U.S. PROVISIONAL PATENT APPLICATION FOR

DEVICE AGNOSTIC MOTION-RELATED METRIC PREDICTION FOR WEARABLE ELECTRONIC DEVICES

Inventors: Agni KUMAR, residing in
Sunnyvale, California

Darren A. FORDE, residing in
San Mateo, California

Di WU, residing in
San Jose, California

En-Tzu YANG, residing in
Cupertino, California

Haraldur T. HALLGRIMSSON, residing in
Seattle, Washington

Hui CHEN, residing in
Los Altos, California

Jaya NARAIN, residing in
Santa Cruz, California

James P. OCHS, residing in
San Francisco, California

Karthik JAYARAMAN RAGHURAM, residing in
Potomac, Maryland

Assignee: Apple Inc.
One Apple Park Way
Cupertino, California 95014

Entity: Undiscounted

BAKERHOSTETLER
600 Anton Blvd., Ste. 900
Costa Mesa, CA 92626-7221
714.754.6600

DEVICE AGNOSTIC MOTION-RELATED METRIC PREDICTION FOR WEARABLE ELECTRONIC DEVICES

TECHNICAL FIELD

[0002] The present description generally relates to device agnostic motion-related metric prediction for wearable electronic devices.

BACKGROUND

[0003] Various physiological parameters of a user can be measured and analyzed to estimate motion-related measures indicative of the user's physiological state. Computer hardware has been utilized to make improvements across different industry applications including applications used to assess and monitor motion-related metrics of the user.

BRIEF DESCRIPTION OF THE DRAWINGS

[0004] Certain features of the subject technology are set forth in the appended claims. However, for purpose of explanation, several embodiments of the subject technology are set forth in the following figures.

[0005] FIG. 1 illustrates an example network environment in accordance with one or more implementations.

[0006] FIG. 2 illustrates an example computing architecture for device agnostic motion-related metric prediction in accordance with one or more implementations.

[0007] FIG. 3 illustrates an example framework of a foundation model and task-specific adapters in accordance with one or more implementations.

[0008] FIG. 4 illustrates an example of a training operational framework for device agnostic motion-related metric prediction in accordance with one or more implementations.

[0009] FIG. 5 illustrates an example of an inference operational framework for device agnostic motion-related metric prediction in accordance with one or more implementations.

[0010] FIG. 6 is a flow chart of an example process that may be performed for device agnostic motion-related metric prediction in accordance with one or more implementations.

[0011] FIG. 7 is a flow chart of an example process that may be performed for device agnostic motion-related metric prediction in accordance with one or more implementations.

[0012] FIG. 8 is a flow chart of an example process that may be performed for device agnostic motion-related metric prediction in accordance with one or more implementations.

[0013] FIG. 9 illustrates an electronic system with which one or more implementations of the subject technology may be implemented.

DETAILED DESCRIPTION

[0014] The detailed description set forth below is intended as a description of various configurations of the subject technology and is not intended to represent the only configurations in which the subject technology can be practiced. The appended drawings are incorporated herein and constitute a part of the detailed description. The detailed description includes specific details for the purpose of providing a thorough understanding of the subject technology. However, the subject technology is not limited to the specific details set forth herein and can be practiced using one or more other implementations. In one or more implementations, structures and components are shown in block diagram form in order to avoid obscuring the concepts of the subject technology.

[0014] Machine learning has seen a significant rise in popularity in recent years due to the availability of training data, and advances in more powerful and efficient computing hardware. Machine learning may utilize models that are executed to provide predictions in particular applications. In an era characterized by the integration of Artificial Intelligence of Things (AIoT) and Internet of Medical Things (IoMT) into daily life, the capability to non-invasively and precisely monitor cardiovascular activities across various situations enables individuals to make informed health-related decisions. Emerging commercial-off-the-shelf measurement devices intended for customers and ongoing research initiatives aim to empower users in self-monitoring their cardiovascular activities.

[0015] Embodiments of the subject technology provide for a sensor-based foundation model configured to generalize motion-related metric predictions, such as calorie consumption, step data, or other motion-related outputs across multiple hardware platforms without reliance on training data specific to each device type. The subject technology can enable transferability of the foundation model to new hardware platforms, such as earbud devices, using a limited amount of device-specific training data. This transferability capability can be applied through a transfer learning process in which the sensor-based foundation model, originally trained using data from phones and/or smartwatches, can be adapted for use with other devices from which the training data was not obtained, such as earbud devices.

[0016] In one or more implementations, inertial measurement unit (IMU) signals obtained on one or more different devices, including gyroscope and accelerometer data, can be pre-processed by removing gravitational components to produce normalized motion signals independent of gravitational bias. The sensor-based foundation model can be trained on these normalized motion signals and heart rate data collected from a large-scale dataset encompassing multiple device types and diverse placement locations on a subject. The sensor-based foundation model can be fine-tuned for different physical activity types through few-shot learning techniques with limited data. Furthermore, task-specific adapters can be applied to the sensor-based foundation model for individual motion-related outputs to extend capabilities of the sensor-based foundation model without retraining of the sensor-based foundation model. The sensor-based foundation model can be configured using generative modeling techniques to generate vector representations of the normalized motion signals independent of manually engineered features. The generated vector representations can be used to estimate various motion-related outputs, including but not limited to caloric expenditure, step count, distance traveled, gait metrics, metabolic equivalents, activity classification, and contextual awareness of the subject.

[0017] FIG. 1 illustrates an example network environment 100 in accordance with one or more implementations. Not all of the depicted components may be used in all implementations, however, and one or more implementations may include additional or different components than those shown in the figure. Variations in the arrangement and type of the components may be made without departing from the spirit or scope of the claims as set forth herein. Additional components, different components, or fewer components may be provided.

[0018] The network environment 100 includes an electronic device 110, an electronic device 112, an electronic device 114, an electronic device 116, an electronic device 118, a server 120, and a group of servers 130. The network 106 may communicatively (directly or indirectly) couple the electronic device 110 and/or the server 120. In one or more implementations, the network 106 may be an interconnected network of devices that may include, or may be communicatively coupled to, the Internet. For explanatory purposes, the network environment 100 is illustrated in FIG. 1 as including the electronic device 110, the electronic device 112, the electronic device 114, the electronic device 116, the electronic device 118, the server 120, and the group of servers 130; however, the network environment 100 may include any number of electronic devices and any number of servers or a data center including multiple servers.

[0019] In the example of FIG. 1, the electronic device 110 is depicted as a smartphone. However, it is appreciated that the electronic device 110 may be implemented as another type of device, such as a wearable device (e.g., a smart watch or other wearable device). The electronic device 110 may be a device of a user (e.g., the electronic device 110 may be associated with and/or logged into a user account for the user at a server). The electronic device 110 may be, for example, a desktop computer, a portable computing device such as a laptop computer, a peripheral device (e.g., a digital camera, headphones), a tablet device, a wearable device such as a watch, a band, and the like. Although a single electronic device 110 is shown in FIG. 1, it is appreciated that the network environment 100 may include more than one electronic device, including more than one electronic device of a user and/or one or more other electronic devices of one or more other users. The electronic device 110 may be, and/or may include all or part of, the electronic system discussed below with respect to FIG. 7.

[0020] By way of example, the electronic device 112 is depicted as a head mountable portable system that includes a display system capable of presenting a visualization of an extended reality environment to a user. The electronic device 112 may be, for example, a desktop computer, a portable computing device such as a laptop computer, a smartphone, a peripheral device (e.g., a digital camera, headphones), a tablet device, or a wearable device such as a watch. The electronic device 112 may be, and/or may include all or part of, the electronic system discussed below with respect to FIG. 7.

[0021] By way of example, the electronic device 114 is depicted as a watch. The electronic device 114 may be, for example, a desktop computer, a portable computing device such as a laptop computer, a smartphone, a peripheral device (e.g., a digital camera, headphones), or a tablet device. The electronic device 114 may be, and/or may include all or part of, the electronic system discussed below with respect to FIG. 7.

[0022] By way of example, the electronic device 116 is depicted as a desktop computer. The electronic device 116 may be, for example, a portable computing device such as a laptop computer, a smartphone, a peripheral device (e.g., a digital camera, headphones), a tablet device, a wearable device such as a watch, a band, and the like. The electronic device 116 may be, and/or may include all or part of, the electronic system discussed below with respect to FIG. 7.

[0023] By way of example, the electronic device 118 is depicted as an earbud. The electronic device 118 may be, for example, a desktop computer, a portable computing device such as a laptop computer, a smartphone, a peripheral device (e.g., a digital camera), a tablet device, a wearable device such as a watch, a band, and the like. The electronic device 118 may be, and/or may include all or part of, the electronic system discussed below with respect to FIG. 7.

[0024] In one or more implementations, one or more of the electronic devices 110-130 may provide a system for training a machine learning model using training data, where the trained machine learning model is subsequently deployed to one or more of the electronic devices 110-130. One or more of the electronic devices 110-130 can collect data (e.g., health-related information) that is then used to train a machine learning model, which will be described with reference to FIGs. 2-7. Further, one or more of the electronic devices 110-130 may provide one or more machine learning frameworks for training machine learning models and/or developing applications using such machine learning models. In an example, such machine learning frameworks can provide various machine learning algorithms and models for different problem domains in machine learning. In an example, the electronic device 110 may include a deployed machine learning model that provides an output of data corresponding to a prediction or some other type of machine learning output. In one or more implementations, training and inference operations that involve individually identifiable information of a user of one or more of the electronic devices 110-130 may be performed entirely on the electronic devices 110-130, to prevent exposure of individually identifiable data to devices and/or systems that are not authorized by the user.

[0025] The server 120 may form all or part of a network of computers or the group of servers 130, such as in a cloud computing or data center implementation. For example, the server 120 stores data and software, and includes specific hardware (e.g., processors, graphics processors and other specialized or custom processors) for rendering and generating content such as graphics, images, video, audio and multi-media files. In an implementation, the server 120 may function as a cloud storage server that stores any of the aforementioned content generated by the above-discussed devices and/or the server 120.

