4. Conceptual drawing of blocking (eclipsing) infrared light from the camera to generate a silhouette of the person.

We can see the silhouetting process in Figure 5. Figure 5(a) shows a standard camera view of someone standing in front of the back-wall projection screen with graphics displayed. In Figure 5(b), we see the same picture but now with infrared light directed through the back screen using six infrared light (840 nm) emitters[3]. By placing an infrared-pass/visible-block filter over the video camera lens, a brightly lit screen (with no graphics visible) and a clean silhouette of the user is seen, as shown in Figure 5(c). The infrared light is not visible to the human visual system and thus one sees only video projected on the display screen (as shown in Figure 5(a)). We point the reader to [4] for further details on the approach. This silhouette form is then further processed to recognize the movements of the user.

3.2. Motion Templates

Recently, we have developed real-time computer vision methods for recognizing large-scale body movements such as aerobic exercises. That work constructs temporally-collapsed motion templates of the participant's silhouette, and measures shape properties of that template to recognize various aerobic exercise (and other) movements in real-time.

To show an example of such motion templates, Figure 6 presents templates generated from the infrared silhouettes for the movements of left-arm-raise (left-side stretch) and fan-up-both-arms (deep-breathing ex-

[2]The system checks every movement cycle to see if the user is complying. A *negative* comment is given every cycle until the user performs the move. If the user is performing the move, a *positive* comment is given approximately every few cycles or some predetermined wait period (whichever is smaller).

[3]The camera has no infrared-blocking filter and is thus sensitive to infrared light.

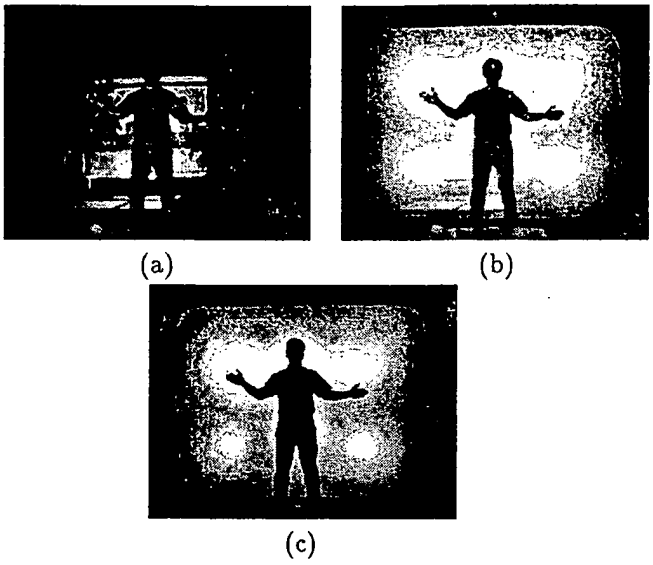


Figure 5. Infrared light. (a) The person in front of a video projection screen. (b) Now with infrared light directed through the screen. (c) The image filtered through a visible-block filter.

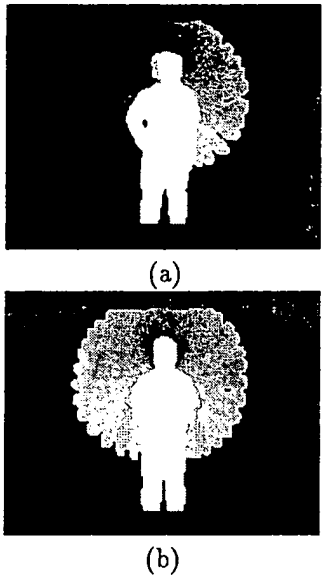


Figure 6. Example motion templates for infrared silhouettes. (a) Motion template for left-arm-raise (left-side stretch). (b) Motion template for fan-up-both-arms (deep-breathing exercise stretch).

ercise stretch). The motion of the user is encoded in the varying graylevels within the template.

