



US 20180348187A1

(19) **United States**

(12) **Patent Application Publication**  
**Fernstrom et al.**

(10) **Pub. No.: US 2018/0348187 A1**  
(43) **Pub. Date: Dec. 6, 2018**

(54) **METHOD, APPARATUS AND SYSTEM FOR FOOD INTAKE AND PHYSICAL ACTIVITY ASSESSMENT**

(71) Applicant: **University of Pittsburgh - Of the Commonwealth System of Higher Education, Pittsburgh, PA (US)**

(72) Inventors: **John Fernstrom, Pittsburgh, PA (US); Madelyn Fernstrom, Pittsburgh, PA (US); Wenyan Jia, Wexford, PA (US); Robert Sciabassi, Gibsonia, PA (US); Mingui Sun, Pittsburgh, PA (US)**

(21) Appl. No.: **15/994,003**

(22) Filed: **May 31, 2018**

**Related U.S. Application Data**

(63) Continuation of application No. 13/676,468, filed on Nov. 14, 2012, now Pat. No. 10,006,896.

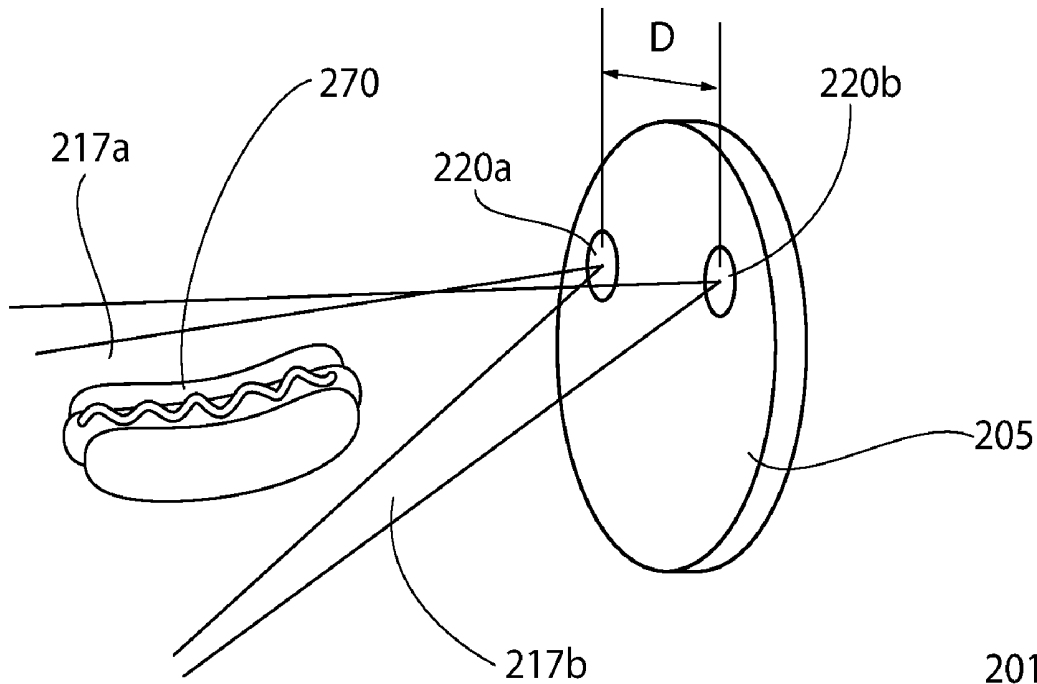
(60) Provisional application No. 61/559,370, filed on Nov. 14, 2011.

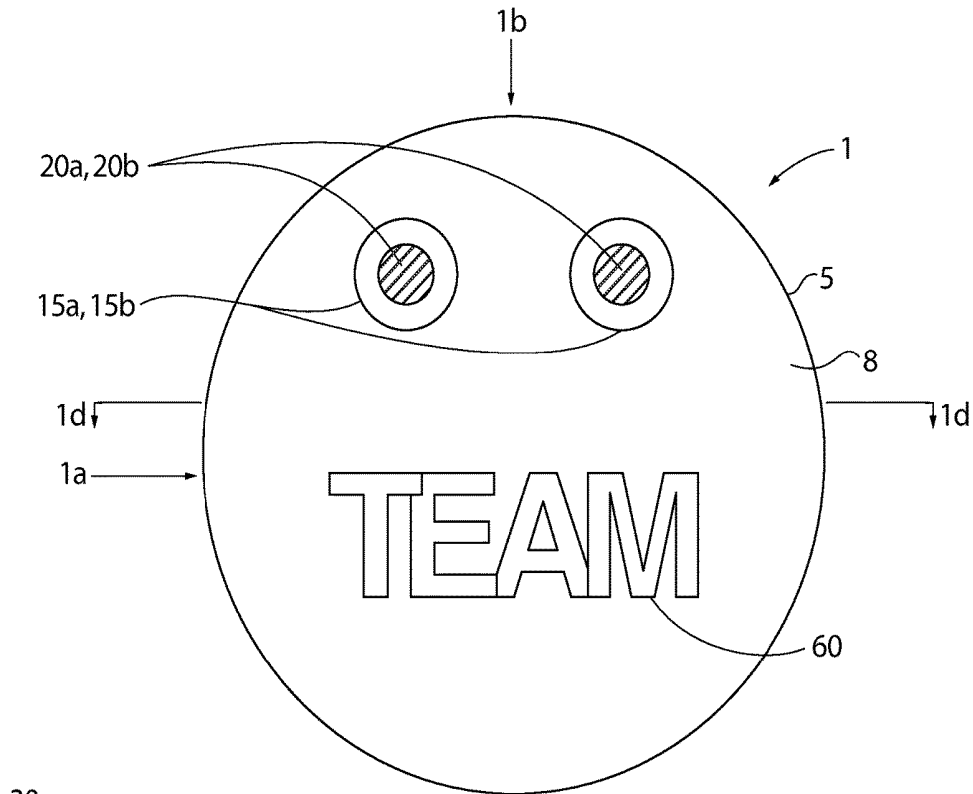
**Publication Classification**

(51) **Int. Cl.**  
*G01N 33/02* (2006.01)  
*A61B 5/00* (2006.01)  
(52) **U.S. Cl.**  
CPC ..... *G01N 33/02* (2013.01); *A61B 5/01* (2013.01); *A61B 5/0022* (2013.01)

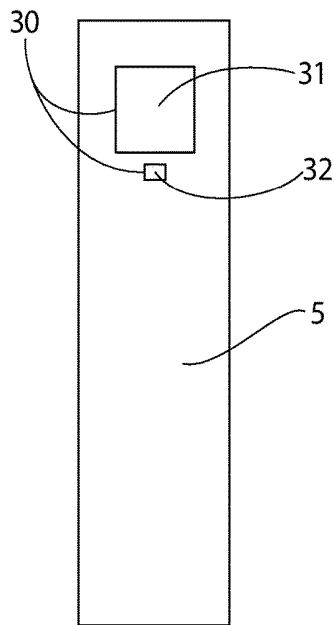
(57) **ABSTRACT**

Electronic systems, devices and methods are provided to accurately record and analyze food intake and physical activity in a subject. A device is provided to be placed on a subject which records video using at least two video cameras positioned to record stereoscopic image pair, as well as other physiological and/or environmental data including, for example, oxygen saturation, heart rate, and environmental factors such as physical location, temperature, and humidity. Video data is analyzed along with other data obtained by the device to determine food consumption and/or physical activity of the subject, much of which is accomplished by automated computer-implemented processes.