[0026] The server 120 and/or the group of servers 130 may provide a system for training a machine learning model using training data, where the trained machine learning model is subsequently deployed to the server 120, the group of servers 130 and/or to one or more of the electronic devices 110-118. In an implementation, the server 120 and/or the group of servers 130 may train a given machine learning model for deployment to a client electronic device (e.g., the electronic device 110, the electronic device 112, the electronic device 114, the electronic device 118). In one or more implementations, the server 120 and/or the group of servers 130 may train portions of the machine learning model that are trained using (e.g., anonymized) training data from a population of users, and one or more of the electronic devices 110-130 may train portions of the machine learning model that are trained using individual training data from the user of the electronic devices 110-118. The machine learning model deployed on the server 120, the group of servers 130 and/or one or more of the electronic devices 110-118 can then perform one or more machine learning algorithms. In an implementation, the server 120 and/or the group of servers 130 provides a cloud service that utilizes the trained machine learning model and/or continually learns over time.

[0027] FIG. 2 illustrates an example computing architecture for device agnostic motion-related metric prediction in accordance with one or more implementations. For explanatory purposes, the computing architecture is described as being provided by an electronic device 200, such as by a processor and/or memory of the server 120, or by a processor and/or a memory of any other electronic device, such as the electronic device 110. Not all of the depicted components may be used in all implementations, however, and one or more implementations may include additional or different components than those shown in the figure. Variations in the arrangement and type of the components may be made without departing from the spirit or scope of the claims as set forth herein. Additional components, different components, or fewer components may be provided.

[0028] As illustrated, the electronic device 200 includes training data 210 that is stored in memory of the electronic device 200 for training a machine learning model. In an example, the server 120 may utilize one or more machine learning algorithms that uses the stored training data 210 for training a machine learning (ML) model 220. The ML model 220 may include one or more neural networks, which will be described in more detail below with reference to FIGs. 3A, 3B, 5A and 5B.

[0029] Training data 210 may include IMU data, gyroscope data, accelerometer data, time-correlated photoplethysmography (PPG) data and/or heart rate data from diverse participants, encompassing subjects with demographic information (e.g., age, gender, BMI, etc.) for a user of the electronic device 110, and/or a population of other users. Participant profiles can vary, with some presenting normal heart rate readings and others with abnormal heart rate readings, for example. In one or more implementations, the training data 210 may include annotations for supervised learning during training of the ML model 220.

[0030] The training data 210 can be used to produce a machine learning model (e.g., ML model 220) that is trained to predict caloric expenditure, among others, of subjects. In one or more implementations, the training data 210 may include calorimetry measurements and time-correlated sensor signals from diverse participants, encompassing subjects with demographic information (e.g., varying ages, from infants to several years old). Participant profiles can vary, with some presenting low calorimetry measurements and others with high calorimetry measurements, for example. In one or more implementations, the training data 210 may include annotations for supervised learning during training of the ML model 220. In one or more other implementations, the training data 210 includes data captured via at least one of the electronic devices 110-118, such as the electronic device 114.

[0031] FIG. 3 illustrates an example framework of a foundation model 310 and task-specific adapters in accordance with one or more implementations. In one or more implementations, the foundation model 310 is employed in combination with a task-specific adapter to generate a task-specific estimation model through parameter-efficient fine-tuning and modular architectural integration. The foundation model 310 may include a large-scale, pre-trained neural network trained on an extensive dataset including broad general data. This foundation model 310 can encode generic feature representations that can be applied across multiple domains and tasks.

[0032] To tailor the foundation model 310 for a particular estimation task, a task-specific adapter such as the caloric expenditure adapter 322, the gait adapter 324, the step count adapter 326, or the metabolic equivalents adapter 328 can be employed. Each adapter represents a lightweight, trainable module that includes a limited number of neural network layers or parameter-efficient components like low-rank adaptation (LoRA), prefix tuning, or bottleneck layers. These adapters can be integrated either within or atop selected layers of the foundation model 310. During the fine-tuning phase, only the parameters of the adapter are updated, while the majority of the foundation model 310 remains frozen, preserving its generalized feature extraction capabilities while enabling efficient adaptation to specific tasks with reduced computational overhead and limited training data.

[0033] The combined architecture, formed by the foundation model 310 and the trainable adapter, operates as a task-specific estimation model (e.g., ML model 220). Input data, stored temporarily in memory (e.g., the permanent storage device 902 of FIG. 9) within the electronic device 110, the electronic device 114, the electronic device 116, or the electronic device 118, can be processed sequentially through the foundation model 310 to generate a latent representation. The task-specific adapter then refines or modifies this latent representation based on task-specific objectives, producing an output aligned with a targeted estimation function, such as classification, regression, or segmentation.

[0034] In one or more implementations, the adapters can refer to specific programs, which can be small AI programs that a processor (e.g., processing units 912 of FIG. 9) attempts to load and execute. For example, a caloric expenditure application 332 may deploy the ML model 220, which can include a foundation model 310 implemented as an LLM and a caloric expenditure adapter 322 loaded on top of, and/or in association with, the foundation model 310. In one or more implementations, the caloric expenditure adapter 322 can modify the output of the foundation model 310 by enhancing or adjusting motion-related metric estimations of the foundation model 310 to better represent the caloric expenditure of the user. Similarly, a gait application 334 may deploy the ML model 220, which can include the foundation model 310 and a gait adapter 324 loaded onto, and/or in conjunction with, the foundation model 310. In one or more implementations, the gait adapter 324 can modify the output of the foundation model 310 by enhancing or adjusting motion-related metric estimations of the foundation model 310 to better represent gait estimations of the user. Similarly, a step count application 336 may deploy the ML model 220, which can include the foundation model 310 and a step count adapter 326 loaded onto the foundation model 310. In one or more implementations, the step count adapter 326 can modify the output of the foundation model 310 by enhancing or adjusting motion-related metric estimations of the foundation model 310 to better represent step count activity of the user. In one or more other implementations, a metabolic equivalents adapter 328 can be loaded onto the foundation model 310. These adapters can be stored in the permanent storage device 902 (FIG. 9). When an application needs to be run, the requested adapters can be loaded and executed by a processor (e.g., processing units 912 of FIG. 9).

[0035] The foundation model 310 can remain constant across different adapters. In one or more implementations, the foundation model 310 includes a single instance to which multiple adapters can be attached. Due to the substantial size of the foundation model 310, the foundation model 310 may be loaded once between different adapters. When running various applications, each task-specific adapter may refer to the already-loaded foundation model 310. In one or more implementations, each task-specific adapter may possess specific dynamic mutable weight values, which are filled into mutable weights of the foundation model 310 upon attachment to the foundation model 310. In one or more other implementations, mutable weights can allow for dynamic patching of a new set of weight values into the foundation model 310 at runtime. This can enable the loading of the foundation model 310 with non-mutable weight values and dynamic addition of mutable weight values during, and/or immediately prior to, inference. Different instances with varied weights can be created, serving as different adapters. For example, by plugging in a set of mutable weight values tailored for a caloric expenditure task, the caloric expenditure adapter 322 can be generated.

[0036] In one or more implementations, a learning framework can be established for accommodating mutable weights and loads. This learning framework may involve segregating weights of the foundation model 310, while adjustments to intermediate weights can be made to accommodate various adapters. In one or more implementations, during training, the foundation model 310 includes specific linear layers within a neural network architecture. The weights of the foundation model 310 may be divided, with one portion designated for general LLM tasks, while another portion is reserved for learned weights specific to each adapter. In one or more implementations, multiple sets of weights may be stored concurrently in memory (e.g., permanent storage device 902 of FIG. 9), retaining previous sets of weights in memory. The mutable weight values can be determined during a model setup phase, in which each adapter is trained with its own set of mutable weight values tailored to its specific task. For example, each adapter may be associated with a separate training dataset distinct from the training dataset of the foundation model 310. The mutable weights can be trained to enhance the predictive capabilities of the adapter for a particular task. Subsequently, these mutable weight values can be packaged into separate data files and deployed for loading onto an electronic device (e.g., the electronic devices 110-118 of FIG. 1).

[0037] In one or more implementations, multiple instances of adapters can be generated, each adapter referring to a single static instance of weights for the foundation model 310. The task descriptors can be loaded once the adapters introduce their custom weights based on fine-tuning. In one or more other implementations, multiple fine-tuned adapters may be loaded concurrently with a single instance of the foundation model 310 with all adapters sharing that same instance of the foundation model 310. Various instances of mutable weight values can then be incorporated into these adapters. For example, there can be distinct sets of mutable weight values for tasks such as caloric expenditure and step count. Despite differences between the sets of mutable weights, the adapters can operate concurrently within a same model instance. For example, separate adapters can be executed simultaneously while relying on the same weights of the foundation model 310. In one or more implementations, multiple sets of these mutable weights can be loaded and used concurrently. As such, the foundation model 310 may allow for the concurrent use of multiple sets of weights without needing to discard existing weights.

[0038] Each adapter may include weight values that are fine-tuned for a specific task. For example, in the task of caloric expenditure, the training of the caloric expenditure adapter 322 may involve fine tuning the weights specifically for caloric expenditure estimation. Similarly, for tasks such as step count estimation, the step count adapter 326 may be fine-tuned for that particular task. When an adapter is loaded for a specific task, the weights of the foundation model 310 can be modified accordingly to optimize performance for that task.

[0039] In one or more other implementations, instead of relying exclusively on the training data 210, which may include indirect calorimetry measurements obtained via metabolic masks to determine oxygen and carbon dioxide exchange during physical exertion, the ML model 220 may utilize large volumes of unlabeled motion data collected across an extended time interval. Such unlabeled data may have been sourced from diverse user interactions over a period of time, encompassing a wide range of exercise and daily movement activities. The ML model 220 can be trained using this motion data, which is captured by inertial measurement units (IMUs), allowing the ML model 220 to infer structural patterns in movement independent of explicit instruction regarding specific activity types. The ML model 220 can learn to differentiate and associate motion sequences based on similarity metrics derived from sensor data distributions.