For recognition of these moves, statistical pattern recognition techniques are applied to moment-based feature descriptors of the templates. The system employs user training to get a measure of the variation which arise from different people. This approach easily extends to multiple camera views of the person; see [3] for details on the algorithm.

4. Future possibilities

Presently, the system is designed for a single user. To increase the usefulness of the system, it could be designed to run with multiple people (as in a class). We could accomplish this by spreading out the people inside the room so the cameras could get an unobstructed view of each person. The system could also periodically switch its focus from one person to the next (by toggling camera views) to save in processing. However, better vision techniques that allow occlusions from multiple people would be desirable. Also having the system understand and reason about other objects (e.g. weights, towels) would increase its awareness and expressiveness of its environment.

To extend the sensing and robustness of the system as a whole, we could incorporate other types of sensors. For instance, we could run a trigger wire across the top of an exercise step (used for step aerobics). Then it would be easy to verify whether the person was in fact on the step. With this, we can also check the periodicity of the stepping and verify synchronization of the user with the instructor. There are also wireless heart-rate monitors available which can transmit their data to remote computers. With this accessory, the user can have his/her heart-rate displayed on the TV screen while progressing through the workout session. A program could be written to check this information and watch for harmful levels of strenuous activity.

The large video screen behind the user, while necessary for the infrared silhouetting procedure, can also be used to enhance and make the experience more rich and immersive to the users. For example, we could add virtual people to the back screen to make the user feel as if taking part in a group workout. There is also an opportunity here for putting entertainment on this screen for benefit of an audience, rather than just the user. For instance, it may be humorous for an audience (or those waiting their turn) to watch the virtual class toying with the user (unbeknownst to the user, of course). If additional video screens were employed, then we could place the user entirely within some virtual world.

Another video window could be shown (adjacent to the instructor window) which contains a live video feed of the user as seen from a camera mounted near the TV screen. This video mirror (with the video stream reflected) would serve the purpose of giving a visual aid to the user involved in the current session. The person can watch themselves and see how they look while performing the moves. We could also augment the video with graphics to better aid in responding to the user (e.g. drawing a straight line down the user's back to emphasis to the user to keep his/her back straight).

Lastly, a record of past sessions for an individual could be kept (e.g. retaining the moves the person performed, and how well he/she executed those moves). With this past information available, the system could be designed to respond with more informed comments like, "Your doing much better today!". This capability would move the PAT system into the highly personalized interaction category.

5. Conclusion

In this paper, we described a prototype system for a virtual Personal Aerobics Trainer (PAT). Users can select which aerobic moves, which songs, and which instructor they want for their workout session. The resulting program guides the user through the workout while watching and commenting on the user's performance.

The system presents various movie clips of the instructor, which can be ordered and synchronized with the music. A computer vision system watches the user throughout the session, and reports to the virtual instructor how the user is doing. With this information, comments are given by the instructor for pushing or complementing the user.

This system moves beyond the highly non-interactive media forms of video tapes and TV shows by watching and responding to the user. We feel that many future exercise systems will be more interactive and less passive, and perhaps one day will be commonplace within the home environment.

6. Acknowledgments

We first would like to especially thank Andy Lippman for playing out the role of the Army Drill Sargent. His sense of fun made that character come to life. We would also like to thank those people who provided their time sweating for us by training up the system. This work is supported in part by the Digital Life consortium of the MIT Media Laboratory.

References

[1] Blumberg, B. Old tricks, new dogs: ethology and interactive creatures. PhD dissertation, MIT Media Lab, MIT, 1996.

[2] A. Bobick, S. Intille, J. Davis, F. Baird, L. Cambell, Y. Ivanov, C. Pinhanez, A. Schutte, and A. Wilson. The KIDSROOM: Action recognition in an interactive story environment. *Presence (to appear).*

[3] Davis, J. and A. Bobick. The representation and recognition of human movement using temporal templates. In *Proc. Comp. Vis. and Pattern Rec.*, pages 928–934, June 1997.