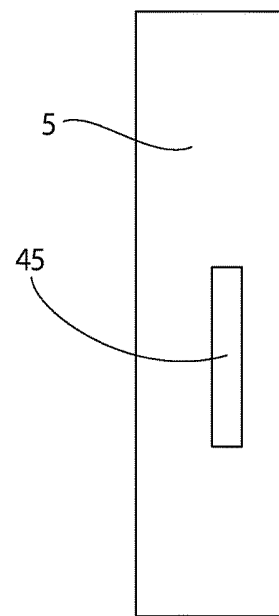




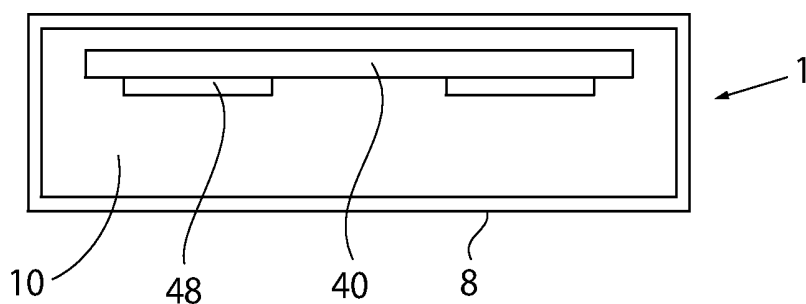
**FIG. 1a**



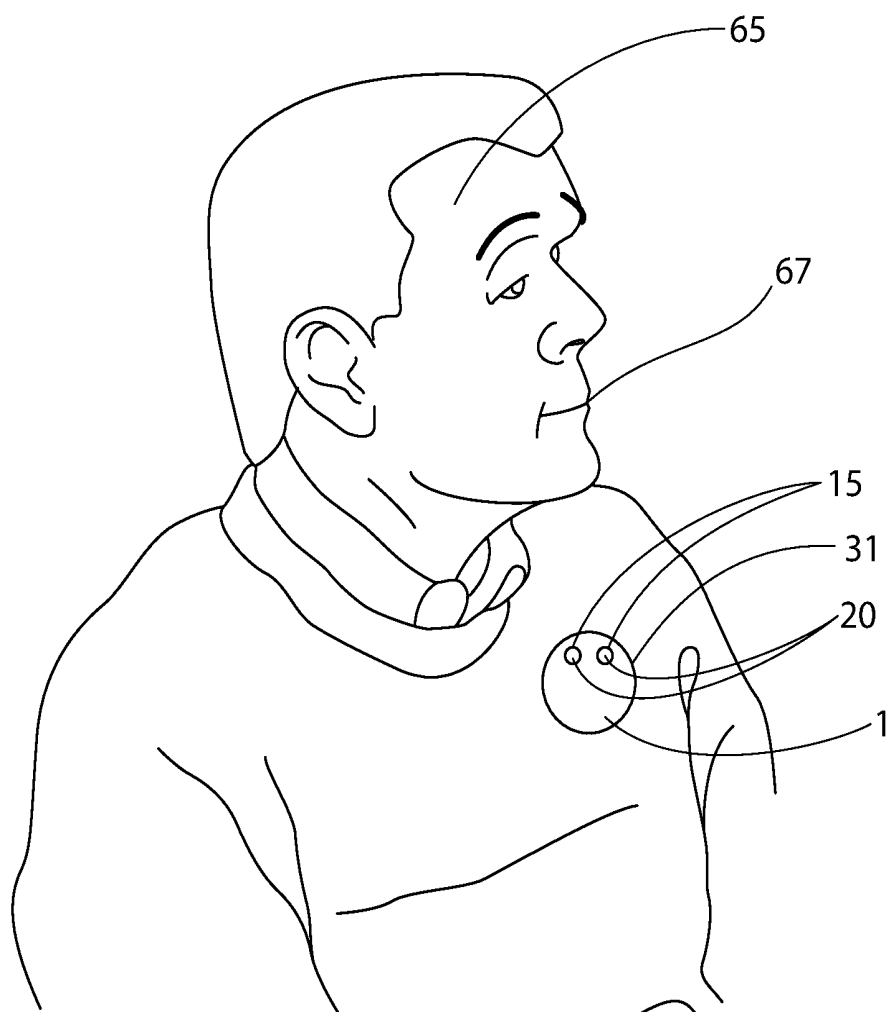
**FIG. 1b**



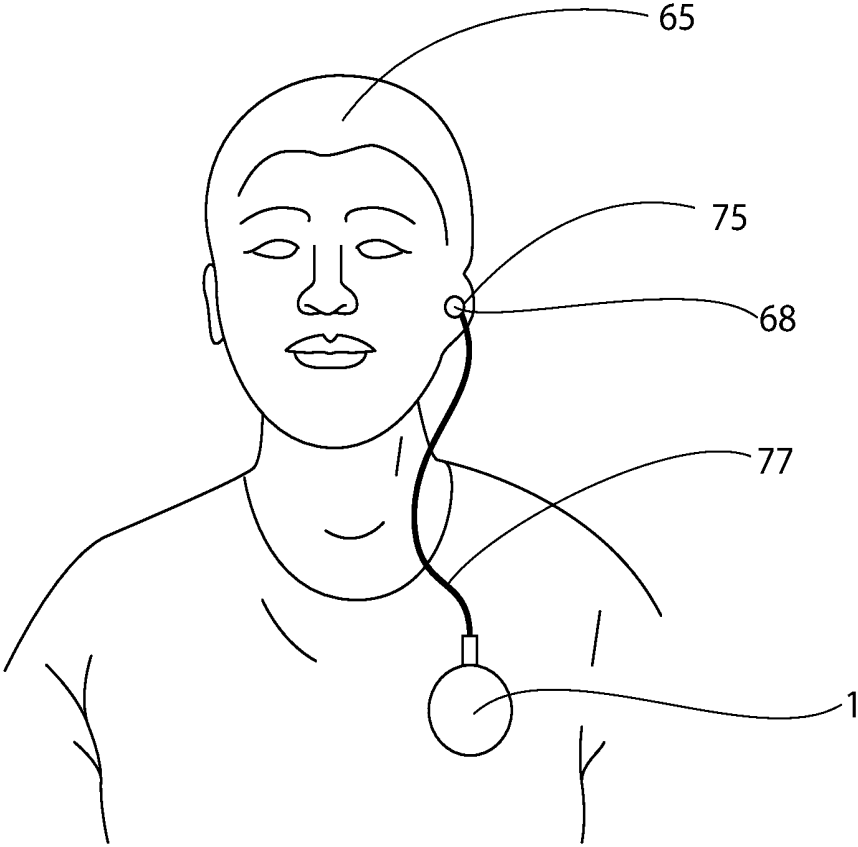
**FIG. 1c**



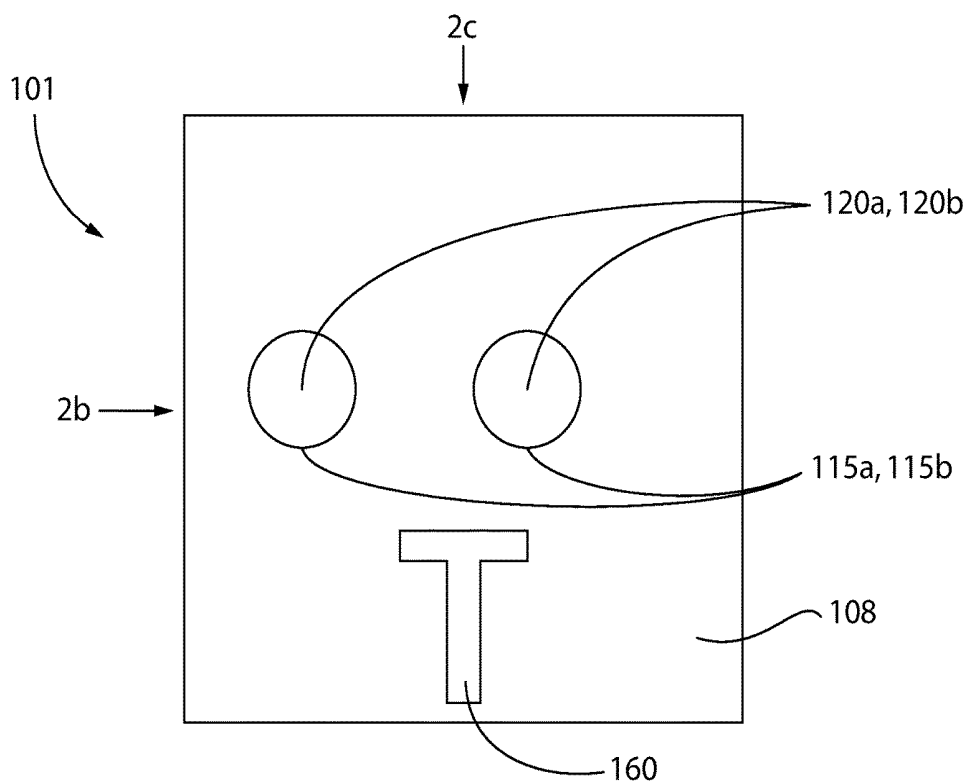
**FIG. 1d**



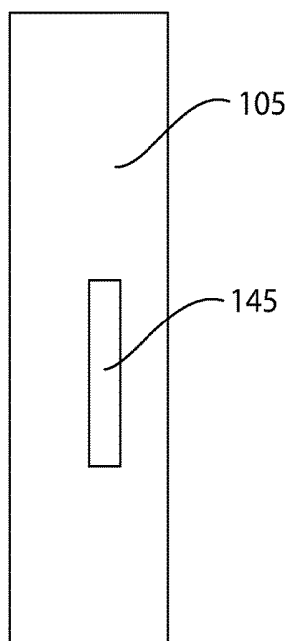
**FIG. 1e**



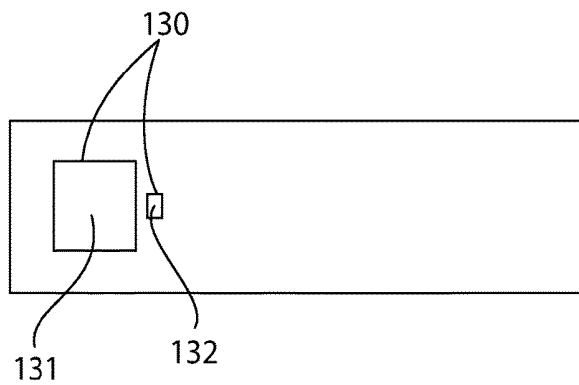
**FIG. 1f**



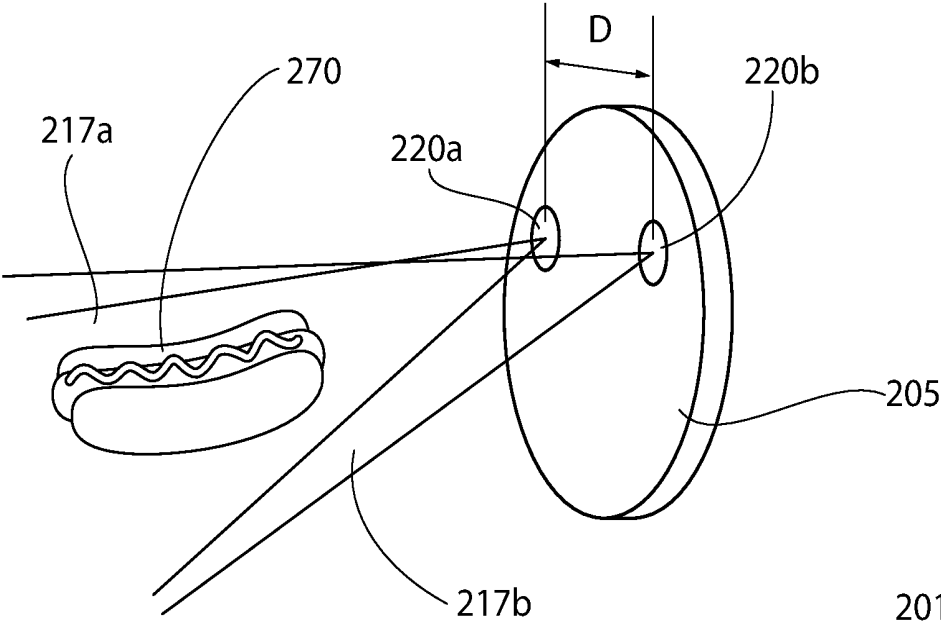
**FIG. 2a**



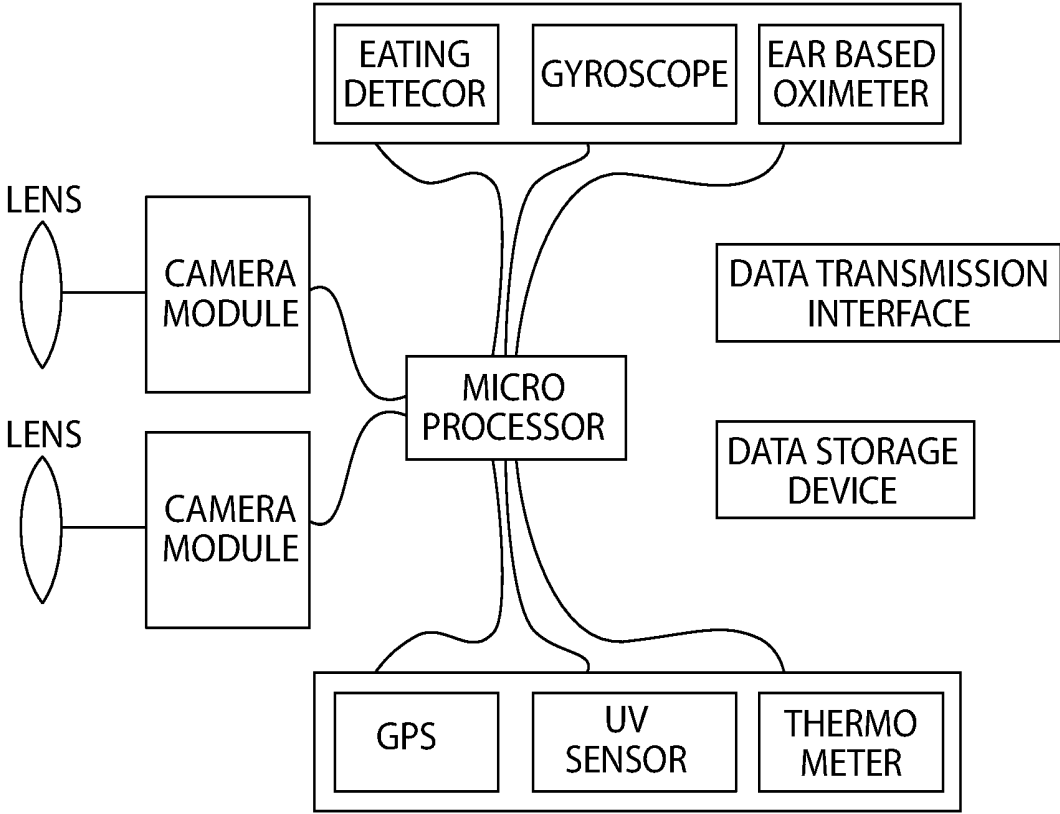
**FIG. 2b**



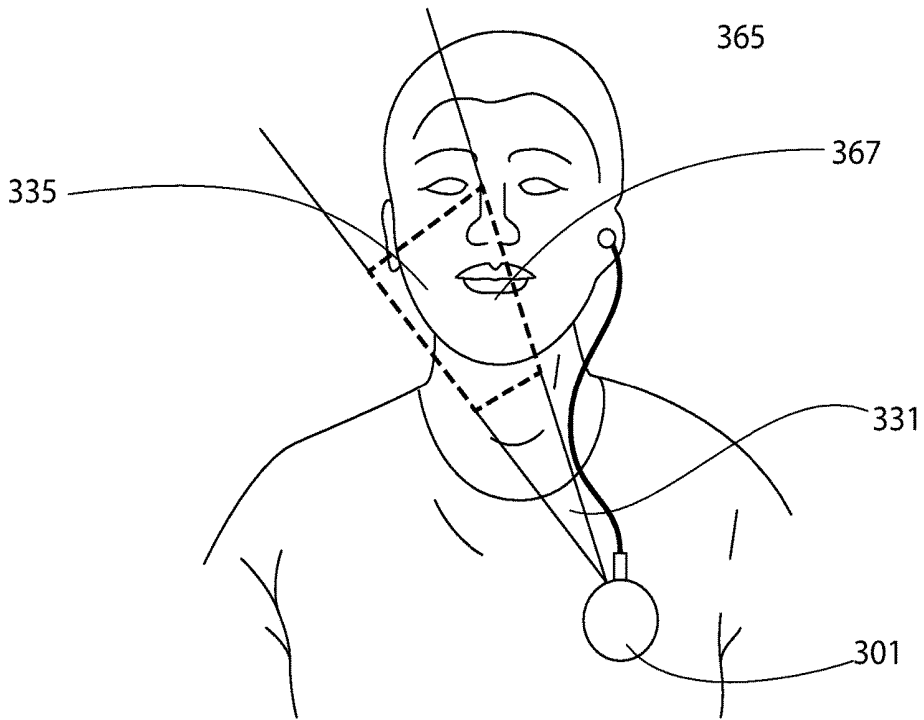
**FIG. 2c**



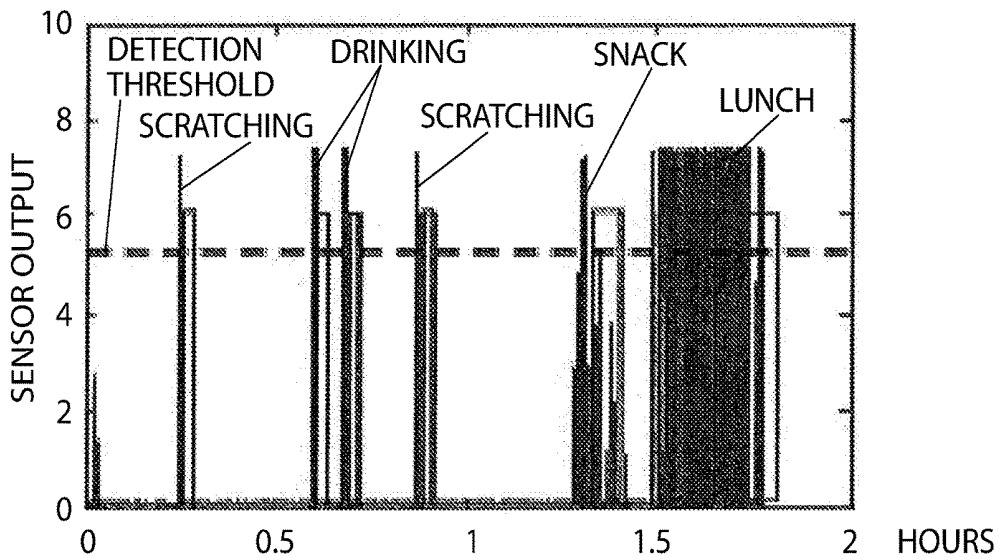
**FIG. 3**



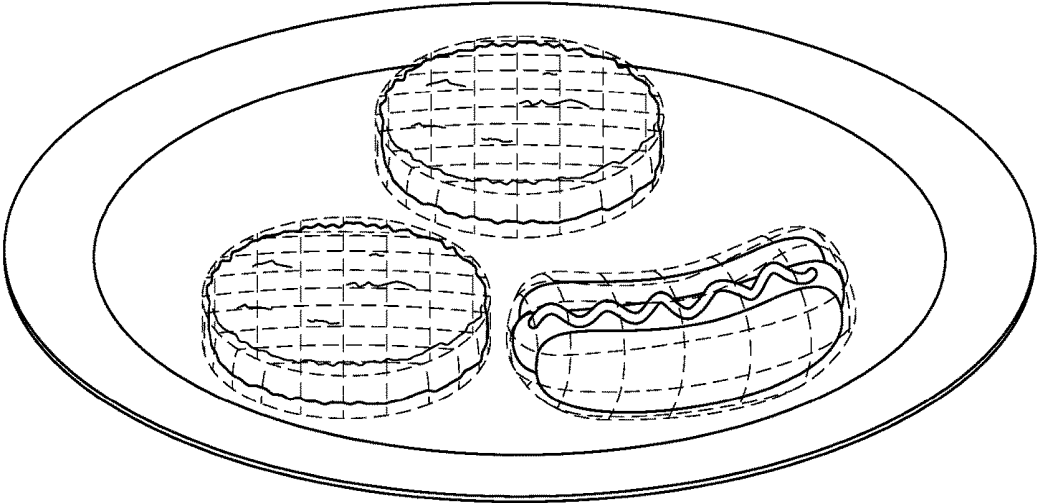
**FIG. 4**



**FIG. 5**



**FIG. 6**



**FIG. 7**

**METHOD, APPARATUS AND SYSTEM FOR  
FOOD INTAKE AND PHYSICAL ACTIVITY  
ASSESSMENT**

**CROSS REFERENCE TO RELATED  
APPLICATIONS**

[0001] This application is a continuation of U.S. patent application Ser. No. 13/676,468 filed Nov. 14, 2012, which claims the benefit of U.S. Provisional Patent Application No. 61/559,370, filed Nov. 14, 2011, each of which is herein incorporated by reference in its entirety.

**STATEMENT REGARDING FEDERALLY  
SPONSORED RESEARCH**

[0002] This invention was made with government support under Grant No. U01, HL091736 awarded by the National Institutes of Health. The government has certain rights in the invention.

[0003] Provided herein are devices and methods relating to personal monitoring, including without limitation, monitoring food intake and physical activity.

[0004] More than 60% of United States adults are overweight, and over one-third are obese. Obesity causes ~300,000 premature deaths each year in the United States. A recent study indicates that the estimated direct and indirect costs of obesity to the United States economy are at least \$215 billion annually. Obesity is caused by an accumulated surplus of energy, e.g., the calories obtained from foods and beverages exceed those expended on physical activity (physical activity) and metabolic functions. Therefore, measurements of energy intake and expenditure are key to the study of obesity and obesity-linked diseases, such as diabetes, cancer, and cardiovascular disease. Currently, the measurements of food intake and physical activity depend primarily on self-reporting methods. Although such methods of assessment are commonly used in studies, they have clear limitations: results rely on the subject's cognitive ability, memory and willingness to disclose personal behaviors, causing inaccuracy and bias in the measurements. While objective methods have been developed for diet and physical activity assessments, these methods have one or more of the following limitations: labor-intensive for human subjects, requiring an institutional setting (e.g., a metabolic kitchen), high cost, and/or questionable accuracy. The research community needs an advanced tool for accurate evaluation of both diet and physical activity in free-living individuals, which would overcome such limitations.

[0005] In recent years, some electronic methods have been developed for use in diet evaluation studies. Questionnaires and self-managed diet-evaluation tools can now be completed with an electronic interface or in an internet setting, where digital pictures assist participants in selecting food and determining portion size. A notable development has been the use by participants of a smart phone to take food pictures and transmit them to the laboratory for analysis in real time. Although these methods advance the field of diet evaluation, participants still must play a dominant role in providing diet information, which can disrupt normal eating habits. To avoid this problem, a recent method uses a smart phone worn below the neck (on a lanyard) that records a picture every six seconds and transmits it to the laboratory automatically. However, this method is also limited: the cell phone does not allow long-term operation (relatively short

battery life), requiring the subject to turn it on and off frequently (to preserve battery life), thereby disrupting normal behavior around eating episodes.