[0040] In one or more implementations, a smaller set of labeled training data corresponding to predefined workout types can be included as supplemental data in the training data 210 and fed to the foundation model 310. Using this supplemental data, the ML model 220 can be calibrated via one or more adapters, including the caloric expenditure adapter 322, which can operate atop the foundation model 310. This adapter can function as a task-specific extension that translates generalized motion representations into caloric estimation outputs. In one or more implementations, the ML model 220 can generate caloric expenditure estimates across a heterogeneous set of activities supported by the electronic device 118. In one or more implementations, the ML model 220 can generate caloric expenditure estimates across a heterogeneous set of activities supported by the electronic device 118. The approach represents a departure from prior task-specific modeling methods used for the electronic device 114 and may be beneficial for applications that may benefit from generalized motion interpretation across diverse platforms.

[0041] In one or more implementations, the electronic device 118 may include additional embedded sensors, such as a heart rate sensor. This configuration can enable the electronic device 118, in addition to the electronic device 114, to contribute data beneficial for estimating caloric expenditure during physical activity. Each instance of the electronic device 114 may be previously configured using the training data 210 associated with a specific workout type, followed by development of a task-specific adapter (e.g., the caloric expenditure adapter 322) tailored to that activity. In one or more other implementations, the implementation involving the electronic device 118 may benefit from compatibility across multiple workout types simultaneously.

[0042] In one or more other implementations, any one of the electronic devices 110-118 may include a heart rate sensor that incorporates (1) physiological signal data acquired from any one of the electronic devices 110-118 and (2) a detection algorithm to provide (3) real-time monitoring and alert notifications. For example, when a motion-related metric is estimated by at least one of the electronic devices 110-130, the system may generate alert notifications to a user depending on user preferences. In one or more other implementations, motion-related metric information can be logged in memory in a privacy-sensitive manner and trackable via an on-device application of the electronic devices 110-118.

[0043] FIG. 4 illustrates an example of a training operational framework for device agnostic motion-related metric prediction in accordance with one or more implementations. Unlabeled motion data 430 may include motion-related signals including IMU data collected via onboard sensors, such as accelerometers and gyroscopes, from these electronic devices, which is then input to the foundation model 310. The foundation model 310 may be trained using unsupervised training such as masked language modeling, in which the task may involve predicting a missing token from a sequence, to encourage the foundation model 310 to form generalized motion representations from the unlabeled motion data 430.

[0044] As part of a pre-training process, a preprocessing module 410 may operate to convert the IMU data into motion descriptors, facilitating downstream processing and enabling effective learning of temporal dynamics. In one or more implementations, the system operation workflow includes the preprocessing module 410 into the data pipeline to address the presence of gravitational components in the IMU data originating from electronic devices, such as the electronic device 118 and/or the electronic device 114. In one or more implementations, the preprocessing module 410 is, or at least in part of, included in the ML module 220. In one or more other implementations, the preprocessing module 410 is separate from the ML module 220. For illustrative purposes, the preprocessing module 410 is outside the ML module 220 as a separate component.

[0045] In one or more implementations, the IMU data may include measurements acquired by accelerometers that measure linear acceleration, which inherently include both dynamic movement and a constant gravitational force vector directed toward the center of the Earth. As a result, raw accelerometer data contains a persistent gravitational bias that varies based on device orientation. This gravitational component can introduce ambiguity in the training of the foundation model 310, as the foundation model 310 would otherwise infer and compensate for gravitational influence implicitly during feature extraction. To mitigate this gravitational influence, the preprocessing module 410 may integrate measurement data from both the accelerometer and the gyroscope onboard the electronic device 114 to estimate and isolate the gravitational vector.

[0046] In one or more implementations, an application programming interface (API) call may be made to a third-party API 420 to pre-process and normalize the IMU data into a structured feature basis. In one or more implementations, the preprocessing module 410 may initiate an API call to the third-party API 420 as part of the gravity removal subprocess. The gravitational vector may then be separated from the IMU data and treated as an independent input, while the motion descriptors containing the normalized linear acceleration and angular rotation are processed as primary inputs to the foundation model 310. One or more gravitational components can be removed from the IMU data to generate normalized motion data (devoid of gravitational bias). The removal of gravitational components may involve estimating a gravitational vector from the IMU data and subtracting this gravitational vector, resulting in normalized motion data including linear acceleration signals and rotation signals independent of gravitational bias. This decomposition allows the foundation model 310 to interpret user movement without the confounding effects of static gravitational force, simplifying the learning process and enhancing the quality of extracted motion features.

[0047] The foundation model 310 may be configured to generalize across different electronic device types by abstracting the sensor modality rather than the hardware source, enabling sensor-based learning independent of whether the motion data originates from a smartwatch (e.g., the electronic device 114) or a mobile phone (e.g., the electronic device 110). This generalization may facilitate the transfer of learned motion representations to, for example, wireless earbuds (e.g., the electronic device 118) irrespective of the limited availability of historical motion data for that wearable electronic device.

[0048] The foundation model 310 may be trained on motion data collected from electronic device 110 or electronic device 114 of device types different than the wearable electronic device 118. For example, when the wearable electronic device is an earbud, the foundation model 310 may be trained using data from a phone (e.g., the electronic device 110) or a smartwatch (e.g., the electronic device 114). The foundation model 310 can be initially trained to extract generalized motion features from the IMU data. The motion profiles captured by the electronic device 110 and the electronic device 114 may vary significantly due to differing wear locations and orientations. The electronic device 118, for example, may exhibit distinct motion characteristics compared to handheld or wrist-worn devices. For example, data collected from the electronic device 110 held in one position or the electronic device 114 worn on the wrist of a user may differ substantially in orientation from data collected by the electronic device 118 placed in the ear of the user. By removing the gravitational bias at the outset of training, the ML model 220 can learn more transferable and orientation-invariant representations, which can then be adapted for use with the electronic device 118 despite the relative scarcity of labeled data for that hardware platform.

[0049] In one or more other implementations, additional pre-training augmentation techniques may be applied, including data augmentation operations that can modify the input signal to the foundation model 310 by adding perturbations to the IMU data. The foundation model 310 may be optimized through perturbations and transformation functions configured to introduce representational invariance properties. These perturbations may include data augmentation components that modify the motion descriptors while preserving semantic motion integrity. These perturbations may increase the robustness of the foundation model 310 by forcing the foundation model 310 to identify consistent and meaningful patterns in the IMU data. The transformation functions may enable the foundation model 310 to generalize across diverse device configurations and user-specific motion characteristics. This process may enable the ML model 220 to produce reliable outputs even when transferred across device modalities with varying sensor configurations and data densities.

[0050] In one or more implementations, the noise introduced into the IMU data pipeline can enhance the robustness and generalization capability of the foundation model 310. The data augmentation procedures can be applied to normalized motion signals that have undergone prior gravitational component removal. This data augmentation is intended to enforce invariance properties consistent with physical laws and to improve the ability of the foundation model 310 to learn meaningful motion representations.

[0051] In one or more implementations, the data augmentation can involve the application of rotational transformations to the IMU signals. After the gravity components have been removed, the orientation of the device—such as the electronic device 110 or the electronic device 114—can vary arbitrarily, but the physical interpretation of motion can remain invariant under such rotation. To encode this rotational invariance into the foundation model 310, the training pipeline may include synthetic rotations of the IMU data. Samples that are equivalent under such transformations may be treated as semantically identical, whereas samples that differ significantly in orientation or motion pattern may be treated as dissimilar. In one or more other implementations, the data augmentation can involve the manipulation of temporal proximity within time series data. For example, consecutive or temporally adjacent segments of motion data can be labeled as similar, whereas segments that are temporally distant can be labeled as dissimilar.

[0052] To adapt the foundation model 310 for application to the electronic device 118, the ML model 220 may be configured to include one or more task-specific adapters, such as the caloric expenditure adapter 322. The aforementioned data augmentation operations collectively may enable the foundation model 310 to learn a representation of motion data that is invariant to orientation and temporally coherent, improving downstream task performance in task-specific adapters such as the caloric expenditure adapter 322.

[0053] Following training of the foundation model 310, a post-training stage may be employed to bridge the learned representations from the foundation model 310 to motion-related estimation outputs. For example, proxy labels, such as oxygen consumption approximations, can be introduced this post-training stage. These proxy labels can serve as intermediate targets, allowing the ML model 220 to begin learning a mapping between IMU-derived features and motion-related metric indicators (e.g., caloric-related indicators).

[0054] In one or more other implementations, an adaptation stage may be employed using a limited dataset containing ground truth motion-related metric data. This dataset can be collected through controlled measurements involving the electronic device 118, where actual motion-related metrics can be measured. During this adaptation stage, the foundation model 310 may undergo fine-tuning to align its outputs with precise motion-related metric values. The adaptation stage may involve varied training sample sizes and a range of loss functions incorporated into the foundation model 310 to improve model generalization and stability. In one or more implementations, a loss function is used to assess model performance on the caloric expenditure task (e.g., the caloric expenditure adapter 322) as well as its potential applicability to other downstream tasks, such as the gait adapter 324, the step count adapter 326, or the metabolic equivalents adapter 328, among others.