[4] Davis, J. and A. Bobick. A robust human-silhouette extraction technique for interactive virtual environments. In *IFIP Workshop on Modelling and Motion Capture Techniques for Virtual Environments*, Geneva, Switzerland, Nov. 1998.

[5] Ishii, H., and B. Ullmer. Tangible bits: towards seamless interfaces between people, bits and atoms. In *CHI'97*, pages 234 – 241, 1997.

[6] C. S. Pinhanez, K. Mase, and A. F. Bobick. Interval scripts: A design paradigm for story-based interactive systems. In *CHI'97*, pages 287–294, Atlanta, Georgia, Mar. 1997.

[7] Rekimoto, J. and N. Matsushita. Perceptual surfaces: towards a human and object sensitive interactive display. In *Workshop on Perceptual User Interfaces (PUI-97)*, pages 30 – 32, October 1997.

[8] Scheirer, E. Tempo and beat analysis of acoustic musical signals. *Acoustic Society of America*, 103(1):588–601, Jan 1998.

[9] Wren, C., Azarbayejani, A., Darrell, T., and A. Pentland. Pfinder: Real-time tracking of the human body. In *SPIE Conference on Integration Issues in Large Commercial Media Delivery Systems*, 1995.

Advanced Search

Related Keywords (1)

computer vision

Related Paper (5)

Kidsroom: Action recognition in an interactive story environment

Tai Chi synthesizer: a motion synthesis framework based on key-postures and motion instructions

Virtual Human Physiotherapist Framework for Personalized Training and Rehabilitation

Motion segmentation and pose recognition with motion history gradients

The KidsRoom: An example application using a deep perceptual interface

Academic Publication Virtual PAT: A Virtual Personal Aerobics Trainer

Edit

Virtual PAT: A Virtual Personal Aerobics Trainer (Citations: 38)

Aaron F. Bobick, James W. Davis

A prototype system for implementing a virtual PersonalAerobics Trainer (PAT) is presented. Unlikeworkout video tapes or TV exercise shows, this systemallows the user to create and personalize an aerobicssession to meet the user's needs and desires. Variousmedia technology and computer vision algorithms areused to enhance the interaction of a virtual instructorby enabling it to watch and talk to the user. Throughoutthe paper, we report the system components and discussthe advantages,...

Published in 1998.

View or Download

(vismod.www.media.mit.edu)

Citation Context (3)

...Related work on VTs can be found in, among others, [18], where a physiotherapist is described with similar functionality as ours, [13] with an interesting Tai Chi application, and [2], reporting about work on an aerobics trainer...

— Dennis Reidsma, et al. Challenges for Virtual Humans in Human Computing

...Davis developed a vision-based motion training system, called Virtual PAT (Personal Aerobics - Trainer) using IR light sources providing manually prerecorded instructive videos and audio feedback [7]...

— Doo Young Kwon, et al. Combining body sensors and visual sensors for motion training

...Other related vision-based interfaces include an interactive playspace for children (The KidsRoom) [4, 5], and a reactive virtual exercise trainer [10]...

— James W. Davis, et al. A perceptual user interface for recognizing head gesture acknowledgeme...

Reference (6)

Tempo and beat analysis of acoustic musical signals (Citations: 274)

E. Scheirer

Published in 1998.

A Robust Human-Silhouette Extraction Technique for Interactive Virtual Environments (Citations: 21)

James W. Davis, Aaron F. Bobick

Conference: Modelling and Motion Capture Techniques for Virtual Environments - CAPTECH, pp. 12-25, 1998

Pfinder: Real-Time Tracking of the Human Body (Citations: 297)

Alex Pentl, Ali Azarbayejani, Christopher Wren, Trevor Darrell

Conference: Storage and Retrieval for Image and Video Databases - SPIE, 1995

The Representation and Recognition of Human Movement Using Temporal Templates (Citations: 114)