[0006] In sharp contrast to diet, where objective tools are rare, many objective tools are available for physical activity evaluation, such as pedometers, heart rate meters, and accelerometers. However, these tools provide little information about other important factors related to physical activity, such as the subject's behavior and physiological response, as well as built environment (this term refers to man-made surroundings for human activity) and natural environment (e.g., natural landscape, weather, air pollution indices). Some tools are also obtrusive (e.g., calorimetry), costly (e.g., doubly labeled water), suited only to laboratory settings (e.g., metabolic chamber), or inaccurate (e.g., some accelerometers).

[0007] In light of the deficiencies of currently available methods and devices for dietary and physical activity analysis, there is a need for a device which is: 1) objective regarding diet and physical activity data collection, 2) low-cost, 3) unobtrusive, 4) a low burden on participants, 5) cosmetically appealing, and 6) multi-functional. Furthermore, there is a need for a device that provides not only data for energy intake and expenditure calculations, but also information about built/natural environment, behavior, and physiological responses of a subject. Lastly, there is a need for a device that accurately and automatically identifies eating and physical activities and that can automatically catalog food items with minimal user input.

**SUMMARY**

[0008] Electronic systems, devices and methods are provided to accurately measure both food intake and physical activity. A novel multimedia method, an electronic device, and an integrated data processing system are described. Methods of use of the device also are described.

[0009] A device is provided for the remote monitoring of food intake and physical activity in a subject. According to one embodiment, the device comprises a housing; a power supply; and in the housing: one or more physiological sensors; two or more video cameras positioned in the housing to produce stereo images which, when viewed together, create a three dimensional perspective; a data storage device; and one or more processors (e.g. microprocessors) connected to the physiological sensors, one or more cameras for both images and videos, the data storage device and the power supply. The microprocessor(s) manage the collection, storage, and transmission of data from the physiological sensors and the video cameras, and other optional components configured within the device. In order to function, the device comprises a program that controls data acquisition from the video cameras and one or more physiological sensors and data storage on the data storage device, and optionally data analysis and conversion. The elements of the device are connected by and the effects of the device are implemented by suitable electrical or electronic elements, (e.g., terminals and connectors, wires, conductive traces, resistors, capacitors, transistors, integrated circuit chips, switches, diodes, protective devices, inductive devices, piezoelectric devices, crystals, resonators, etc).

[0010] In one non-limiting embodiment, the data storage medium is a flash-type memory card such as a MicroSD card or a USB thumb drive for recording data. The card or drive

is removable so that recorded data can be transferred to a computer or other external device.

**[0011]** In addition to a data recorder, optionally, the device further comprises a data transmission interface connected to the microprocessor(s) and/or data storage device which transfers data from the sensors, processors or data storage device to and/or from an external device. The transmission interface comprises either a wireless data interface, e.g., a transmitter, receiver and/or adaptor such as a Bluetooth, WiFi, cellular network or infrared interface or a hard-wire connection such as Universal Serial Bus (USB) connector. The power supply, for example a battery, is located either within or external of the housing.

**[0012]** Non-limiting examples of the one or more sensors of the device include an accelerometer (two-dimensional (2D) or three-dimensional (3D)), a gyroscope (2D or 3D), an indoor/outdoor detector, a temperature sensor, an audio system, an eating detector, an oxygen saturation sensor such as an ear-based oximeter, and an electrode system. According to one embodiment, the eating detector is an optical eating detector that projects a light beam in the region in front of the mouth of the subject and measures when an object, such as a food item, passes through the region. Examples of light beams include an infrared light beam, a laser beam, or a beam of light omitted by a light emitting diode (LED).