[0055] Each task-specific adapter is configured to enable adaptation of the foundation model 310 to different user activity types by training one or more parameters of the adapter utilizing few-shot learning with a subset of motion data associated with the wearable electronic device 118. In one or more implementations, few-shot learning can be used to facilitate the incorporation of additional task-specific adapters such as the gait adapter 324, the step count adapter 326, or other motion classification tasks. The adaptation process relies on fine-tuning mechanisms whereby the adapter layers learn to translate motion features extracted by the foundation model 310 into task-specific outputs. For example, a user may perform a particular motion, and the task-specific adapter may capture the corresponding embeddings generated by the foundation model 310 from the input sensor data. These embeddings can represent latent features of the motion instance. For instance, repeated identification of a specific embedding pattern may be classified as "waving an arm," "waving a leg," or "playing a sport." The embeddings may then be used as input to lightweight code modules or classifiers that associate specific motion signatures with identified activities. In one or more implementations, the task-specific adapter may be applied to the output embeddings of the foundation model 310. This process may be performed independently of retraining the foundation model 310. The task-specific adapter may be trained on a limited training dataset from the electronic device 118 including devices of device types similar to the electronic device 118. For example, the caloric expenditure adapter 322 may be fine-tuned using a combination of historical labeled data in the training data 210 and limited available data from the electronic device 118. In one or more implementations, a small subset of labeled training data 210 may be utilized to supervise the outputs of the foundation model 310, which are then used as input features to train a downstream task-specific adapter such as the caloric expenditure adapter 322 or the step count adapter 326.

[0056] In one or more other implementations, personalization may be achieved by incorporating user-specific data into the foundation model 310 through one of the task-specific adapters (e.g., 322-328). For example, for long-term users of the electronic device 114, historical data collected and stored in memory of the electronic device 114 may be utilized to refine the ML model 220 output, potentially enhancing accuracy for that particular user. This approach may involve on-device training of a task-specific adapter (e.g., 322-328) using personalized data to generate a customized version of the task-specific adapter tailored to characteristics of the user or particular user activity.

[0057] FIG. 5 illustrates an example of an inference operational framework for device agnostic motion-related metric prediction in accordance with one or more implementations. In one or more implementations, the system operational framework includes a pipeline in which motion-related data is used to produce caloric expenditure estimations via a multi-stage machine learning process. The system operational framework allows raw IMU signals collected from the electronic device 118 to be streamed to the electronic device 110. The electronic device 110 may execute inference via the ML model 220, generating a compact and semantically rich feature vector that encapsulates the motion event. This feature vector is then utilized as input to the caloric expenditure adapter 322 to estimate caloric expenditure.

[0058] This system operational framework can enable real-time estimation of energy expenditure directly from motion signals captured during user activity. The subject technology is constructed to perform robustly in a deployment scenario where minimal historical data is available for a target device (e.g., the electronic device 118). The system architecture facilitates scalability and adaptability across diverse hardware platforms by leveraging generalized representations learned from other device data (e.g., from the electronic device 114 and/or the electronic device 110).

[0059] In one or more implementations, wearable sensor data associated with a subject is received from a wearable electronic device (e.g., the electronic device 114, the electronic device 118). The wearable sensor data may include inertial measurement unit (IMU) data, gyroscope data, accelerometer data, and heart rate data obtained from the wearable electronic device 118. In one or more implementations, IMU data can be acquired from the electronic device 118. This IMU data can be transmitted to the electronic device 110 (paired with the electronic device 118), which may serve as a local inference environment for executing the ML model 220.

[0060] In one or more implementations, the foundation model 310 can produce a feature vector that includes a motion representation. For example, the foundation model 310 processes the normalized motion data to generate one or more vector representations. At least a portion of these vector representations is then processed through at least one task-specific adapter, such as a caloric expenditure adapter 322 (FIG. 3), gait adapter 324 (FIG. 3), step count adapter 326 (FIG. 3), or metabolic equivalents adapter 328 (FIG. 3), associated with the foundation model 310 to produce one or more motion-related metric predictions.

[0061] For illustrative purposes, the intermediate output from the foundation model 310 can be processed by the caloric expenditure adapter 322. The caloric expenditure adapter 322 may be trained to map outputs of the foundation model 310 to an estimation of caloric expenditure based on motion features corresponding to various physical activity types. The motion-related metric predictions produced may include at least one of gait metric estimation, metabolic equivalent of task (MET) computation, step count tracking, caloric expenditure estimation, activity classification, and contextual awareness of the physical state and surroundings of the subject.

[0062] In one or more implementations, the foundation model 310 may support multiple downstream tasks by generating a general-purpose representation of inertial sensor data that can be reused across a range of applications. These applications may include detection of physical activity types, such as distinguishing between a yoga session and a high-intensity interval workout, and estimation of discrete movement events, such as step counting via the step count adapter 326.

[0063] A benefit of utilizing the foundation model 310 in conjunction with task-specific adapters such as the caloric expenditure adapter 322 is that the system may generalize to previously unseen activity types without retraining the foundation model 310 with new motion data. For example, if a physical activity such as “pickleball” becomes widely adopted, the system may provide reasonable caloric expenditure estimates for that activity without requiring full-scale collection of new motion data into the training data 210 or retraining the ML model 220. While limited validation data collection (e.g., using indirect calorimetry) may still be beneficial for evaluating the ML model 220, the amount of data and processing needed to adapt the system may be significantly reduced.

[0064] In operation, such as during exercise sessions involving the electronic device 118, a user may also carry the electronic device 110 or wear the electronic device 114. In scenarios where the electronic device 110 is stationary, such as being placed on a table during treadmill use, and the electronic device 114 is not worn, the ML model 220 may continue to generate outputs based solely on data from the electronic device 118. The subject technology can be configured to function without reliance on the electronic device 114, providing activity inference and caloric estimation capabilities from data collected only from the electronic device 118 and optionally from the electronic device 110. In one or more implementations, sensor embeddings from both the electronic device 114 and the electronic device 118 may be used concurrently within the ML model 220 via a shared input adapter (not shown) configured to process embeddings from multiple electronic devices. This can enable integration across the electronic device 114 and the electronic device 118, facilitating combined use of embeddings derived from each electronic device.

[0065] For illustrative purposes, inference performed by the ML model 220 may be executed on the electronic device 110 during real-time data collection. In cases where the electronic device 110 may not be present or connected to either of the electronic device 114 or the electronic device 118, such as when the user leaves behind the electronic device 110 at home and listens to music via the electronic device 118 paired with the electronic device 114, real-time inference may be interrupted. The electronic device 118 may have limited storage capacity, which may reduce local data retention capabilities. To mitigate this, sensor readings from the electronic device 118 may be recorded and temporarily stored on other connected devices, such as the electronic device 114, allowing deferred processing of stored sensor data. Recorded sensor data stored in memory of these electronic devices may be transferred to the electronic device 110 (when the electronic device 110 is locally available) to enable post hoc inference of caloric expenditure or activity estimation by the electronic device 110.

[0066] FIG. 6 is a flow chart of an example process that may be performed for device agnostic motion-related metric prediction in accordance with one or more implementations. For explanatory purposes, the process 600 is primarily described herein with reference to the electronic device 110 of FIG. 1. However, the process 600 is not limited to the electronic device 110 of FIG. 1, and one or more blocks (or operations) of the process 600 may be performed by one or more other components of other suitable devices and/or servers. Further for explanatory purposes, some of the blocks of the process 600 are described herein as occurring in serial, or linearly. However, multiple blocks of the process 600 may occur in parallel. In addition, the blocks of the process 600 need not be performed in the order shown and/or one or more blocks of the process 600 need not be performed and/or can be replaced by other operations.

[0067] As illustrated in FIG. 6, at block 602, an apparatus (e.g., the electronic device 110 of FIG. 1; ML model 220 of FIG. 2; foundation model 310 of FIG. 3; the caloric expenditure adapter 322, the gait adapter 324, the step count adapter 326, or the metabolic equivalents adapter 328 of FIG. 3) receives wearable sensor data associated with a subject from a wearable electronic device, such as the electronic device 114 (FIG. 1) and/or the electronic device 118 (FIG. 1). The wearable sensor data may include one or more of IMU data, gyroscope data, accelerometer data, and/or heart rate data from the electronic device 118.

[0068] At block 604, the apparatus removes one or more gravitational components from the wearable sensor data to generate normalized motion data. Removing the gravitational components includes estimating a gravitational vector from the wearable sensor data, in which the gravitational vector includes the gravitational components. The gravitational vector can be subtracted from the wearable sensor data to generate normalized motion data having linear acceleration signals and rotation signals that are independent of gravitational bias.

[0069] At block 606, one or more vector representations of the normalized motion data are generated using a foundation model 310. The foundation model 310 may be trained using motion data collected from one or more electronic devices of types different from the electronic device 114. For example, the motion data used to train the foundation model 310 may have been collected from the electronic device 110 or the electronic device 114.

[0070] At block 608, the apparatus produces one or more motion-related metric predictions by processing at least a portion of the one or more vector representations through at least one task-specific adapter associated with the foundation model 310. At least a portion of the one or more vector representations is processed using at least one task-specific adapter associated with the foundation model 310 to produce one or more motion-related metric predictions. The motion-related metric predictions may include, for instance, estimations of gait metrics, calculations of metabolic equivalents of task (MET), step count tracking, caloric expenditure estimation, activity classification, or contextual awareness relating to a user’s physical state and environmental conditions. The at least one task-specific adapter—such as the caloric expenditure adapter 322, GAIT adapter 324, step count adapter 326, or metabolic equivalents adapter 328—is applied to output embeddings of the foundation model 310. Each of these adapters is configured to adapt the foundation model 310 to a corresponding user activity type. This adaptation is accomplished by training one or more parameters of the respective adapter using few-shot learning and a subset of motion data specific to the electronic device 114, without retraining the foundation model 310 itself. The task-specific adapters may be trained using a limited dataset derived from electronic devices similar to the electronic device 118, for example.

[0071] FIG. 7 is a flow chart of an example process that may be performed for device agnostic motion-related metric prediction in accordance with one or more implementations. For explanatory purposes, the process 700 is primarily described herein with reference to the electronic device 110 of FIG. 1. However, the process 700 is not limited to the electronic device 110 of FIG. 1, and one or more blocks (or operations) of the process 700 may be performed by one or more other components of other suitable devices and/or servers. Further for explanatory purposes, some of the blocks of the process 700 are described herein as occurring in serial, or linearly. However, multiple blocks of the process 700 may occur in parallel. In addition, the blocks of the process 700 need not be performed in the order shown and/or one or more blocks of the process 700 need not be performed and/or can be replaced by other operations.