James W. Davis, Aaron F. Bobick

Conference: Computer Vision and Pattern Recognition - CVPR, pp. 928-934, 1997

Tangible bits: towards seamless interfaces between people, bits and atoms (Citations: 1080)

Hiroshi Ishii, Brygg Ullmer

Conference: Computer Human Interaction - CHI, pp. 234-241, 1997

Citation (38)

SENSITIVITY ANALYSIS OF HAND MOVEMENT CLASSIFICATION TECHNIQUE USING MOTION TEMPLATES

Sanjay Kumar, Dinesh K Kumar, Arun Sharma

Published in 2010.

Design of Experience and Flow in Movement-Based Interaction

Anton Nijholt, Betsy Van Dijk, Dennis Reidsma

Published in 2008.

Shadow reaching: a new perspective on interaction for large displays (Citations: 6)

Garth Shoemaker, Anthony Tang, Kellogg S. BoothConference: User Interface Software and Technology - UIST, pp. 53-56, 2007

Nonverbal and Bodily Interaction in Ambient Entertainment (Citations: 2)

Anton NIJHOLT, Dennis REIDSMA, Zsofia RUTTKAY, Herwin van WELBERGEN, Pieter BOS
Published in 2007.

Performative surface: double sided interaction

Hugues Bruyère, Thierry Giles
Conference: ACM Multimedia Conference - MM, pp. 493-496, 2007

USING HUMAN BODY GESTURES AS INPUTS FOR GAMING VIA DEPTH ANALYSIS

Yong Wang, Tianli Yu, Larry Shi*, Zhu Li**

Motorola Labs, Motorola Corporation, Schaumburg, IL, 60196, USA

ABSTRACT

Natural ways of input greatly enhance the entertainment experience for emerging gaming systems represented by the Sony Playstation 2 EyeToy and the Nintendo Wii Console. In this paper we present a novel method of using human body gestures depth image as gaming application input. Depth images have natural advantages over grayscale or color images in terms of robustness against illumination change, texture complexity, and background interference. Our proposed method consists of three major components: depth image acquisition, mean shift based preprocessing, and HMM-based gesture recognition. We validate our method by applying it to a boxing game scenario to distinguish boxing gestures such as dodge, jab, hook, and uppercut. The experiment results indicate that our method can efficiently distinguish the subtle differences among these gestures and yield excellent accuracy (up to about 98%). The potential usage of the proposed method on gaming applications and generic human computer interaction is very promising.

Index Terms— Video Gaming, Human Computer Interaction, Stereo Image, HMM

1. INTRODUCTION

Natural ways of input, especially body gesture-based input, not only simplify the interaction between a computer and a human, but also greatly enhance the entertainment experience for emerging gaming systems represented by Sony Playstation 2 EyeToy and Nintendo Wii Console. Gesture recognition based on stereo depth video is still an active area of research, despite the substantial body of work addressing the issue of understanding human gestures leveraging computer vision and pattern recognition techniques [1]. Most research focuses on recovering the human body pose using depth image [2, 3]. These body poses can then be used for action recognition[4]. Active sensing technologies such as structured lighting [2] or time-of-flight sensor [5] can also be used to capture the necessary depth map, with both advantages (accurate depth, less computation) and disadvantages (requires active illumination).

In this paper we present a novel framework for learning human body gesture based on its depth image. Depth images have natural advantages over grayscale or color images in terms of robustness against factors such as illumination change, texture complexity, and background interference. Our proposed method consists of three major components: data acquisition, pre-processing, and gesture recognition. First, a stereo camera is set up to capture calibrated image pairs from which the depth images are computed in real-time. Second, we apply mean-shift based clustering on depth images to eliminate depth noises, separate the background/foreground parts, and extract the gestures of interest represented by feature vectors. Lastly, a Hidden Markov Model (HMM)-based machine learning system is used to classify different gestures. We validate our proposed method by applying it to a boxing game scenario to distinguish boxing gestures such as jab, hook, and uppercut. The experiment delivered excellent results and our method can reliably recognize these gestures with superior accuracy (about 92% overall and up to 98%). The potential usage of the proposed method on gaming applications and generic human computer interaction is very promising.