**[0013]** In one embodiment, the device further comprises one or more geographical and environmental sensors. Non-limiting examples of these sensors include a global positioning system (GPS) sensor, an ultra violet (UV) radiation sensor, a thermometer, and a humidity sensor.

**[0014]** The device optionally further comprises one or more indicia and/or aesthetic features. The indicia or aesthetic features are adaptable for the tastes of individual subjects. An exemplary aesthetic feature for such an embodiment of the device is a logo of a favorite team, university, or other entity with which the subject is affiliated. The aesthetic feature typically is on an outward-facing surface of the device.

**[0015]** In one non-limiting embodiment, the device is adapted to be worn as a button wherein the housing comprises a shell, such as a cylindrical shell, which contains the electronic components, including visual, geographical, motion, orientation, and/or environmental sensors, two video cameras in a stereoscopic arrangement, a recording device, and a power supply. Optionally, a lead (e.g., a wire) attached to the device connects the device to an ear-based oximeter for measuring oxygen saturation and/or electrodes to measure physiological variables, such as the heart rate and the respiration rate.

**[0016]** According to another embodiment, a system is provided for remote monitoring of food intake and physical activity in a subject. In one example, the system comprises: a device for monitoring food intake and physical activity of a subject and a computer. Optionally, the computer has a reader for receiving data from the device and inputting the data into the computer for processing. The reader is one or more of a card reader for receiving a data storage card from the device, a wireless data link comprising the following sensors in separate chips or a combined chip such as a Bluetooth network chip, a Wi-Fi chip, a service-activated cellular network chip, and a GPS receiver. According to one non-limiting embodiment, the wired information link consists of a hardwired connection such as a USB port. In one

embodiment, the system comprises one or more computer programs for determining a dimension of a food item in video data obtained from the video cameras and for determining a physical activity of a subject wearing the device. In one example of the system, the program for determining the dimension of a food item from video data obtained with the video camera is a food model fitting type analysis using a stereo image from the two video cameras. Similarly, the program for determining physical activity of the subject wearing the device is an automated process using an algorithm performed by a computer or a manual review process.

**[0017]** In another embodiment, a method is provided for remote monitoring of food intake and physical activity in a subject. In one example, the method comprises obtaining video data in the form of a stereo image from two or more video cameras suitably placed on a subject to show food consumption by the subject and physical activity of the subject; and determining from the video data one or both of food consumption by the subject and physical activity of the subject. The method optionally comprises limiting times when data regarding food intake is recorded to times when the subject is eating by using an optical eating detector to determine when eating occurs. The method also optionally comprises altering the frame rate and resolution of the cameras as well as turning one or both of the cameras off at times when the subject is not eating or is sedentary as detected by the accelerometer and/or gyroscope. The cameras can also be controlled selectively at different geographic locations (determined by the GPS sensor and/or indoor/outdoor sensor).

**[0018]** Also provided in the present application are methods of monitoring behavior, condition or activity of a subject. Such parameters may be monitored by recording output of a device comprising one or more sensors such as without limitation oximeters, heart rate monitors, accelerometers, and gyroscopes. The sensors may also include environmental sensors such as GPS sensors, magnetometers, electronic compasses, thermometers, VOC sensors, ambient sound recorders, humidity sensors, and UV sensors, two or more imaging sensors, and an eating detector. The device may held in a housing and may include a bus, a microprocessor connected to the bus, and a data storage device including data storage medium and a program for managing data acquisition and storage. The device also includes a power supply.

**[0019]** In non-limiting embodiments, the method may be used to collect information on behavioral, condition or activity data for specific populations. For example and without limitation, the methods may be used to collect information from police officers, astronauts, or consumers in a shopping environment (e.g. in a particular store or mall). The method may also be used to assess the behavior, condition or activity of animals, for example livestock on a farm or grazing on grazing lands.

#### BRIEF DESCRIPTION OF THE DRAWING

**[0020]** The figures are provided for illustrative purposes only and are not intended to limit the scope of the present inventions.

**[0021]** FIGS. 1a-1d depict schematically one non-limiting embodiment of the device for remote monitoring of food intake and physical activity in a subject. FIGS. 1a, 1b and 1c show top and two side views, respectively of the device. FIG. 1d depicts a cross section of the device.

[0022] FIG. 1e is a schematic drawing of the device shown in use on a subject.

[0023] FIG. 1f depicts a non-limiting embodiment of the device in use by a subject having an ear based oximeter attached to the ear lobe of the subject, comprising a lead running from the device to the photosensor.

[0024] FIGS. 2a-2c depict schematically a second embodiment of the device for monitoring of food intake and physical activity in a subject

[0025] FIG. 3 depicts schematically a non-limiting embodiment of the invention which highlights the placement of the two video cameras.

[0026] FIG. 4 is a block diagram showing the relationship of electronic components in one non-limiting embodiment of the device.

[0027] FIG. 5 depicts a schematic drawing showing the optical eating detector in use.

[0028] FIG. 6 is a graph displaying data collected from an optical eating detector worn by a subject over a two hour period.

[0029] FIG. 7 is a photograph of three food items arranged on a plate. A computer generated template is laid over the food item images to represent how dimensions and volume of the food items are estimated according to one process for estimating food dimensions.

#### DETAILED DESCRIPTION

[0030] The use of numerical values in the various ranges specified in this application, unless expressly indicated otherwise, are stated as approximations as though the minimum and maximum values within the stated ranges are both preceded by the word “about”. In this manner, slight variations above and below the stated ranges can be used to achieve substantially the same results as values within the ranges. Also, unless indicated otherwise, the disclosure of these ranges is intended as a continuous range including every value between the minimum and maximum values. For definitions provided herein, those definitions refer to word forms, cognates and grammatical variants of those words or phrases.

[0031] As used herein, the term “patient” refers to members of the animal kingdom, including but not limited to human beings, and implies no relationship between a doctor or veterinarian and a patient.

[0032] As used herein, the term “comprising” is open-ended and may be synonymous with “including”, “containing”, or “characterized by”. The term “consisting essentially of” limits the scope of a claim to the specified materials or steps and those that do not materially affect the basic and novel characteristic(s) of the claimed invention. The term “consisting of” excludes any element, step, or ingredient not specified in the claim. As used herein, embodiments “comprising” one or more stated elements or steps also include, but are not limited to embodiments “consisting essentially of” and “consisting of” these stated elements or steps. “A” or “an” refers to one or more.

[0033] A device is provided for monitoring food intake and physical activity in a subject, which includes monitoring the physical surroundings of the subject as well as physiological parameters and activities of the subject. In one embodiment, the subject is a patient, e.g., an obese patient or a patient suffering from diabetes, seeking to collect objective data regarding eating habits and physical activity for treatment. Patients suffering from other chronic diseases

whose effects are exacerbated by environmental factors may also use the device to collect objective data about diet, environmental conditions, and physical activity relevant for their treatment and condition. In another embodiment, the device assists blind persons in moving around by tracking location and alerting the user of objects in their path. In another embodiment, individuals, especially senior citizens, wear the device as a personal safety device to signal for help in the case of physical trauma (e.g., heart attack, stroke, traumatic physical injuries). In other embodiments, the subject is an athlete collecting data to assist in creating a physical training and nutrition regimen; and employees, so that the employer can collect data regarding physical demands of a job, active duty police officers and military, or astronauts during space missions. Similarly, the subject could be a shopper or consumer, allowing the store to collect data about shopper habits. In another embodiment of the device, a subject is a wild animal. For example, ecologists or biologists may use the device to collect data to better understand eating habits, responses to environmental stimuli, and physical activity of animals in the wild.

[0034] In a most general sense, the device comprises a housing. The housing is manufactured from any suitable material or combination of materials, including, for example: plastic ((co)polymer(s)), glass, ceramic, or metal. The housing comprises an interior cavity. The housing optionally includes one or more indicia and/or aesthetic features, such as a decal depicting a sports team’s logo, affixed to an outer surface of the housing, such as an outward-facing surface of the housing. Any element of the device that is contained wholly or partially within the housing, such as a battery within the housing or a lens protruding through a hole in the housing, are considered to be “in the housing”. The device comprises two or more imaging sensors, such as cameras or video cameras, arranged for stereoscopic imaging, a data storage device, a power supply, and, optionally, other types of sensors, such as: an eating detector (e.g., for determining when the subject is eating and for controlling the collection of data with regard to eating habits), an accelerometer (e.g., for measuring body acceleration), a gyroscope (e.g., for measuring body orientation), a GPS sensor (e.g., for determining the subject’s geographical location), an oxygen saturation sensor (e.g., for measuring blood oxygen level and/or heart rate), an electrode (e.g., for measuring for skin impedance) and an audio sensor/player (e.g., one or both of a microphone and a speaker or a plug/adaptor for a speaker or amplifier/speaker combination). Suitable sensors capable of being adapted for use with the device are available from a variety of commercial sources.

[0035] The components of the device, such as cameras/imaging sensors and sensors are combined in various manners with various analog and digital circuitry, including controllers, filters, ADCs (analog-digital chips), memory, communication devices and adaptors. As devices become smaller and processors become more powerful and use less energy, it is possible to integrate many more sensors, such as NEMS (microelectromechanical or nanoelectromechanical systems), onto single chips. MEMS accelerometers, gyroscopes, gas sensors, thermometers, humidity sensors, magnetometers, etc. are readily available from commercial sources and/or are abundantly described in the art. Technologies such as package on package (PoP) and system on a chip (SoC) integrated circuit packages allow manufacture

of very small devices with significant capacities. For example, smart phones use PoP technologies to stack memory and processors in a very small volume. One example of a SoC is a microcontroller (MCU), which is a small computer on a single integrated circuit typically containing a processor core, memory, and programmable input/output peripherals. Microcontrollers also may include timer module(s) and analog-to-digital converter(s) for, e.g., converting analog sensor output to a digital signal. High definition video cameras are available that measure less than 10 mm on a side. For example, a prototype device having eight PCB layers, an ARM 11 microprocessor, 32 GB (gigabytes) of flash memory, a video camera, a multi-axis accelerometer, a daylight sensor, a thermometer, a digital compass, an ambient sound recorder and a lithium ion battery was manufactured into a circular disc (pin) that is 61 mm in diameter, 10 mm thick, and weighs 40 g.

**[0036]** A “bus” is a subsystem that transfers electrical signals, e.g., data, between components inside a computer, or in the present instance within the device and often is a PCB, but can be any connector capable of transferring electrical signals between components of the devices described herein, including connections or connection points in PoP and SoC structures. Further, by a stated component being “connected to the/a bus” it is meant that component is electrically connected to the bus and in functional electrical communication (including data communication as is applicable) with other components of a device connected to the bus according to the design of the device.

**[0037]** An accelerometer measures acceleration. Most accelerometers can also measure tilt. The accelerometer was originally a large device, but with the continued advances in the field of micro-electromechanical systems (MEMS) technology, they now are available in sizes as small as 2.9×2.9×0.92 mm with 3-axis measurements. The architecture of the MEMS device initially used a suspended cantilever beam. Recently, a newer architecture has been developed using a heated gas bubble with thermal sensors. When the accelerometer is tilted or accelerated, the location of the gas bubble causes variations in the output. Although most applications of accelerometers are found in the automobile industry, such MEMS devices also are useful in biomedical applications, such as in physical activity, swallowing, and respiratory measurements. Although other types of accelerometers may find use in the devices described herein, a 3-axis accelerometer may be preferred in most instances for characterization of physical activity.