[0072] As illustrated in FIG. 7, at block 702, an electronic device (e.g., the electronic device 110 of FIG. 1; ML model 220 of FIG. 2; foundation model 310 of FIG. 3; the caloric expenditure adapter 322, the gait adapter 324, the step count adapter 326, or the metabolic equivalents adapter 328 of FIG. 3) that includes one or more sensors is configured to receive wearable sensor data from a wearable electronic device (e.g., the electronic device 114, the electronic device 118) associated with a subject.

[0073] At block 704, the electronic device includes a processor that is configured to remove gravitational components from the wearable sensor data to produce normalized motion data. The processor may be further configured to estimate a gravitational vector from the wearable sensor data and subtract the estimated gravitational vector, resulting in normalized motion data comprising linear acceleration and rotation signals that are not influenced by gravity.

[0074] At block 706, the processor generates vector representations of the normalized motion data using the foundation model 310. The foundation model 310 may be trained using motion data collected from electronic devices of a type different from the electronic device 114. For example, if the electronic device is the electronic device 118, then the training data 210 may have been sourced from the electronic device 110 or the electronic device 114. The processor may adapt the foundation model 310 to different user activity types by training one or more parameters of the task-specific adapter (e.g., caloric expenditure adapter 322, GAIT adapter 324, step count adapter 326, metabolic equivalents adapter 328) using few-shot learning on a limited amount of motion data from the electronic device 114. The adaptation occurs without retraining the foundation model 310.

[0075] At block 708, the processor may process at least a portion of the vector representations by applying at least one of the task-specific adapters associated with the foundation model 310 to produce one or more motion-related metric predictions. The foundation model 310 may be associated with multiple task-specific adapters, each configured to produce at least one of the motion-related metric predictions.

[0076] FIG. 8 is a flow chart of an example process that may be performed for device agnostic motion-related metric prediction in accordance with one or more implementations. For explanatory purposes, the process 800 is primarily described herein with reference to the electronic device 110 of FIG. 1. However, the process 800 is not limited to the electronic device 110 of FIG. 1, and one or more blocks (or operations) of the process 800 may be performed by one or more other components of other suitable devices and/or servers. Further for explanatory purposes, some of the blocks of the process 800 are described herein as occurring in serial, or linearly. However, multiple blocks of the process 800 may occur in parallel. In addition, the blocks of the process 800 need not be performed in the order shown and/or one or more blocks of the process 800 need not be performed and/or can be replaced by other operations.

[0077] As illustrated in FIG. 8, at block 802, an apparatus may be a non-transitory computer-readable medium that stores instructions which, when executed by one or more processors of an electronic device (e.g., the electronic device 110 of FIG. 1; ML model 220 of FIG. 2; foundation model 310 of FIG. 3; the caloric expenditure adapter 322, the gait adapter 324, the step count adapter 326, or the metabolic equivalents adapter 328 of FIG. 3) that causes the electronic device to receive wearable sensor data from a wearable electronic device (e.g., the electronic device 114, the electronic device 118) associated with a subject. For example, the wearable sensor data used for these processes may include IMU, gyroscope, accelerometer, and heart rate data collected from the electronic device 118.

[0078] At block 804, the apparatus may further cause the electronic device to remove gravitational components from the data to produce normalized motion data. The instructions stored on the non-transitory computer-readable medium may further cause the processor to estimate a gravitational vector from the wearable sensor data and subtract the gravitational vector such that the resulting normalized motion data comprises linear acceleration signals and rotational signals devoid of gravitational bias.

[0079] At block 806, the apparatus may further cause the electronic device to generate one or more vector representations of the normalized motion data using the foundation model 310. The foundation model 310 may be trained on motion data collected from electronic devices that differ in type from the electronic device 114. Furthermore, the instructions may cause the processor to adapt the foundation model 310 to various user activity types by updating parameters of the task-specific adapter via few-shot learning using motion data from the electronic device 114, without requiring retraining of the foundation model 310.

[0080] At block 808, the apparatus may further cause the electronic device to provide at least a portion of the one or more vector representations to at least one task-specific adapter associated with the foundation model 310 to produce one or more motion-related metric predictions. The foundation model 310 may be associated with a plurality of task-specific adapters, each configured to produce one or more of the motion-related metric predictions, such as gait estimation, metabolic equivalent computation, step counting, caloric expenditure determination, activity classification, and assessment of user physical context.

[0081] FIG. 9 illustrates an electronic system 900 with which one or more implementations of the subject technology may be implemented. The electronic system 900 can be, and/or can be a part of, the electronic device 110, and/or the server 120 shown in FIG. 1. The electronic system 900 may include various types of computer readable media and interfaces for various other types of computer readable media. The electronic system 900 includes a bus 908, one or more processing unit(s) 912, a system memory 904 (and/or buffer), a ROM 910, a permanent storage device 902, an input device interface 914, an output device interface 906, and one or more network interfaces 916, or subsets and variations thereof.

[0082] The bus 908 collectively represents all system, peripheral, and chipset buses that communicatively connect the numerous internal devices of the electronic system 900. In one or more implementations, the bus 908 communicatively connects the one or more processing unit(s) 912 with the ROM 910, the system memory 904, and the permanent storage device 902. From these various memory units, the one or more processing unit(s) 912 retrieves instructions to execute and data to process in order to execute the processes of the subject disclosure. The one or more processing unit(s) 912 can be a single processor or a multi-core processor in different implementations.

[0083] The ROM 910 stores static data and instructions that are needed by the one or more processing unit(s) 912 and other modules of the electronic system 900. The permanent storage device 902, on the other hand, may be a read-and-write memory device. The permanent storage device 902 may be a non-volatile memory unit that stores instructions and data even when the electronic system 900 is off. In one or more implementations, a mass-storage device (such as a magnetic or optical disk and its corresponding disk drive) may be used as the permanent storage device 902.

[0084] In one or more implementations, a removable storage device (such as a flash drive, and its corresponding solid-state drive) may be used as the permanent storage device 902. Like the permanent storage device 902, the system memory 904 may be a read-and-write memory device. However, unlike the permanent storage device 902, the system memory 904 may be a volatile read-and-write memory, such as random-access memory. The system memory 904 may store any of the instructions and data that one or more processing unit(s) 912 may need at runtime. In one or more implementations, the processes of the subject disclosure are stored in the system memory

904, the permanent storage device 902, and/or the ROM 910. From these various memory units, the one or more processing unit(s) 912 retrieves instructions to execute and data to process in order to execute the processes of one or more implementations.

[0085] The bus 908 also connects to the input device interface 914 and output device interface 906. The input device interface 914 enables a user to communicate information and select commands to the electronic system 900. Input devices that may be used with the input device interface 914 may include, for example, alphanumeric keyboards and pointing devices (also called "cursor control devices"). The output device interface 906 may enable, for example, the display of images generated by electronic system 900. Output devices that may be used with the output device interface 906 may include, for example, printers and display devices, such as a liquid crystal display (LCD), a light emitting diode (LED) display, an organic light emitting diode (OLED) display, a flexible display, a flat panel display, a solid state display, a projector, or any other device for outputting information. One or more implementations may include devices that function as both input and output devices, such as a touchscreen. In these implementations, feedback provided to the user can be any form of sensory feedback, such as visual feedback, auditory feedback, or tactile feedback; and input from the user can be received in any form, including acoustic, speech, or tactile input.

[0086] Finally, as shown in FIG. 9, the bus 908 also couples the electronic system 900 to one or more networks and/or to one or more network nodes, such as the electronic device 110 shown in FIG. 1, through the one or more network interface(s) 916. In this manner, the electronic system 900 can be a part of a network of computers (such as a LAN, a wide area network ("WAN"), or an Intranet, or a network of networks, such as the Internet. Any or all components of the electronic system 900 can be used in conjunction with the subject disclosure.

[0087] Implementations within the scope of the present disclosure can be partially or entirely realized using a tangible computer-readable storage medium (or multiple tangible computer-readable storage media of one or more types) encoding one or more instructions. The tangible computer-readable storage medium also can be non-transitory in nature.

[0088] The computer-readable storage medium can be any storage medium that can be read, written, or otherwise accessed by a general purpose or special purpose computing device, including any processing electronics and/or processing circuitry capable of executing instructions. For example, without limitation, the computer-readable medium can include any volatile semiconductor memory, such as RAM, DRAM, SRAM, T-RAM, Z-RAM, and TTRAM. The computer-readable medium also can include any non-volatile semiconductor memory, such as ROM, PROM, EPROM, EEPROM, NVRAM, flash, nvSRAM, FeRAM, FeTRAM, MRAM, PRAM, CBRAM, SONOS, RRAM, NRAM, racetrack memory, FJG, and Millipede memory.

[0089] Further, the computer-readable storage medium can include any non-semiconductor memory, such as optical disk storage, magnetic disk storage, magnetic tape, other magnetic storage devices, or any other medium capable of storing one or more instructions. In one or more implementations, the tangible computer-readable storage medium can be directly coupled to a computing device, while in other implementations, the tangible computer-readable storage medium can be indirectly coupled to a computing device, e.g., via one or more wired connections, one or more wireless connections, or any combination thereof.

[0090] Instructions can be directly executable or can be used to develop executable instructions. For example, instructions can be realized as executable or non-executable machine code or as instructions in a high-level language that can be compiled to produce executable or non-executable machine code. Further, instructions also can be realized as or can include data. Computer-executable instructions also can be organized in any format, including routines, subroutines, programs, data structures, objects, modules, applications, applets, functions, etc. As recognized by those of skill in the art, details including, but not limited to, the number, structure, sequence, and organization of instructions can vary significantly without varying the underlying logic, function, processing, and output.

[0091] While the above discussion primarily refers to microprocessor or multi-core processors that execute software, one or more implementations are performed by one or more integrated circuits, such as ASICs or FPGAs. In one or more implementations, such integrated circuits execute instructions that are stored on the circuit itself.