The rest of this paper is organized as follows. Section 2 describes the details of the proposed system. The experiment results are presented in Section 3. Section 4 discusses our conclusion.

2. LEARNING GESTURE FROM DEPTH MAP

Fig. 1 provides an overall diagram of the proposed system, which comprises of three steps: acquisition, preprocessing, and recognition.

2.1. Stereo Input Acquisition

We used a calibrated stereo camera to capture the stereo video of users performing gestures. Each captured stereo pair was first corrected for geometric and photometric distortion. The correspondence problem between the two images was formulated in the Markov Random Field framework with L1 smoothness prior, and optimized globally (over the entire image) using belief propagation [6, 7]. The result is a depth (or disparity) map that measures how far away each pixel is with respect to the camera. Thanks to recent algorithmic and hardware advances [8, 7], the depth computation can be finished in real-time during capturing.

*: This work was done when these authors were at Motorola Labs.

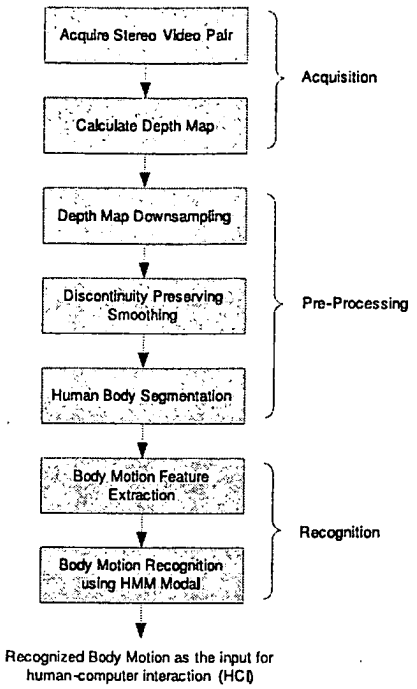


Fig. 1. Diagram of the proposed system

2.2. Preprocessing of Depth Map

Preprocessing is necessary in order to reduce the level of noise in the depth map and to extract reliable features for gesture recognition.

2.2.1. *Depth Map Downsampling*

The first preprocessing step is to downsample the obtained depth map in order to reduce complexity for real-time processing. Though more advanced downsampling approaches can be considered, our experiments indicate that an image sized down to 32 by 24 pixels suffices to yield satisfactory performance.

2.2.2. *Discontinuity Preserving Smoothing*

Depth images from the stereo algorithm usually contain noise that would considerably influence feature extraction performance. Therefore the depth images need to be pre-processed by discontinuity-preserving smoothing before they are fed into the down-stream processing modules. The purpose of this operation is to smooth out the variance in the pixels having close enough depth levels (statistically these pixels usually come from a semantically meaningful object, such as a human body), while keeping as still distinguishable the pixels having large enough depth difference . Specifically this can be done through mean shift filtering [9]. I.e., each pixel $\mathbf{x}_i = \left(\mathbf{x}_i^s, x_i^d\right)$ at location i in the depth map is assigned to a value that is the convergence point $\mathbf{y}_i^c$ of the sequence $\left\{\mathbf{y}_{i,j} = \left(\mathbf{y}_{i,j}^s, y_{i,j}^d\right)\right\}, j = 1, 2, ...,$ obtained iteratively through (note $\mathbf{y}_{i,1} = \mathbf{x}_i$):