**[0038]** A gyroscope measures orientation. In one embodiment of the device, a gyroscope is used to determine changes in the orientation of the subjects’ body to help identify the physical activity performed. A mechanical gyroscope was originally a large device consisting of a rotor mounted within a frame structure of concentric circular members. The device measures changes in angular momentum. Gyroscopes based on MEMS technology are now widely commercially available. Available MEMS devices use one or more mechanisms for measuring orientation including tuning forks, vibrating wheels, or resonant solids of various designs. Major manufacturers of gyroscope chips are Panasonic, Robert Bosch GmbH, Seiko Epson, STMicroelectronics, and Analog Devices. Commercially available products have dimensions of about 4×4×1 mm. Commercial chips that combine a 3-axis accelerometer and a 3-axis gyroscope within one package are now available.

**[0039]** An oxygen saturation sensor indirectly measures oxygen saturation of arterial blood. For example, a pulse oximeter is a medical device that indirectly measures oxygen saturation by passing light in the red or infrared wavelengths through the skin of a subject and measuring the change in absorbance due to pulsing arterial blood. In one embodiment, the oximeter is an ear-based oximeter having a lead extending from the device housing to the ear of the subject. Commercially available ear-based oximeters adaptable for use with the present device include the BCI ear sensor clip for the BCI 3301 oximeter, the Nonin PureLight ear clip sensor manufactured by Nonin Medical, Inc., Plymouth, Minn., the ChipOX SPO2 Pulse Oximeter Module, and various ear sensor clips manufactured by ENVITEC-Wismar GmbH, Wismar, Germany. Ear-based oximeters may either attach to the ear lobe (so as not to impede hearing) or be inserted directly to the ear canal. Currently available oximeters measure both heart rate and oxygen saturation (SaO<sub>2</sub>). A newer enhanced ear-based oximeter is expected to become commercially available which also measures maximal oxygen consumption (VO<sub>2</sub>2). It is anticipated that the enhanced sensor may also be adapted to work with the device.

**[0040]** The imaging sensors, for instance a camera or video camera, may be any useful device for generating and subsequently recording useful image data. Although designs from different vendors are different, a video camera usually consists of a charge-coupled device (CCD) or complementary metal-oxide-semiconductor (CMOS) imaging sensor, a lens, a multifunctional video control/DSP chip, and a set of discrete components (e.g., capacitor, resistors, and connectors). The lenses either have a focus range useful for imaging as described herein or they comprise an auto-focus feature. Likewise, the lenses may be equipped with a zoom functionality. While the video control component on the chip performs a number of image acquisition tasks, the DSP component on the same chip implements data processing algorithms, such as noise reduction and simple forms of data compression and encryption. The digital output from the video control/DSP chip may be in either a parallel or a serial form, depending on the particular chip design and the input configuration in the next data processing or interface stage.

**[0041]** In one example, the imaging sensor is a video camera device manufactured by Sanyo Electric Co., Ltd., Tokyo, Japan (Part No. IGT99267J-ST). The size of the module is less than about a half an inch in its longest axis. Its size is adequately small for the devices described herein. The lens is an auto-focus camera lens that can be selected for its focal length. Choice of focal length represents a design element and is optional. A shorter focal length is helpful in terms of privacy protection, as only images within a certain proximity of the camera and the individual will be seen clearly. Beyond this focal distance, images would appear more blurred and perhaps indistinguishable, which would be desirable, for instance, with respect to preventing individuals from being identified with whom the subject volunteer comes into contact. The video camera also comprises a video control/DSP chip, a number of passive components, a printed circuit board (PCB), and a flexible wire strip including a connector for data and power transfer. This particular camera can record 354×288 color video (a resolution comparable to that of a VHS video tape) at 15 frames per second with a peak power consumption of 45 mW. Besides this device, newer commercial chips (e.g., part number OV6130-

C00A, available from OmniVision, Sunnyvale, Calif.) with comparable resolution exist, but are smaller in size and demonstrate lower power consumption (<20 mA when active). Higher-end chips supporting much higher resolution (e.g., DVD resolution or higher) and frame rates ( $\geq 30$  fps) also exist that are equipped with sophisticated imaging control and programming functions. However, these are achieved at the expense of higher power consumption. In configuring the systems described herein, factors such as the total power consumption of the device (which preferably is less than 50 mW (on average) for a one-day (12-18 hours) operation), video quality, complexity in circuit design, and the available space will have to be considered. At such low-level power consumption, the heat generated by the device should be negligible.

**[0042]** Aside from its usefulness in producing images that can be analyzed as described herein for food intake and activity assessment, the stereoscopic nature of the imaging system serves the purpose of rangefinding and sizing of imaged items such as food items. In an earlier system described in US-2009-0012433-A1, a laser rangefinder is used to automatically identify distances so that the size of food items might be calculated. This is not only obtrusive, but requires an entirely separate electronic subsystem in the device, and separate software implementations. Also described in that publication is use of a reference item that is inserted in the image of the food—requiring the user to carry the reference item and display it every time food is used. This, too, is cumbersome, and adds an additional level of complexity to the end-use of the device. To the contrary, stereoscopic imaging requires no user input, reference cards or obtrusive laser beams.

**[0043]** Imaging sensors, such as video cameras or cameras, can be said to be positioned (e.g., in a housing) to record a stereo image pair or to record a stereo image. This means that the cameras, typically two cameras, are arranged and positioned in the housing so that images produced by the cameras are able to produce a three-dimensional, stereoscopic image. A stereoscopic image creates the illusion of depth (three dimensions) by presenting the two offset images separately, e.g., to the left and right eye of a viewer or to a computer viewing/analysis program. Based on the three dimensional image, as shown below, it is possible to accurately estimate dimensional data including the actual distance from the camera and dimensions of objects depicted in the image. Dimensional estimates from stereo images are obtained without using reference cards or reference objects to estimate size and distance—a substantial advance over prior methods. In the context of the devices, systems and methods described herein, so long as there is sufficient distance between the imaging sensors and their field of view overlaps in a useful distance range for creating a stereoscopic image that is useful for the function of the device (for example and without limitation, from 1 to 20 feet and any range therebetween), the cameras are considered to meet this criteria.

**[0044]** A stereoscopic image is considered to be and can be referred to as a “representation of a stereo image” in certain instances in recognition that programs can vary widely in how they utilize output data from the stereo-arranged cameras to achieve a particular operation goal, and may not generate a stereo image per se as part of the process, but some representation of that image pair that can be evaluated by a computer program to achieve a desired

programming goal—irrespective of whether a stereoscopic image that is viewable by an individual is actually generated. For example, stereo images will greatly improve the system’s ability in food portion size measurement. Currently, a reference object (e.g., a checkerboard) must be present within the food image. With the stereo imaging system, this limitation can be removed since the distance between the two cameras can then be used as a reference. Stereo images are also essential in many other applications, such as navigation for the visually impaired persons. In this case, the distances of obstacles can be computationally estimated from such images. In one non-limiting embodiment, the imaging sensor is a video camera is capable of recording images at a variable frame rate from at least 30 images per second to as low as 0.1 frame per second. A stereoscopic image pair is a pair of images that can be viewed together to form a stereoscopic image.

**[0045]** In the most general sense, the device further comprises a data storage device. In one non-limiting embodiment, the data storage device is a digital data recorder such as a disk drive which records data onto a data storage medium. In another embodiment, the data storage medium is flash memory. The data storage medium is any type of non-volatile memory, for example, magnetic data storage media such as a hard disk drive or magnetic tape, or flash-based memory. Flash memory is a non-volatile computer storage chip using NAND or NOR type memory as found in MicroSD cards, USB flash drives, or solid-state drives. File systems optimized for flash memory (solid state media) include Embedded Transactional File System (ETFS), exFat and FFS2 systems. The data storage medium can be random access memory (RAM) or read only memory (ROM). The memory may be removable from the device or permanently installed within the housing and transferable to an external device through the data transmission interface.

**[0046]** In one embodiment, the device further comprises one or more power supplies such as a battery. A battery comprises one or more electrochemical cells that convert stored chemical energy into electrical energy. One non-limiting example of a useful battery is a lithium-ion battery. A lithium-ion battery is a rechargeable battery often used in electronic devices. It is preferable that the capacity of the lithium-ion battery provides enough power for the device to operate continuously for 15-16 hours (entire daytime). In some cases where the device is not operated continuously, however, a battery of smaller capacity is more appropriate for reduced device size and weight. Other types of batteries, adaptable for use in the device include nickel cadmium (NiCd), and nickel metal hydride (NiMH) batteries. Preferably the battery is rechargeable and, in that case, the device further comprises a battery recharge port.

**[0047]** The device described herein is controlled by one or more single-core and/or multiple-core microprocessors, essentially turning the device into a wearable computer. A microprocessor is a chip containing one or more integrated circuits which receives data and processes the data according to instructions stored in the chip’s memory. A microprocessor typically, and along with other functions, manages the collection of data from the various sensors and the video cameras, directs the storing of data by the data storage system and typically allocates system resources between the electronic components to reduce power consumption and decrease the need for duplicative electronic systems. One type of microprocessor which can be adapted for use with

the present device is an ARM Cortex 9 processor running a simplified UNIX operating system. This type of processor is commercially available from numerous sources. In order to save power and allow for parallel data processing, MCUs, which consume much less power than the main processor, can be used when the computing demand is low (the main processor can be powered down), and when the computing demand is high (all processors are activated). This multi-processor strategy is similar to the case of the hybrid car in which two types of engines are used to meet different power demands.

**[0048]** As mentioned above, one sensor optionally included with the device is an eating sensor. The eating sensor is used to determine when the subject eats and to ensure that eating events are recorded for later analysis. Based on the data collected, information including portion size, nutritional value, and caloric intake is determined. Previous eating detectors relied on an accelerometer physically and intimately attached to the subject (e.g., attached to the jugular notch of the subject using an adhesive) to determine when the subject swallowed. Attaching a device to the subject is not optimal. According to one embodiment of the device described herein, the eating detector is an optical detector comprising a diode, which projects a beam of light, such as infrared light, or a laser beam, to a region in front of the mouth of the subject, and a position detector. The beam of light may be distributed in a fan shape, creating a trapezoidal detection zone in front of the mouth of the subject. The position detector comprises a photodetector, such as a sensitive phototransistor. The photodetector and associated circuitry measure changes in light or radiation intensity, thereby determining when an object passes through the mouth region. Light emitting diodes and photosensors, such as phototransistors, are commercially available from numerous sources including Sharp, Inc. In use, an object passes through the detection area when the patient eats, drinks, or smokes. The detector optionally uses a processing algorithm, typically in conjunction with the microprocessor, to distinguish false positives such as when a subject yawns or touches her face as opposed to actual eating events. According to one embodiment, when the sensor determines that the subject is eating, a signal is sent to the video cameras to signal the cameras to record at a high frame rate and high resolution to ensure that the eating event is adequately recorded. At other times, such as when the subject is engaged in sedentary activity as detected by the accelerometer (e.g., watching TV), only a single camera records at a reduced frame rate. In this way, the eating detector regulates power consumption, reduces the amount of redundant data, and conserves storage capacity.