[0092] Those of skill in the art would appreciate that the various illustrative blocks, modules, elements, components, methods, and algorithms described herein may be implemented as electronic hardware, computer software, or combinations of both. To illustrate this interchangeability of hardware and software, various illustrative blocks, modules, elements, components, methods, and algorithms have been described above generally in terms of their functionality. Whether such functionality is implemented as hardware or software depends upon the particular application and design constraints imposed on the overall system. Skilled artisans may implement the described functionality in varying ways for each particular application. Various components and blocks may be arranged differently (e.g., arranged in a different order, or partitioned in a different way) all without departing from the scope of the subject technology.

[0093] It is understood that any specific order or hierarchy of blocks in the processes disclosed is an illustration of example approaches. Based upon design preferences, it is understood that the specific order or hierarchy of blocks in the processes may be rearranged, or that all illustrated blocks be performed. Any of the blocks may be performed simultaneously. In one or more implementations, multitasking and parallel processing may be advantageous. Moreover, the separation of various system components in the implementations described above should not be understood as requiring such separation in all implementations, and it should be understood that the described program components and systems can generally be integrated together in a single software product or packaged into multiple software products.

[0094] As used in this specification and any claims of this application, the terms “base station”, “receiver”, “computer”, “server”, “processor”, and “memory” all refer to electronic or other technological devices. These terms exclude people or groups of people. For the purposes of the specification, the terms “display” or “displaying” means displaying on an electronic device.

[0095] As used herein, the phrase “at least one of” preceding a series of items, with the term “and” or “or” to separate any of the items, modifies the list as a whole, rather than each member of the list (i.e., each item). The phrase “at least one of” does not require selection of at least one of each item listed; rather, the phrase allows a meaning that includes at least one of any one of the items, and/or at least one of any combination of the items, and/or at least one of each of the items. By way of example, the phrases "at least one of A, B, and C" or "at least one of A, B, or C" each refer to only A, only B, or only C; any combination of A, B, and C; and/or at least one of each of A, B, and C.

[0096] The predicate words "configured to", "operable to", and "programmed to" do not imply any particular tangible or intangible modification of a subject, but, rather, are intended to be used interchangeably. In one or more implementations, a processor configured to monitor and control an operation or a component may also mean the processor being programmed to monitor and control the operation or the processor being operable to monitor and control the operation. Likewise, a processor configured to execute code can be construed as a processor programmed to execute code or operable to execute code.

[0097] Phrases such as an aspect, the aspect, another aspect, some aspects, one or more aspects, an implementation, the implementation, another implementation, some implementations, one or more implementations, an embodiment, the embodiment, another embodiment, some implementations, one or more implementations, a configuration, the configuration, another configuration, some configurations, one or more configurations, the subject technology, the disclosure, the present disclosure, other variations thereof and alike are for convenience and do not imply that a disclosure relating to such phrase(s) is essential to the subject technology or that such disclosure applies to all configurations of the subject technology. A disclosure relating to such phrase(s) may apply to all configurations, or one or more configurations. A disclosure relating to such phrase(s) may provide one or more examples. A phrase such as an aspect or some aspects may refer to one or more aspects and vice versa, and this applies similarly to other foregoing phrases.

[0098] The word "exemplary" is used herein to mean "serving as an example, instance, or illustration". Any embodiment described herein as "exemplary" or as an "example" is not necessarily to be construed as preferred or advantageous over other implementations. Furthermore, to the extent that the term "include", "have", or the like is used in the description or the claims, such term is intended to be inclusive in a manner similar to the term "comprise" as "comprise" is interpreted when employed as a transitional word in a claim.

[0099] All structural and functional equivalents to the elements of the various aspects described throughout this disclosure that are known or later come to be known to those of ordinary skill in the art are expressly incorporated herein by reference and are intended to be encompassed by the claims. Moreover, nothing disclosed herein is intended to be dedicated to the public regardless of whether such disclosure is explicitly recited in the claims. No claim element is to be construed under the provisions of 35 U.S.C. § 112(f) unless the element is expressly recited using the phrase “means for” or, in the case of a method claim, the element is recited using the phrase “step for”.

[0100] The previous description is provided to enable any person skilled in the art to practice the various aspects described herein. Various modifications to these aspects will be readily apparent to those skilled in the art, and the generic principles defined herein may be applied to other aspects. Thus, the claims are not intended to be limited to the aspects shown herein, but are to be accorded the full scope consistent with the language claims, wherein reference to an element in the singular is not intended to mean “one and only one” unless specifically so stated, but rather “one or more”. Unless specifically stated otherwise, the term “some” refers to one or more. Pronouns in the masculine (e.g., his) include the feminine and neuter gender (e.g., her and its) and vice versa. Headings and subheadings, if any, are used for convenience only and do not limit the subject disclosure.

CLAIMS

What is claimed is:

1. A method comprising:

receiving wearable sensor data associated with a subject from a wearable electronic device;

removing one or more gravitational components from the wearable sensor data to generate normalized motion data;

generating one or more vector representations of the normalized motion data using a machine learning model; and

producing one or more motion-related metric predictions by processing at least a portion of the one or more vector representations through the machine learning model.

2. The method of claim 1, wherein the wearable sensor data comprises one or more of inertial measurement unit (IMU) data, gyroscope data, accelerometer data, and heart rate data from the wearable electronic device.

3. The method of claim 1, wherein removing the one or more gravitational components comprises:

estimating a gravitational vector from the wearable sensor data, the gravitational vector comprising the one or more gravitational components; and

subtracting the gravitational vector from the wearable sensor data, wherein the normalized motion data comprises linear acceleration signals and rotation signals that are independent of gravitational bias.

4. The method of claim 1, wherein the machine learning model is trained on motion data collected from electronic devices of device types different than the wearable electronic device.

5. The method of claim 4, wherein the wearable electronic device is an ear bud and the electronic devices of device types different than the wearable electronic device is a phone or a smartwatch.

6. The method of claim 1, wherein the machine learning model includes a foundation model and at least one task-specific adapter associated with the foundation model, further comprising applying the at least one task-specific adapter to output embeddings of the

foundation model, wherein each of the at least one task-specific adapter is configured to enable the foundation model to adapt to different user activity types by training one or more parameters of the at least one task-specific adapter utilizing few-shot learning with a subset of motion data associated with the wearable electronic device independent of retraining the foundation model.

7. The method of claim 6, wherein the at least one task-specific adapter is trained on a limited training dataset from electronic devices of device types similar to the wearable electronic device.

8. The method of claim 1, wherein the one or more motion-related metric predictions comprises at least one of gait metric estimation, metabolic equivalent of task (MET) computation, step count tracking, caloric expenditure estimation, activity classification, and contextual awareness of a physical state and surroundings of the subject.

9. An electronic device comprising:

one or more sensors configured to receive wearable sensor data from a wearable electronic device associated with a subject; and

at least one processor configured to:

remove one or more gravitational components from the wearable sensor data to generate normalized motion data,

generate one or more vector representations of the normalized motion data using a foundation model, and

apply at least one task-specific adapter associated with the foundation model to at least a portion of the one or more vector representations to produce one or more motion-related metric predictions.

10. The electronic device of claim 9, wherein the at least one processor configured to remove the one or more gravitational components is further configured to:

estimate a gravitational vector from the wearable sensor data, the gravitational vector comprising the one or more gravitational components, and

subtract the gravitational vector from the wearable sensor data, wherein the normalized motion data comprises linear acceleration signals and rotation signals that are independent of gravitational bias.

11. The electronic device of claim 9, wherein the foundation model is trained on motion data collected from electronic devices of device types different than the wearable electronic device.

12. The electronic device of claim 11, wherein the at least one processor is further configured to adapt the foundation model to different user activity types by training one or more parameters of the at least one task-specific adapter through few-shot learning with a subset of motion data associated with the wearable electronic device independent of retraining the foundation model.

13. The electronic device of claim 11, wherein the wearable electronic device is an ear bud and at least one of the electronic devices of device types different than the wearable electronic device is a phone or a smartwatch.

14. The electronic device of claim 9, wherein the foundation model is associated with a plurality of task-specific adapters, and wherein each of the plurality of task-specific adapters is configured to produce at least one of the one or more motion-related metric predictions.

15. A non-transitory computer-readable medium storing instructions that, when executed by one or more processors of an electronic device, cause the electronic device to:

receive wearable sensor data from a wearable electronic device associated with a subject;

generate normalized motion data by removing one or more gravitational components from the wearable sensor data;

generate one or more vector representations of the normalized motion data using a machine learning model; and

provide at least a portion of the one or more vector representations to the machine learning model to produce one or more motion-related metric predictions.

16. The non-transitory computer-readable medium of claim 15, wherein the instructions, when executed by the one or more processors, further cause the electronic device to:

estimate a gravitational vector from the wearable sensor data, the gravitational vector comprising the one or more gravitational components, and

subtract the gravitational vector from the wearable sensor data, wherein the normalized motion data comprises linear acceleration signals and rotation signals that are independent of gravitational bias.

17. The non-transitory computer-readable medium of claim 15, wherein the machine learning model includes a foundation model and at least one task-specific adapter associated with the foundation model, wherein the foundation model is trained on motion data collected from electronic devices of device types different than the electronic device, and wherein the instructions, when executed by the one or more processors, further cause the electronic device to adapt the foundation model to different user activity types by training one or more parameters of the at least one task-specific adapter through few-shot learning with a subset of motion data associated with the wearable electronic device independent of retraining the foundation model.

18. The non-transitory computer-readable medium of claim 17, wherein the foundation model is associated with a plurality of task-specific adapters, and wherein each of the plurality of task-specific adapters is configured to produce at least one of the one or more motion-related metric predictions.

19. The non-transitory computer-readable medium of claim 15, wherein the wearable sensor data comprises one or more of inertial measurement unit (IMU) data, gyroscope data, accelerometer data, and heart rate data from the wearable electronic device.

20. The non-transitory computer-readable medium of claim 15, wherein the one or more motion-related metric predictions comprises at least one of gait metric estimation, metabolic equivalent of task (MET) computation, step count tracking, caloric expenditure estimation, activity classification, and contextual awareness of a physical state and surroundings of the subject.