$$\mathbf{y}_{i,j+1} = \frac{\sum_{n=1}^{N} \mathbf{x}_n g\left(\left\|\frac{\mathbf{x}_n^s - \mathbf{y}_{i,j}^s}{h_s}\right\|^2\right) g\left(\left\|\frac{x_n^d - y_{i,j}^d}{h_d}\right\|^2\right)}{\sum_{n=1}^{N} g\left(\left\|\frac{\mathbf{x}_n^s - \mathbf{y}_{i,j}^s}{h_s}\right\|^2\right) g\left(\left\|\frac{x_n^d - y_{i,j}^d}{h_d}\right\|^2\right)} \quad (1)$$

where N is the number of pixels in the depth map, $i = 1, 2, ..., N$ is the index of the pixel, $\mathbf{x}_i^s$, x_i^d are the 2-D coordinates and the depth level of the pixel respectively, h_s, h_d are the window width controlling the smoothing intensity, and $\mathbf{y}_{i,j}$ is the mean value weighted by the kernel function $g(x)$, which is defined as:

$$\begin{aligned} g(x) &= -k_E'(x) \quad (2) \\ k_E(x) &= \begin{cases} 1-x & 0 \le x \le 1 \\ 0 & x > 1 \end{cases} \end{aligned}$$

where k_E' is the derivative of k_E.

2.2.3. *Human Body Segmentation*

The smoothed depth map is further segmented so that only the objects of interest are kept. In our proposed method these levels include human body and the hand closest to the camera (if any). For this purpose a linked list is constructed in depth descending order, with depth levels close enough merged into the same entry in the list. The merging criterion is a predefined threshold, usually the window width for depth level used in the smoothing step. Each entry in the list records the depth level and the amount of pixels in the entry. Therefore we have: $\mathbf{L} = \{(d_k, p_k)\}, k = 0, 1, 2, ...K-1, d_i > d_j$ when $i < j$, where K is the number of merged depth levels, d_k and p_k are the depth level and the pixel number for the k^{th} entrance respectively. A classifier for depth level $C(k)$ works as follows. A depth level k is considered for the hand when it is the first depth level in the list and its pixel amount $A(k)$ is less than certain threshold T_h. Similarly the first depth level whose area $R(k)$ is larger than certain threshold T_b is consdiered for the body. All other levels are treated as background or noise and are therefore ignored. I.e.,

$$C(k) = \begin{cases} hand & k = 0 \text{ and } A(k) < T_h \\ body & R(k) > T_b, R(i) \le T_b \text{ when } i < k \\ ignored & others \end{cases} \quad (3)$$

The whole pre-processing procedure is illustrated in Fig. 2.

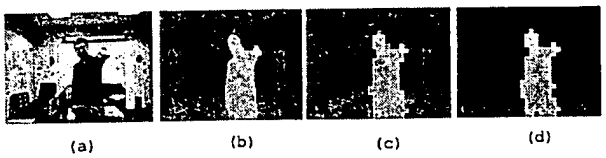


Fig. 2. Pre-processing of depth images. (a): original input frame; (b): depth map; (c): down-sampled map; (d): Segmented map (including both discontinuity-preserving smoothing and human body segmentation)

2.2.4. Body Feature Extraction

The following features are used in our system:

1) Body centroid (2-D):

$$\mathbf{x}_b^s = \frac{\sum_{n=1}^{N} \mathbf{x}_n^s \mathbf{1}\left\{x_n^d \in \text{Depth level of body}\right\}}{\sum_{n=1}^{N} \mathbf{1}\left\{x_n^d \in \text{Depth level of body}\right\}} \quad (4)$$

2) Hand displacement (2-D):

$$\mathbf{x}_h^s = \mathbf{x}_b^s - \frac{\sum_{n=1}^{N} \mathbf{x}_n^s \mathbf{1}\left\{x_n^d \in \text{Depth level of hand}\right\}}{\sum_{n=1}^{N} \mathbf{1}\left\{x_n^d \in \text{Depth level of hand}\right\}} \quad (5)$$

3) Relative depth level of hand (1-D):

$$d_h^r = x_b^d - x_h^d \quad (6)$$

where x_b^d and x_h^d are the depth level (grayscale value in depth map images) of the body and hand respectively.