**[0049]** In addition to the sensors described above, adapted to measure physiological factors, additional embodiments of the invention include one or more environmental sensors to record information about the subject's surroundings. Non-limiting examples of environmental monitoring sensors include: a global positioning or Global Positioning System (GPS) sensor (e.g., for determining and recording the subject's location), an ultraviolet (UV) light sensor (e.g., for measuring whether the subject is indoors or outdoors and for measuring exposure to sun radiation), a thermometer and a humidity sensor. In one embodiment of the invention, location data obtained from the GPS is used in combination with environmental data obtained from a public database, such as the air quality database maintained by the United States

Environmental Protection Agency (EPA). The EPA database includes environmental information collected from a dense network of monitoring stations located throughout the country. Measurements relevant to the analysis of physical activity and well being include measurements of: weather condition, atmospheric particulates (e.g., PM2.5), carcinogenic particles, allergens, toxins and ozone. Optionally, an algorithm is used to estimate environmental conditions at the subject's exact location by extrapolating, interpolating, differencing, etc. conditions at two or more nearby or nearest monitoring stations to, e.g., account for other environmental factors, such as wind strength and directions. In one embodiment, data is downloaded to the device using the data transfer interface for real-time comparison and monitoring. An example of such a database is the EPA's Air Quality System (AQS) Data Mart. Information on this database is available at <http://www.epa.gov/ttn/airs/aqsdatamart/index.htm>.

**[0050]** In one embodiment of the device or system described herein, the device further comprises a data transmission interface for sending and receiving data to and from external sources, in effect, creating a personal area network (PAN) comprising the device, the data transmitter and an external receiver attached to an external source. A PAN is a computer network used for communication (e.g., data transmission) among computer devices including telephones and personal digital assistants (PDAs) in close proximity to the user's body. PANs can be used for communication among the personal devices themselves (intrapersonal communication), or for connecting to a higher level network and the Internet (an uplink). Networks may be wired, using, e.g., USB, ethernet and FireWire protocols. A wireless personal area network (WPAN) is made possible with wireless network technologies such as Bluetooth, WiFi, Z-Wave and ZigBee. WiFi (e.g., IEEE 802.11a, b, g, n) networking protocols may be used, which advantageously have a greater transmission range than (e.g.,) Bluetooth, but consequently also have greater power consumption. Suitable external sources for receiving data transmitted from the device and optionally processing the data include a computer, tablet PC, or smart phone and/or an external hard drive or other device for backing up stored data.

**[0051]** The device optionally comprises one or more privacy features. In one embodiment, data within the device and/or transmitted to and from the device is protected by encryption, according to any useful method. Privacy of the subject wearing the device or others can be protected by hardware or software components that allow for turning off recording from one or more of the cameras or when present a microphone, at any time when privacy issues are raised by recordation of data by the device, for instance at times when the subject is at the work site (as determined by the GPS sensor), being sedentary (as determined by the motion sensors), or anywhere either indoor or outdoor (as determined by the ultraviolet light sensor), etc. In addition, according to one embodiment, the frame rate of the video recording is reduced at times which the cameras are unlikely to record relevant data so that personal identifying information is not recorded. Additional privacy measures include face-recognition and obliteration programs that automatically recognize human faces and blur them or otherwise prevent recognition.

**[0052]** In another embodiment, efforts are taken to ensure that the privacy of the subject is protected as the data

obtained by the device is analyzed. Multi-layered protection may be utilized to guard against any possible failures to preserve privacy of both the subject and any bystanders. Each security layer is independent and, consequently the device, may include all, some or none of these layers. For example, a first security layer may involve installing hardware on the device specially designed electronic components to limit the scope (depth of field) of the video camera. A second security layer may use advanced encryption of all multimedia data. In one embodiment of the privacy protocol, no human is allowed to access any information during data recording, transmission, and transportation. A third security layer may comprise extensive filtering implemented by software. For example, all images of humans appearing in the entire video record may be automatically identified and destroyed. Again, no human needs to be involved in this security layer. In a further example of a security layer, passwords are used to allow authorized persons (e.g., who are called "honest brokers" as defined by HIPAA US Federal privacy regulation), who are required to sign a legal document of non-disclosure, to access data stored on the device. This access is possible only after all the previous security layers are automatically implemented. In an additional layer of privacy protection, a convenient mechanism can be provided by the systems to switch-off or block the camera in sensitive locations or public places where photographing or video-taking is prohibited.

**[0053]** The device comprises computer one or more programs (e.g., computer programs) for implementation of the functionality of the device. As used herein and as described in further detail below, a "program" comprises processes, modules, routines, subroutines, algorithms, software, executable instructions, etc. for implementing an action in a computer, such as a microprocessor-based device as described above. This is achieved by processing executable instructions by a processor.

**[0054]** In another non-limiting embodiment, a system is provided for remote monitoring of food intake and physical activity in a subject. A system refers to a combination of elements including a device for the remote collection of food intake and physical activity data as described herein and a computer for running programs, processes, algorithms, etc. such as programs for classifying, analyzing, modifying, converting and/or organizing the data. A computer can be any device for processing data including a smart phone, tablet PC, calculator, or standard laptop or desktop computer. A computer program is software that implements a certain specified function in the device, and can include algorithms or other structural features, and which can work with hardware in implementation of the certain task. A process is an active implementation of a program or a portion of a program. For instance, a "program for facial recognition and obliteration" is a program (e.g., software) that resides in the device and which has the capacity to recognize a face in a video image data and render the face unrecognizable in one or more images stored on the device that include the face. Programs for implementation of specified tasks are typically readily developed by one of ordinary skill in the art and can take a variety of physical/logical forms. Programs are stored on any useful data storage device (e.g., ROM or other non-volatile memory) as are broadly known in the computer arts. A database is an organized collection of data typically stored in digital form. A system may comprise a database, with includes either a local

database stored on data storage medium in the computer and/or the device or an active or activatable link providing access to data in a remote, e.g., public, database such that information required for activity of the program is obtained by remote connection to the remote database.

**[0055]** The system comprises a communications interface which receives data from the device and inputs the data into the computer. In another embodiment, the communications interface is a card reader such as a flash memory or MicroSD disk drive which receives a removable storage disk from the device and transfers the recorded data to the computer. In other instances, the communications interface is a wireless receiver such as a Bluetooth network card.

**[0056]** Dimension of food item: In one embodiment, the system includes a program for determining dimensions of a food item. The computer program includes function(s), e.g., algorithms, that identify food dimensions according to a form fitting model that estimates food volume from images recorded by the video cameras. As an example, a pair of stereo images, such as the images recorded by the two or more video cameras of the device, can only reconstruct a portion of the 3D food surface. Fortunately, most human foods have roughly known geometries (e.g., a piece of pizza is roughly a triangular prism). Therefore, food item volume can be estimated even when only part of the 3D surface is observable. An exemplary form fitting algorithm operates according to a virtual reality (VR) method wherein a computer-generated template is scaled, rotated, and deformed to virtually fit the food, as shown in the recorded stereo image, in the virtual image space in FIG. 7. The template is fit to the image of the food item using a least-square fit algorithm. When the camera is well calibrated, a wireframe virtual template provides an accurate measure of the portion size.

**[0057]** Determining physical activity: Where physical activity of a subject is determined, data is recorded as the subject engages in physical activity; the data is later reviewed either manually or by an automated process using a computer to identify and classify the physical activity performed by the subject. In an example of a manual method for determining physical activity, screenshots are viewed directly by a data assistant. The data assistant prepares an event summary package using software developed for that purpose. Using the software, the assistant reviews each screen shot utilizing an extremely fast "multi-touch" interface to review and label screenshots in quick succession. The assistant optionally also combines physical activity, diet, and behavior information into the event summary package.

**[0058]** In other embodiments, physical activity is determined by an automated process using computational algorithms. In one example of a method utilizing an automated process for identifying physical activity, a modeling computer program is used to simplify activity classification. According to one embodiment, numerous potential physical activities are decomposed so as to define a set of Eigen activities. Eigen activities are much easier to classify by the computer compared to the innumerable physical activities which could be performed by a subject. The Eigen value concept is parallel to a powerful mathematical concept where a large number of functions can be approximated using a much smaller set of basic functions. Briefly, the data is first broken down using an improved version of a segmentation algorithm. Then, the measured physical activity data within each "screenshot" is pre-processed individually

to extract “features”. Next, a classifier is used to classify the feature vectors into Eigen activities (EAs). This EA sequence is then analyzed by the second classifier that determines physical activity based on a hidden Markov model or a linguistic model. Under a hidden Markov model, a given series of observations (data from the video camera and other sensors) is used as an input, and according to the algorithm, the most likely corresponding sequence of states (the Eigen activity) is computed. In addition to using video data to identify physical activity, data from the other sensors may also be relied upon to further distinguish and classify various physical activities. For example, the accelerometer identifies trunk acceleration, the gyroscope indicates trunk orientation, the GPS identifies physical location (near buildings or in an outdoor environment), the oximeter evaluates heart rate and O<sub>2</sub> saturation, and the UV sensors to indicate whether the subject is indoors or outdoors. In certain non-limiting embodiments, recording of data by the system is triggered by thresholds that are detectable by one or more of the sensors. For example and without limitation, recording may be triggered by movement of the subject at a certain threshold speed (i.e. indicative of walking, cycling or running) for a set period of time (i.e. for at least 30 seconds, 15 seconds, or 5 seconds).

**[0059]** FIGS. 1a-1d depict schematically one non-limiting embodiment of the device for remote monitoring of food intake and physical activity in a subject. FIGS. 1a, 1b and 1c show top and two side views, respectively of the device 1. As shown in FIGS. 1a-1d, device 1 comprises a housing 5 (cylindrical in this non-limiting embodiment). The housing 5 comprises two holes 15a, 15b through which lenses 20a and 20b of video cameras are exposed. Holes 15a and 15b and lenses 20a and 20b are positioned to record a stereo image. In FIG. 1b, an eating detector 30 is depicted, comprising a light beam projector 31 and a light sensor 32. FIG. 1c depicts a flash memory card slot 45 for receiving a flash memory card, exposed through a hole in the housing 5. FIG. 1d is a cross section of the device 1, showing outer surface 8 and inner cavity 10 of the housing 5 in which electronic components are located. A printed circuit board (PCB) 40 is depicted. The PCB 40 is shown schematically and in reality may comprise any useful electronic structure(s) and configuration. Sensors, data storage, processors, data connectors, e.g., flash memory card slots and USB adaptors, device bus, such as an accelerometer and gyroscope are also located in the inner cavity 10. For example, a data storage device comprising a recorder and a data storage medium, such as a flash memory card, is located within the inner cavity. A microprocessor 48 is located in the inner cavity 10 and is connected to the PCB 40. A battery (not shown) is similarly housed within the interior cavity 10. In one embodiment, a battery is connected to the microprocessor so that the microprocessor manages power consumption of the battery. As shown in FIG. 1a, the device 1 also comprises an aesthetic feature 60, such as the logo of a sports team located on the outer surface 8 of the housing 5. FIG. 1e is a schematic drawing of the device 1 shown in use on a subject 65. The device comprises lenses 20 positioned in the housing 5 to record a stereo image. FIG. 1f depicts a non-limiting embodiment of the device 1 in use by a subject 65 having an ear based oximeter 75 attached to the ear lobe 68 of the subject 65. A lead 77 is shown connecting the device 1 to the oximeter 75. The device 1 is depicted as being worn as a button attached to the shirt of the subject 65.

**[0060]** FIGS. 2a-2c depict schematically a second embodiment of the device 101 for remote monitoring of food intake and physical activity in a subject. The device 101 comprises a housing 105 having a substantially rectangular shape and having an outer surface 108. The two video cameras are positioned to record a stereo image. The lenses 120a, 120b of the video cameras are exposed through holes 115a, 115b in the housing 105. As depicted in FIG. 2b, a card slot 145 is located on the side of the device for receiving a flash memory card. An optical eating detector 130 comprising a light beam projector 131 and a light sensor 132 is located on a side of the device 101 as shown in FIG. 2c. The device 101 also comprises an aesthetic feature 160 such as the logo of a sports team located on the outer surface 108 of the housing 105.