ABSTRACT

The subject technology provides for device agnostic motion-related metric prediction for wearable electronic devices. Wearable sensor data from a wearable electronic device is processed by removing gravitational components to generate normalized motion data. A foundation model, trained on motion data from electronic devices of different types, generates vector representations of the normalized data. These representations are refined by one or more task-specific adapters, such as caloric expenditure or gait adapters, to produce motion-related metric predictions. The adapters enable adaptation to various user activities by training a limited number of parameters through few-shot learning on motion data specific to the wearable device, without retraining the foundation model. The motion-related metrics may include gait estimation, metabolic equivalent of task computation, step count, caloric expenditure, activity classification, and contextual awareness of the subject's physical state and environment. This modular approach facilitates efficient, scalable, and device-agnostic motion analysis using sensor data.

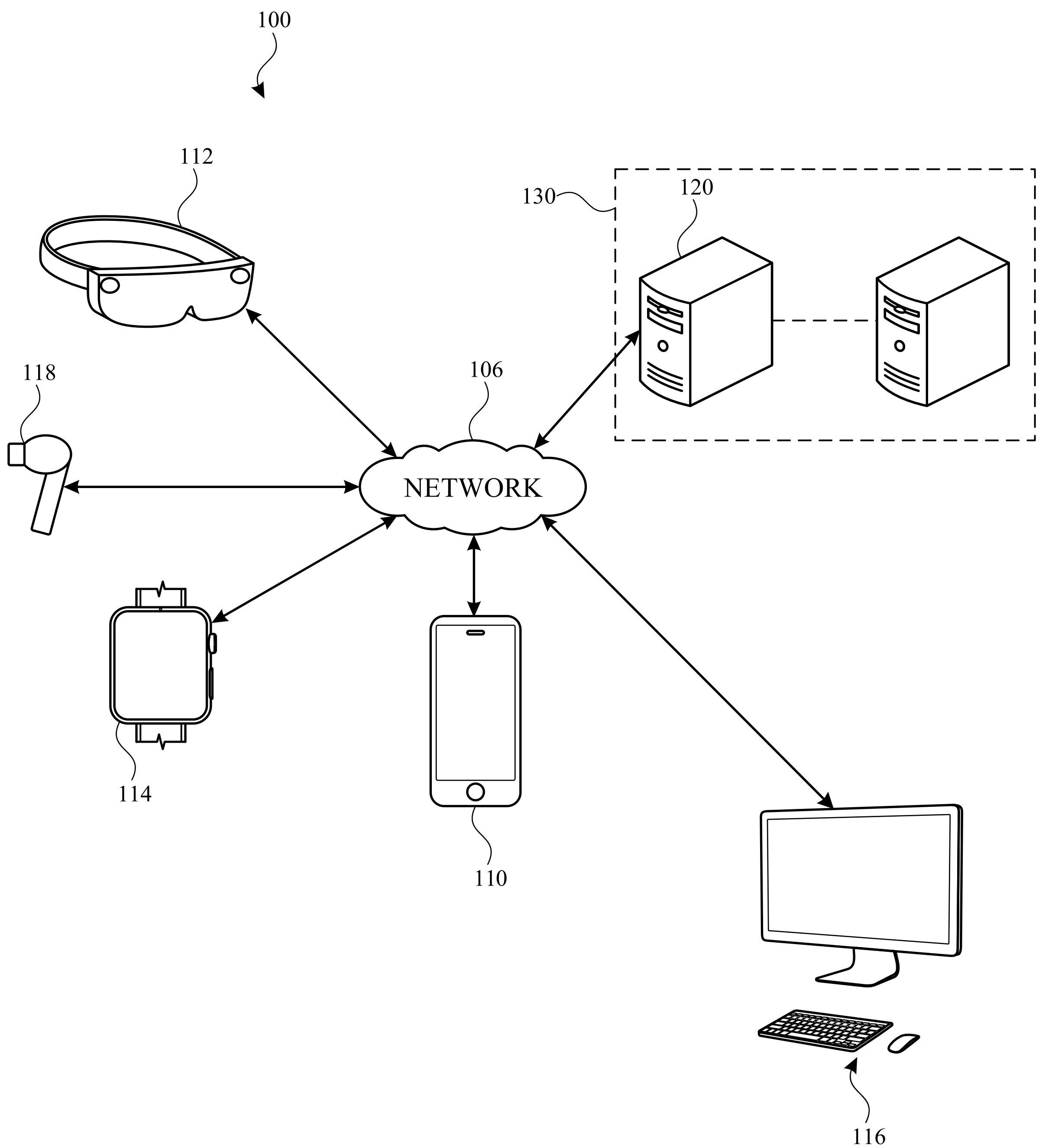

FIG. 1

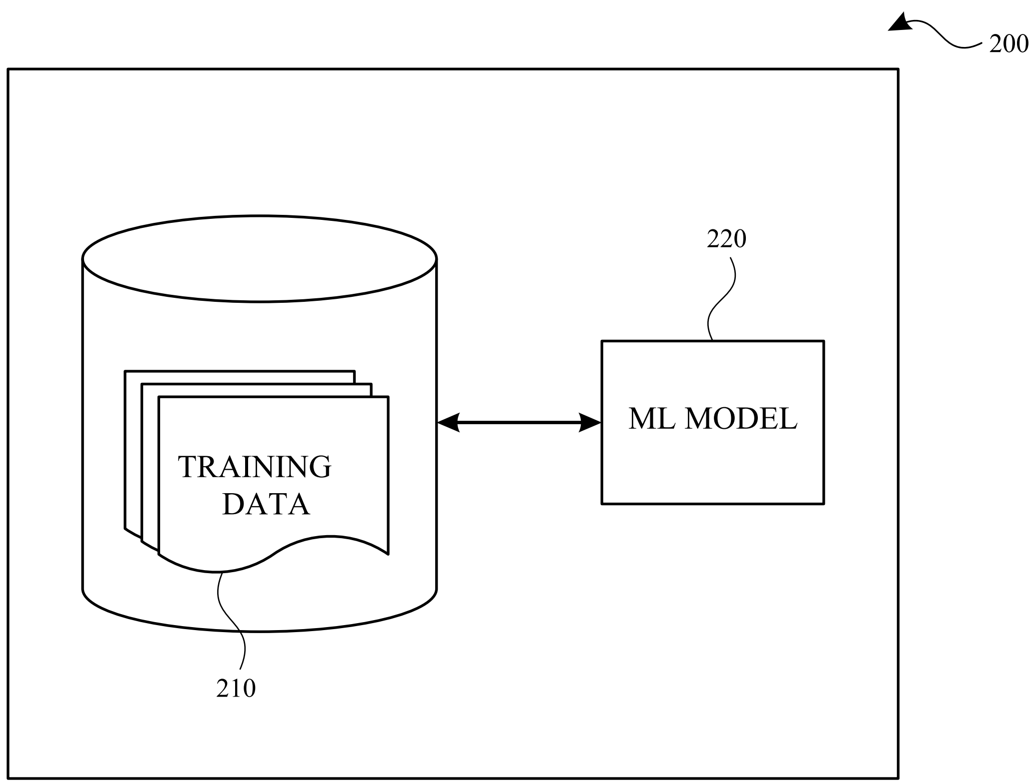

FIG. 2

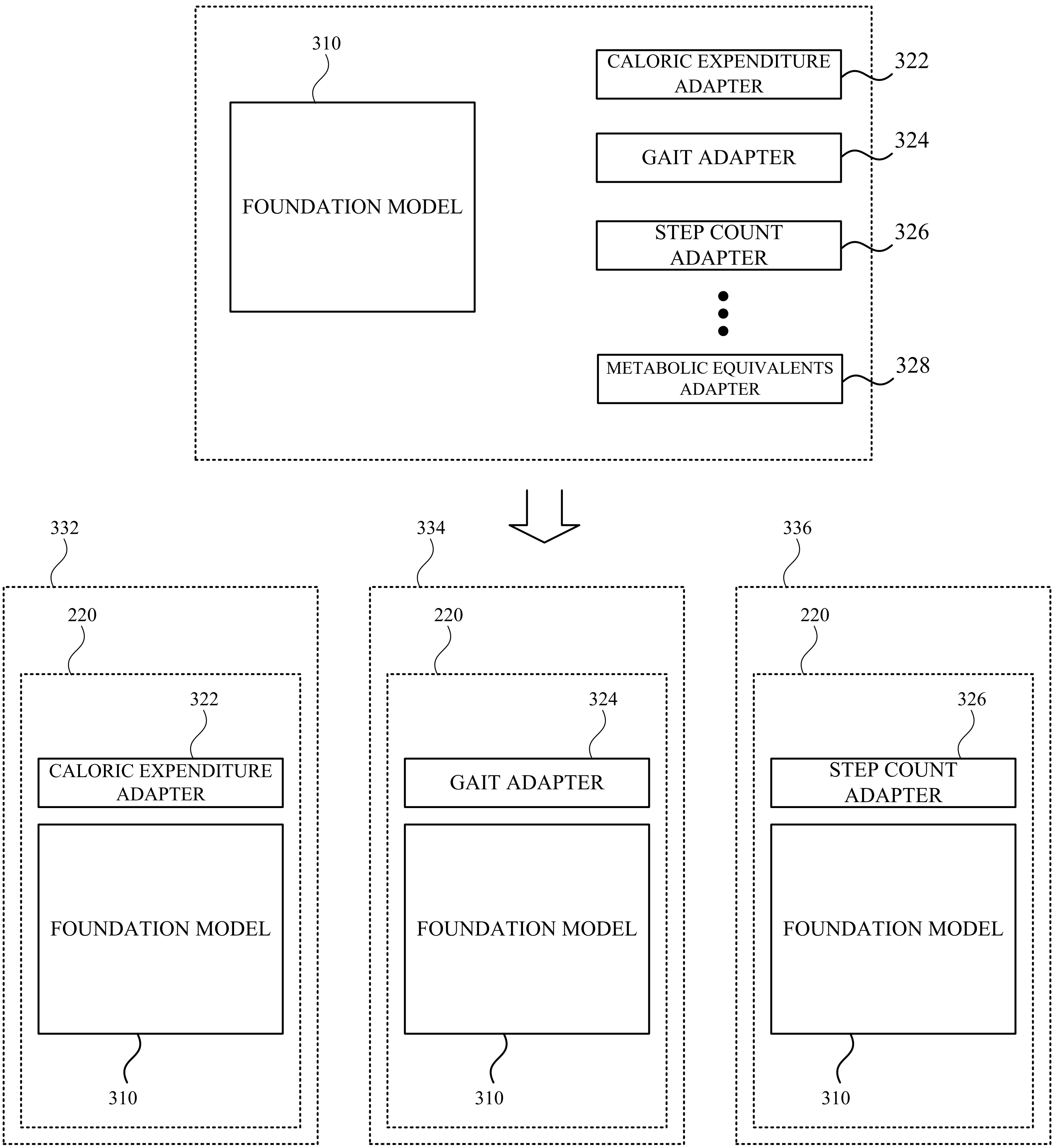
310
FOUNDATION MODEL
CALORIC EXPENDITURE ADAPTER
322
GAIT ADAPTER
324
STEP COUNT ADAPTER
326
METABOLIC EQUIVALENTS ADAPTER
328
332
220
322
CALORIC EXPENDITURE ADAPTER
FOUNDATION MODEL
310
334
220
324
GAIT ADAPTER
FOUNDATION MODEL
310
336
220
326
STEP COUNT ADAPTER
FOUNDATION MODEL
310