The full set of the feature vector is therefore represented as: $\mathbf{F} = \left(\mathbf{x}_b^s, \mathbf{x}_h^s, d_r^h\right)$. In order to reduce computational complexity, not all features are used for every gesture. Specifically, for body gestures with hand motion (e.g., hook), a subset of the feature vector $\mathbf{F}_h = \left(\mathbf{x}_h^s, d_r^h\right)$ is adopted; for body gestures without hand motion (e.g., dodge), $\mathbf{F}_b = \left(\mathbf{x}_b^s\right)$ is adopted. This way the dimension of the feature vector can be reduced to up to 3 from 5, which significantly reduces the computational cost for both training and testing. Our experiment indicates that this simplification does not significantly affect the accuracy.

2.3. Body Gesture Recognition

Gaussian Hidden Markov Model (GHMM) is applied on our system to train the model to recognize body gestures. Particularly a four-stage GHMM is applied to each body gesture, where stage 1 represents the moment when the gesture is about to be executed and stage 4 represents the moment when the gesture is fully executed.

In total there are M such models, each for a specific body gesture. Every frame captured at time t_i by the stereo camera is an observation represented as $\mathbf{F}_i$. All observations in certain stage are considered to obey a Gaussian distribution and the parameters can be efficiently estimated through the algorithm of Gaussian Mixture Model (GMM). A video clip containing a gesture is represented by a feature sequence $\mathbf{S} = (\mathbf{F}_0, \mathbf{F}_1, \mathbf{F}_2, ..., \mathbf{F}_{L-1})$, where L is the number of sampling (frames) of the clip. To recognize a gesture $G(\mathbf{S})$ is a classic evaluation problem of HMM [10]:

$$G(\mathbf{S}) = \arg\max_n \left\{P(\mathbf{S}|\lambda_n)\right\}, n = 0, 1, 2, ...N-1 \quad (7)$$

where $P(\mathbf{x})$ is the probability of $\mathbf{x}$, and λ_n is the model for the n^{th} gesture, obtained through a typical HMM learning procedure.

3. EXPERIMENT RESULTS

The stereo camera we used is a PointGrey Bumblebee camera with a 12 cm baseline. The depth image was captured at 10 fps, with a resolution of 512 × 384 and 60 disparities.

Experimental data was collected in an ordinary meeting room with a sofa, a coffee table, and several chairs (i.e., regular furniture in a typical living room). The setup of the room and one sample frame taken by the stereo camera are shown in Fig. 3. Eight subjects with different skin colors, clothes (textures), heights, and weights participated in the experiment. Table 1 lists the boxing actions used in our experiment. Except for defense, each gesture included both left and right side actions. In total there were (M=)9 gestures. Each subject repeated each action for 2 to 6 times. Each action was represented as a feature vector as described in Section 2.3. We used the mean shift code by Chris Christoudias, *et.al.* [11] with $h_s = 32$ and $h_d = 20$.

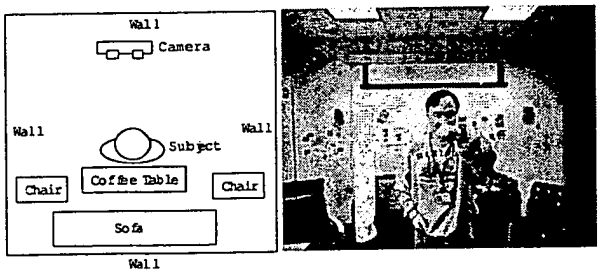


Fig. 3. Experiment setup and a sample frame

After the data was collected and the feature space was constructed, all trials were divided into K=8 subgroups. Thereafter a leave-one-out cross-validation procedure was conducted. K-1 subgroups were training sets and the remaining subgroup is the testing set. This procedure was repeated until each subgroup was used as the testing set. We used the HMM codes by Kevin Murphy [12]. The results are shown in Table 2. It is very clear that our proposed system can