**[0061]** FIG. 3 depicts schematically a non-limiting embodiment of the device 201 which highlights the placement of lenses 220a and 220b of two video cameras in the housing 205 configured to record a stereo image. The lenses 220a and 220b are separated by a distance D. The field of view 217a and 217b for each video camera is depicted, showing overlap of the respective fields of view on food item 270.

**[0062]** FIG. 4 is a block diagram showing the relationship of electronic components in one non-limiting embodiment of the device described herein. Lenses from each of two video cameras receive an image and, if optionally equipped with an auto-focus mechanism, focus their lenses for recording by the camera. Similarly, data is collected by physiological sensors, including the accelerometer, gyroscope, eating detector, oxygen saturation sensor, optical sensors, heart rate monitor, and electrode. Signal received from any analog signal source, such as a sensor or camera is converted into digital form by suitably configured analog-to-digital (A/D) converter(s). A microprocessor along with suitable software programs/processes coordinates data acquisition from the camera and the physiological sensors, storage of the data, any required conversion or modification of the data, privacy functions such as data encryption, allocation of system resources for the device as well as any additional processing functions required by the device or system. With suitable software according to certain embodiments of the device, the microprocessor controls the usage of the physiological sensors and the video cameras by turning the devices on and off or, in the case of the video cameras, reducing the frame rate and resolution at times when the sensors and cameras are less likely to record relevant information regarding food consumption and physical activity. The microprocessor is also shown in electronic connection with the data storage device which comprises the recorder and data storage medium. The microprocessor controls and optimizes data storage and ensures efficient use of storage space. The device also comprises a data transmission interface comprising an antenna for sending and receiving data from the device to, e.g., a computer using an acceptable transmission protocol such as Bluetooth. The data transmission interface is designed to be in wireless communication with an external data receiver. The data receiver could be attached to an external source for backing up data such as a hard drive or to a computer for processing data, etc.

**[0063]** FIG. 5 illustrates an optical eating detector device 301 e.g., the device 1 depicted in FIG. 1a, in use on a subject 365. A beam of light 331 (either infrared or laser) is projected from the device 301. The beam 331 passes in front

of the mouth 367 of the subject 365. A light detection zone 335 is defined as the region in front of the mouth 367 of the subject, which is depicted as being trapezoidal in shape in a two-dimensional view. The device 301 includes a light sensor for measuring when an object passes within the detection zone.

[0064] FIG. 6 is a graph showing data obtained from the optical eating detector. The graph presents data recorded while the device is worn by a subject for a two hour period. Each spike indicates activity in the detecting zone. The amplitude of the spike indicates the intensity of the reading by the sensor and corresponds to the volume of the object which passes through the detecting region or, more specifically, to the percentage of the light beam which is obscured. The video camera is set to begin recording when one of the spikes reaches threshold intensity as denoted according to the dashed line. The video camera is set to continue recording for a pre-determined period of time after the sensor is triggered. Any additional triggering activity will re-start timing. The rectangles imposed over the spikes indicate times when the eating detector signals that the video cameras should be recording. The eating detector may also record other events when an object passes in front of the mouth of a subject such as smoking.

[0065] FIG. 7 depicts a computer generated template laid over an image of a food item to demonstrate how the dimensions and volume of the food item are estimated using an automated process using recorded stereo images. While in general, a pair of stereo images can only reconstruct a portion of the 3D food surface, since most food items have roughly known geometries (e.g. a piece of pizza is roughly a triangular prism) it is possible to estimate volume of the entire food item even though the 3D surface of only a small portion of the food item is actually known. Researchers recently developed a virtual reality (VR) method to model food shape from a partial observation. According to the recently developed automated process, a computer-generated template is scaled, rotated, and deformed to virtually fit the food in the image space. In FIG. 7, three food items are arranged on a plate. A computer generated template is oriented to overlay each food item. When the camera is well calibrated, the wireframe virtual template provided an accurate measure of the portion size.

[0066] Also provided in the present application is a method of monitoring behavior, condition or activity of a subject. Such parameters may be monitored by recording output of a device comprising a housing; a bus inside the housing; one or more sensors connected to the bus; two or more imaging sensors connected to the bus and positioned in the housing to record a stereo image; an eating detector in the housing and connected to the bus, eating detector comprising a light beam source configured in the device to pass a light beam in front of a subject's mouth when the device is worn by a subject and a light sensor able to measure changes of light intensity in a path of a light beam from the light beam source; a microprocessor connected to the bus; a data storage device connected to the bus and comprising a data storage medium; a program for managing one or more operations of the device including imaging, data acquisition and storage; and a power supply connected to the bus.

[0067] For monitoring behavior, in non-limiting embodiments the sensors included in the housing may be any suitable sensors that can collect data sufficient to determine the behavioral patterns of a human or animal. For example,

and without limitation, any sensor described herein, such as oximeters, heart rate monitors, accelerometers, and gyroscopes. The sensors may also include environmental sensors like GPS sensors, VOC sensors, magnetometers, electronic or digital compasses, thermometers, humidity sensors, ambient sound recorders, and UV sensors.

[0068] In non-limiting embodiments, the method may be used to collect information on behavioral, condition or activity data for specific populations. For example and without limitation, the methods may be used to collect information from police officers, astronauts, or consumers in a shopping environment (e.g. in a particular store or mall). The method may also be used to assess the behavior, condition or activity of animals, for example livestock on a farm or grazing on grazing lands. The livestock may include any domesticated animal including, but not limited, to cattle.

#### EXAMPLES

[0069] The examples described below are examples of validation studies, proposed validation studies or proposed real world uses for the device, system, and method for food intake and physical activity assessment. The examples presented are purely illustrative of potential functionality and uses for various non-limiting embodiments of the device.

##### Example 1: Determination of Food Dimension Using a Form Fitting Model and Nutritional Content Estimation

[0070] A preliminary experiment was conducted to validate the stereo vision approach. Two web cameras were placed side-by-side on a support. The intrinsic parameters of the cameras were obtained from the manufacturer and a once-for-all calibration was performed. Stereo pictures of three foods were taken simultaneously. Correspondences of pixels between each stereo image pair were built using a density matching algorithm. Then, real-world 3D coordinates were computed using the aforementioned numerical optimization process. Utilizing the virtual shape models (sphere, cuboid, and cylinder for orange, cheesecake and hamburger, respectively), food volumes were calculated. The average volumetric error for the three cases was 7.3%, a reasonable error for food objects. Preliminary results suggested that the stereo vision approach is feasible in measuring food portion size.

[0071] Based on volume estimations described above, calorie and nutrition information could be determined if the food item could be identified. One database that is used for this purpose is the Food and Nutrient Database for Dietary Studies (FNDDS) database. Using the FNDDS, the best match for the recorded food item was determined and calorie and nutrient values were retrieved. Although FNDDS is a vast database (over 13,500 foods), some foods are not listed or absent from volume information. In these cases, under the proposed method, a dietitian studies the composition of the food from various information sources (e.g., recipe and the cultural origin) and provides caloric and nutrient values

##### Example 2: Physical Activity Assessment: Manual Identification Method

[0072] To assess the accuracy of physical activity (PA) identification, an evaluation study with human subjects was conducted. Ten healthy adults wore a prototype device during the daytime. A total of 26 days of data (total about

240 hours) was recorded. The average rate of picture taking was approximately one picture for every two seconds (frame rate 0.5 per sec). In order for the subjects to maintain their normal lifestyle without modifying behavior, they were not told about our physical activity study. Sadly, little intentional physical exercise was performed by these subjects in their normal, but busy daily routines. During data analysis, an automatic algorithm to break the raw image sequences into individual “screenshots” was utilized. Then, the subjects were asked to watch these screenshots and identify their own PAs. The results were then saved as gold standards. Next, all of the labeled screenshots from the different human subjects were randomized and incorporated into a single pool of data. Finally, images from the randomized data pool were shown to a new group of participants to assess the accuracy of identifying the physical activity being performed. Each time, a screenshot was randomly drawn from the pool for the observer to identify the event. His/her entry was then compared against the gold standard. Despite the complexity of the daily-life activities, there were on average of 85.7% cases where the observers’ results agreed with the gold standard. Although these results have not been compared with conventional methods (e.g., using accelerometers), it would be difficult for these methods to achieve a comparable accuracy with respect to these complex real-life activities.

#### Example 3: Physical Activity Assessment Automated Method

**[0073]** An automated process has been proposed whereby a computer analyzes video data according to a categorization algorithm to identify physical activities performed by the subject. Under the proposed approach, the set of potential physical activities is decomposed into a set of Eigen activities (EAs), which can be more easily classified by a computer. The EA concept is parallel to a powerful mathematical concept where a large number of functions can be approximated using a much smaller set of basic functions. Briefly, the data is first filtered and broken down using an improved segmentation algorithm. Then, the measured PA data within each “screenshot” is pre-processed individually to extract “features”. Next, a classifier will be used to classify the feature vectors into Eigen activities (EAs). This EA sequence will be analyzed by the second classifier that determines physical activities based on a hidden Markov model or a linguistic model. Methods and preliminary data are described in more detail below.

**[0074]** Recently [L. Li, H. Zhang, W. Jia Z.-H. Mao, and M. Sun, “Indirect Activity Recognition Using a Target-Mounted Camera,” Proc. IEEE 4th Int. Conf. Image and Signal Processing (CISP’11), Shanghai, China, 2011.], we reported an automatic method for identifying six basic EAs (walking, bowing, crouching, waist turning, sitting-still, sitting-up) by processing video data was reported. Since the wearer of the device never appears in the video, EAs were recognized indirectly by examining the motion vectors in the image sequence. Recognition involved three steps: (1) identifying pixel correspondences that produced motion vectors between adjacent images; (2) building pattern features from the histograms of the motion vectors in each pair of images; (3) recognizing the PA using a classifier. Classifiers tested were the K-Nearest Neighbor (KNN), Naive Bayesian and Support Vector Machine (SVM) classifiers. Results showed that the classification accuracy depended on the quantization levels of motion vector directions. The SVM was the most

accurate classifier in our test (95%), while the KNN was the least accurate one (85%). In another work, five different EAs were recognized to recognize PAs expressed as the combinations of these activities.

**[0075]** Future study seeks to identify PA when the video recording frame rate varies considerably. Therefore, the motion vector may become ambiguous when the frame rate decreases. However, certain of the devices described herein for monitoring food intake and physical activity are constructed to provide multiple forms of sensor data including accelerometer, gyroscope, GPS, oximeter, and UV sensor data. Data is collected from these multiple sensors and used, in combination with the video data, to identify the physical activity performed. Data from multiple sensors is processed either utilizing normalized features directly to determine EAs or to implement a dimension reduction algorithm, such as the one based on the principal component analysis (PCA). The performances of the two cases are compared experimentally. During the research, sets of EAs are chosen according to the following criteria: 1) they are suitable to the data produced by our sensors; 2) they are as independent as possible—in parallel to the concept of using a basis in mathematics—thus their combinations can represent a larger set of physical and sedentary activities; and 3) the number of EAs is reasonably small to facilitate data processing. The relationships between feature patterns and EAs will be established using a well-established forward algorithm.

**[0076]** Once the set of EAs is chosen, it is applied for the recognition of physical activity in video data. In a preliminary experiment, the class of “indoor activities” was represented as a temporal sequence of events (represented by feature patterns) in an ordered form  $H_1, H_2, \dots, H_T$ . Using the temporal sequence, a hidden Markov model (HMM) with parameters  $\Theta_m$  was obtained from the aforementioned forward algorithm. Given  $k$  classes of activities, the observed data was classified by the following optimization process: where  $P(H_1, H_2, \dots, H_T | \Theta_m)$  is the HMM density function and  $c$  is the determined class. Experimentally, five types of common indoor activities were reviewed from field-acquired real-world videos acquired by the wearable camera.