Table 1. Boxing actions used in the experiment

Action Name	Description
Defense	Cross both fists in front of the head
Dodge	Move or shift quickly aside
Dash	straight, arm-length punches thrown from the leading hand
Hook	swing the arm, which is bent at an angle near or at 90 degrees, into the opponent.
Uppercut	A swinging blow directed upward, as to a boxing opponent's chin.

recognize each body gesture efficiently with high precision. The overall accuracy is about 92%. Note that although our subjects vary considerably in skin color, clothes (texture), weight, and height, our system is still robust enough because the depth map is insensitive to these changes.

Table 2. Experiment results

Action Name	Test Trials	Recognized	Accuracy
Defense	40	39	97.50%
Left Dodge	32	30	93.75%
Right Dodge	40	38	95.00%
Left Dash	40	35	87.50%
Right Dash	40	36	90.00%
Left Hook	40	35	87.50%
Right Hook	16	15	93.75%
Left Uppercut	40	34	85.00%
Right Uppercut	48	47	97.92%
Overall	336	309	91.96%

4. CONCLUSIONS AND FUTURE WORK

In this paper we present a novel method of using depth image of human body gestures as inputs for gaming applications. Stereo cameras are adopted to capture stereo image pairs where depth map is retrieved, content features are extracted from the depth map, and HMM model is used to train models recognizing human body gesture. The proposed system is validated through gesture data for a boxing game. Experimental results indicate that our system performs very well and yields excellent precision. The potential usage of the proposed method on gaming applications and generic human computer interaction is very promising.

Future work includes considering more types of gestures to accommodate general gaming (or computer input) environment rather than just those used in a boxing game, and improving the robustness of the recognition by automatically segmenting gestures from a long clip containing various gestures.

5. REFERENCES

[1] Ying Wu and Thomas Huang, "Vision based gesture recognition:a review," in *Proc. 3rd Gesture Workshop*, 1999, pp. 103–115.

[2] M. Bray, E. Koller-Meier, and L. Van Gool, "Smart particle filtering for 3d hand tracking," in *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, 17-19 May 2004, pp. 675–680.

[3] Hee-Deok Yang and Seong-Whan Lee, "Reconstruction of 3d human body pose from stereo image sequences based on top-down learning," *Pattern Recognition*, vol. 40, pp. 3120–3131, 2007.

[4] Eric Huber, "3-d real-time gesture recognition using proximity spaces," in *Third IEEE Workshop on Applications of Computer Vision (WACV '96)*, Los Alamitos, CA, USA, 1996, vol. 00, p. 136, IEEE Computer Society.

[5] Pia Breuer, Christian Eckes, and Stefan Mller, "Hand gesture recognition with a noval ir time-of-flight range camera - a pilot study," 2007.

[6] Pedro F. Felzenszwalb and Daniel P. Huttenlocher, "Efficient belief propagation for early vision," *Int. J. Comput. Vision*, vol. 70, no. 1, pp. 41–54, 2006.

[7] Tianli Yu, Ruei-Sung Lin, Boaz J. Super, and Bei Tang, "Efficient message representations for belief propagation," in *Proceedings of ICCV, 2007*, 2007.

[8] Qingxiong Yang, Liang Wang, Ruigang Yang, Shengnan Wang, Liao Miao, and David Nister, "Real-time global stereo matching using hierarchical belief propagation," in *BMVC, 2006*, 2006.

[9] Dorin Comaniciu and Peter Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[10] Lawrance Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," vol. 77, no. 2, pp. 257–286, 1989.

[11] Chris Christoudias and Bogdan Georgescu, "Mean shift analysis library," in *http://www.caip.rutgers.edu/riul/research/code/.*

[12] Dorin Comaniciu and Peter Meer, "Hidden markov model (hmm) toolbox," in *http://www.ai.mit.edu/ murphyk/Software/hmm.html.*