**[0077]** In addition to the uses described above (namely as a way to obtain objective data about diet and exercise as part of a treatment for diabetes, obesity, or during athletic training), alternate embodiments of the invention are useful for additional purposes.

#### Example 4: Police Protection

**[0078]** A device as described herein is expected to be useful to police departments for automatically documenting important events by individual officers. According to one embodiment, the device comprises two cameras with either auto-focusing capability or a focus range between, e.g., two and twenty feet. The device also optionally comprises accelerometer, gyroscope, and GPS sensors. Notably, the GPS function could be used to track the location of officers throughout the day. Face-recognition software can be included to focus on individuals surrounding the officer. The software optionally triggers high-resolution imaging when a face comes into the view of the cameras. The device also optionally comprises a “panic button” to indicate that the officer needs assistance. Data, such as GPS location information and images are optionally transmitted by a wireless transmitter in the device, such as WiFi or Bluetooth and/or

over cellular or other networks to a central location such as a police station so that the location of each officer is known in real time and, optionally, images and sound in the field when transmission of these data are necessary. The physiological sensors could also track what sorts of physical activities each officer performs on a daily basis, to better understand how often officers must run, sprint, or walk during the performance of their duties. In some cases, physical data, such as that an officer is running for an extended period could be a sign that the officer is chasing a perpetrator and is in need of back up. Computer monitoring of these variables may trigger automatic acquisition of field images and alert personnel in the central data collection facility. The central facility could automatically send additional officers.

#### Example 5: Astronaut Exploration

**[0079]** A device as described herein is expected to be useful to monitor daily activities of an astronaut during a space mission. Information about the eating habits, daily physical exertion, and environmental factors such as radiation exposure could be used to ensure that astronauts remain in good health throughout the mission. As stated in previous sections, individuals who must keep track of the information by themselves typically under report the calories consumed each day and over report physical activity. The device of the present invention would provide more accurate data regarding astronaut calorie consumption to better modify astronaut diets in light of the physical demands of the job. Because food typically is packaged in single-meal containers, barcodes or similar 2D indicators, such as QR codes (e.g. JAGTAGs), can be used to identify food items.

#### Example 6: Consumer Expectations and Habits Data

**[0080]** The device also monitors shopping habits of consumers such as consumers in a super-market. For example, a supermarket could enlist a number of shoppers to wear the device while shopping at the store. The GPS sensor tracks the shopper's location to better understand the path the shopper uses while navigating through the store. In addition, the physiological sensors could be used to determine how fast the shopper is moving in each area of the store. The store may use this data to determine which advertising efforts are effective (i.e. causing the shopper to slow down and more closely examine the products) and which efforts are less successfully (the shopper quickly walks through the area without examining the merchandise). Similarly, the video cameras could be used to determine which products the shopper looks at and for how long. The device may be used to study the shopping process involving interactions between two or more people. For example, it may monitor how a shopper's decision is influenced by his/her spouse or children. In addition, the video camera or another sensor on the device, such as a bar code scanner, could be used to keep track of which products the shopper picks up, examines, or places in his or her cart. This data regarding the consumers interaction with specific goods could be used to more quickly restock shelves or, potentially, to shorten the check-out process by keeping a real time record of which goods are currently in the shopper's cart. Physiological sensors such as the heart rate monitor would provide additional data concerning the consumer's response to various products and

advertising techniques. This data regarding the way shoppers move around the store and how shoppers respond to different advertising techniques could be used to improve the physical layout of stores and to present products to consumers in effective ways.

#### Example 7: Data Collection for Animals

**[0081]** A device as described herein is expected to be useful for measuring physiological health, eating habits, and location of animals in the wild. In one example, the device is attached to an animal for a period of months or years. The data provided, concerning physical health, eating activities, and location, could help to better understand the impact of changes to the environment such as deforestation or climate change on the physiological health of individual species. For this embodiment, the device could be equipped with a data storage recorder capable of recording data for several weeks at a time. Researchers could locate the animal and recover a removable disk from the device as necessary or download data via wireless communication. The device could also be equipped with a long range wireless transmitter to continuously transmit data to a research facility.

**[0082]** A device as described herein can be used to monitor farm animals. For example, by affixing a device to the head of a cow, the activities, such as eating, resting, drinking, moving, and interacting with other animals, can be monitored for a long period of time without disturbing the animal. When both the calorie/nutrient intake and calorie expenditure of the cow are known, cattle raising methods can be optimized accordingly to maximize the production of milk with high quality.

**[0083]** Whereas particular embodiments of this invention have been described above for purposes of illustration, it will be evident to those skilled in the art that numerous variations of the details of the present invention may be made without departing from the invention as defined in the appended claims.

#### 1-10. (canceled)

**11.** A device configured to be worn on a subject for the remote monitoring of food intake and physical activity in a subject, comprising:

- a housing;
- a bus inside the housing;
- a physical or physiological sensor connected to the bus;
- two or more imaging sensors connected to the bus and positioned in the housing to record a stereo image;
- an eating detector in the housing and connected to the bus, the eating detector comprising a light beam source configured in the device to pass a light beam in a path in front of the subject's mouth and a light sensor able to measure light intensity data representative of movement of an object in the path of the light beam;
- a microprocessor connected to the bus;
- a data storage device connected to the bus and comprising a data storage medium;
- programming instructions, stored on the data storage device, that when executed by the microprocessor cause the microprocessor to:
  - receive the light intensity data from the light sensor;
  - identify based on changes in the received light intensity data when the object is in the path of the light beam to determine when the subject is eating;

- generate a signal to control the imaging sensors to obtain images of a food item being eaten by the subject when it is determined that the subject is eating; and
- generate, from the images, an image or a stereo image of the food item; and
- a power supply connected to the bus.
- 12.** The device of claim **11**, in which data stored on the data storage device is encrypted.
- 13.** The device of claim **11**, in which the data storage device comprises a removable flash memory card.
- 14.** The device of claim **11**, in which the device further comprises a data transmission interface connected to the bus.
- 15.** The device of claim **14**, further comprising programming instructions, stored on the data storage device, that when executed by the microprocessor cause the microprocessor to receive data from an external computer and to compare the received data with data obtained from the sensors.
- 16.** The device of claim **14**, in which the data transmission interface operates under WiFi, cellular network or Bluetooth protocols.
- 17.** The device of claim **14**, in which the data transmission interface comprises a wired data communications interface.
- 18.** The device of claim **11**, wherein the physical or physiological sensor comprises one or more of an accelerometer, a gyroscope, an ultraviolet sensor, a magnetometer, an audio sensor, or an oximeter.
- 19.** The device of claim **11**, in which output from the light sensor of the eating detector controls recording from the cameras by either modifying a frame rate of the cameras or turning the video cameras off or on.
- 20.** The device of claim **11**, in which the light beam source of the eating detector is a laser.
- 21.** A method of monitoring behavior, condition or activity of a subject comprising, using the device of claim **11**:
- recording output of the device affixed to the subject in a data storage device in a computer comprising a processor connected to a data storage medium using executable instructions executed by the processor; and
  - processing, using executable instructions to be executed by a processor output of the device, to produce an output relating to a behavior or activity of the subject.
- 22.** The method of claim **21**, comprising:
- recording image data from the two or more imaging sensors suitably positioned on a subject to record a stereo image pair of food consumed by the subject and an image of the physical activity of the subject; and
  - determining from the image data one or both of food consumption by the subject and physical activity of the subject.
- 23.** The method of claim **21**, in which the device comprises one or two cameras, an accelerometer, a gyroscope, a

heart rate monitor, an oximeter, and a global positioning system (GPS) sensor and in which the subject is a police officer and the data recorded includes environmental data and physiological data.

**24.** The method of claim **21**, in which the device comprises a GPS sensor, a heart rate monitor, and an accelerometer, and in which the subject is a consumer in a shopping environment and the data recorded includes geographical and physiological data.

**25.** The method of claim **21**, in which the device comprises one or two cameras, a GPS sensor, a heart rate monitor, an accelerometer, and an oximeter and in which the subject is livestock and the data recorded includes geographical and physiological data.

**26.** The method of claim **21**, in which the device comprises one or two cameras, an accelerometer, a gyroscope, a GPS sensor, an oximeter, and a UV sensor.

**27.** The device of claim **20**, in which the fitting algorithm for determining the dimension of the food item causes the microprocessor to generate a representation of the stereo image of the food item in a three dimensional perspective, scales and orients the virtual template shape to fit the food item depicted in the stereo image representation and estimates a volume of the food item by determining the volume of the template shape.

**28.** The device of claim **11**, further comprising a global positioning sensor and the system comprises a database comprising environmental quality information for different locations and a computer program for estimating environmental exposure of the subject based on location data obtained from the global positioning sensor, weather data obtained from environmental monitoring sensors on the device, and the database comprising environmental quality information.

**29.** The device of claim **11**, wherein the device comprises one or more programs for:

determining a dimension of the food item based on: the stereo image of the food item, the distance between the two or more imaging sensors, and a fitting algorithm which fits a virtual template shape to the stereo image of the food item; and

determining a physical activity of the subject wearing the device based on data from at least one of the physiological sensors, or the imaging sensors.

**30.** The device of claim **11**, wherein the housing comprises a front surface, a rear surface, and a side surface extending between the front surface and the rear surface, and wherein a field of view of the one, two, or more imaging sensors passes through at least one opening on the front surface of the housing to capture images of items in front of the subject, and wherein the light beam source directs the light beam through an opening on the side of the housing, such that the light beam passes in front of the subject's mouth.

\* \* \* \* \*

专利名称(译)	用于食物摄入和身体活动评估的方法，装置和系统		
公开(公告)号	<a href="#">US20180348187A1</a>	公开(公告)日	2018-12-06
申请号	US15/994003	申请日	2018-05-31
[标]申请(专利权)人(译)	匹兹堡大学		
申请(专利权)人(译)	匹兹堡大学 - 英联邦系统的高等教育的		
当前申请(专利权)人(译)	匹兹堡大学 - 英联邦系统的高等教育的		
[标]发明人	FERNSTROM JOHN FERNSTROM MADELYN JIA WENYAN SCLABASSI ROBERT SUN MINGUI		
发明人	FERNSTROM, JOHN FERNSTROM, MADELYN JIA, WENYAN SCLABASSI, ROBERT SUN, MINGUI		
IPC分类号	G01N33/02 A61B5/00		
CPC分类号	G01N33/02 A61B5/0022 A61B5/01 A61B5/02055 A61B5/14551 A61B5/1118 A61B5/1112 A61B5/04 G01N21/00 G16H20/30 G16H20/60 G16H40/67 A61B5/6816 A61B5/6805 A61B5/6802 A61B5/0077 A61B5/0024 A61B7/00 Y02A90/26		
优先权	61/559370 2011-11-14 US		
外部链接	<a href="#">Espacenet</a> <a href="#">USPTO</a>		

摘要(译)

提供电子系统，设备和方法以准确地记录和分析受试者的食物摄入和身体活动。提供一种设备放置在主体上，该设备使用至少两个定位记录立体图像对的摄像机记录视频，以及其他生理和/或环境数据，包括例如氧饱和度，心率和环境因素。例如物理位置，温度和湿度。分析视频数据以及由设备获得的其他数据以确定受试者的食物消耗和/或身体活动，其中大部分是通过自动计算机实现的过程完成的。