FIG. 3

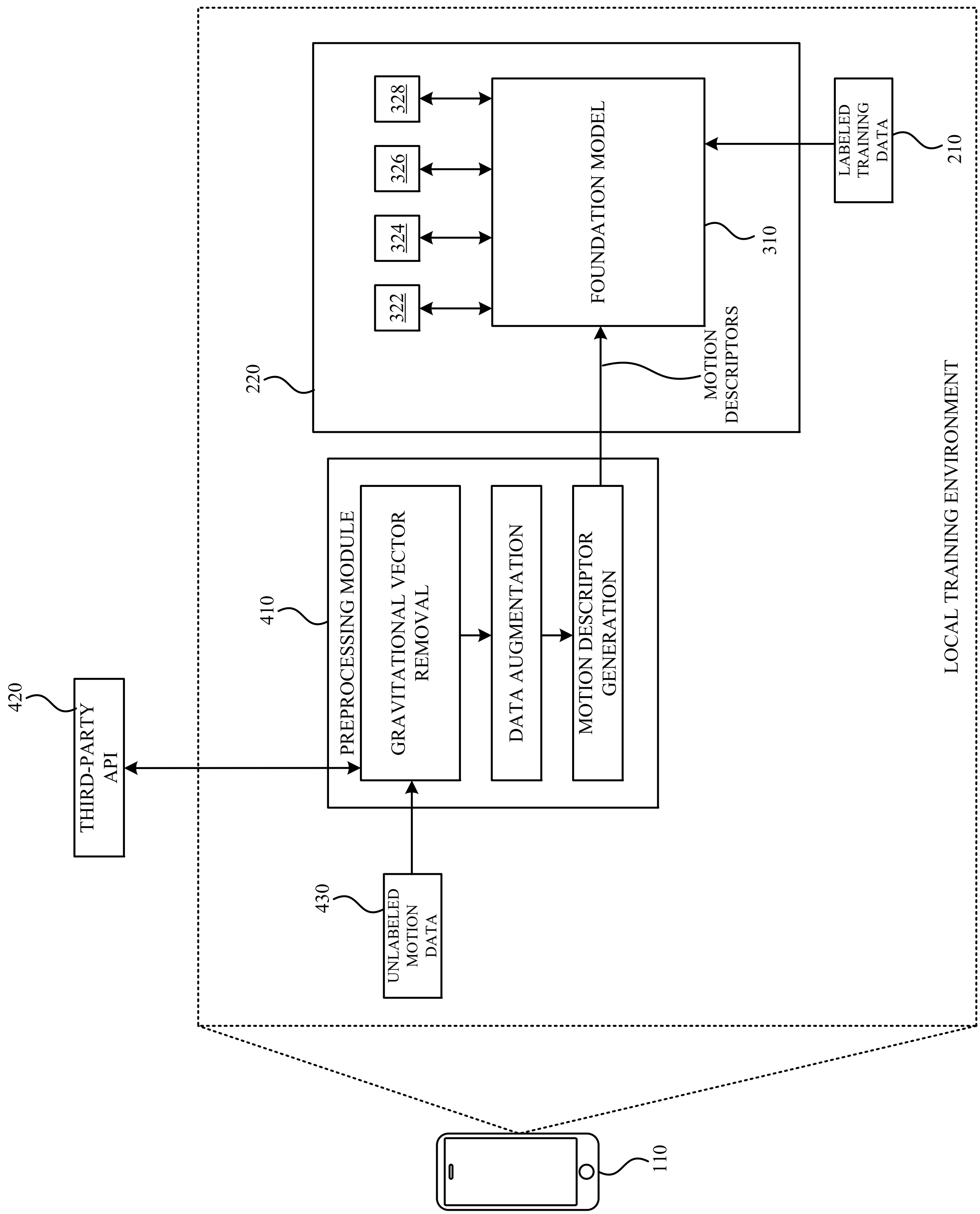
FOUNDATION MODEL
328
326
324
322
310
LABELED TRAINING DATA
210
220
MOTION DESCRIPTORS
PREPROCESSING MODULE
410
GRAVITATIONAL VECTOR REMOVAL
DATA AUGMENTATION
MOTION DESCRIPTOR GENERATION
LOCAL TRAINING ENVIRONMENT
THIRD-PARTY API
420
UNLABELED MOTION DATA
430
110

FIG. 4

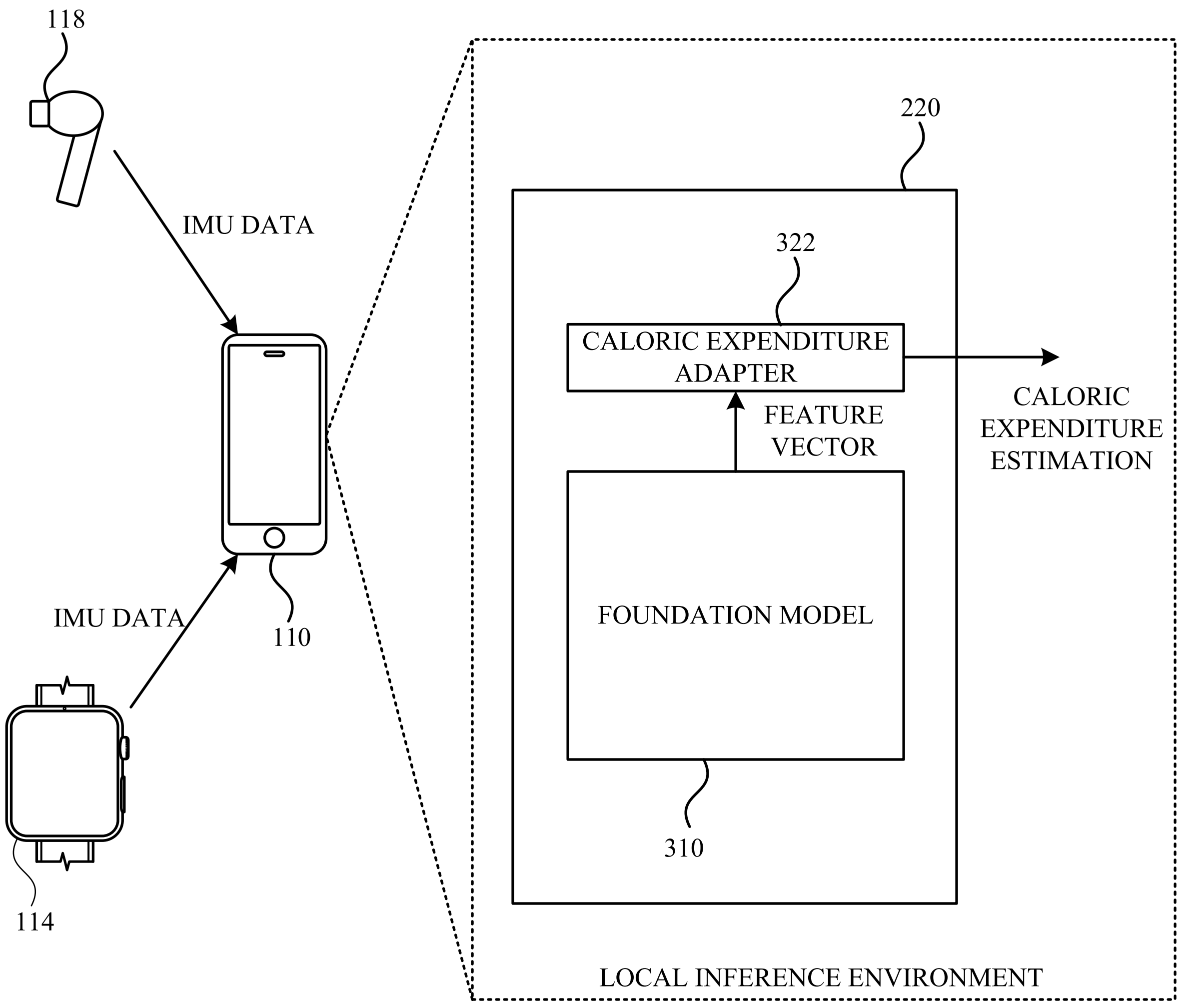
118
IMU DATA
110
IMU DATA
114
220
322
CALORIC EXPENDITURE ADAPTER
FEATURE VECTOR
CALORIC EXPENDITURE ESTIMATION
FOUNDATION MODEL
310
LOCAL INFERENCE ENVIRONMENT

FIG. 5

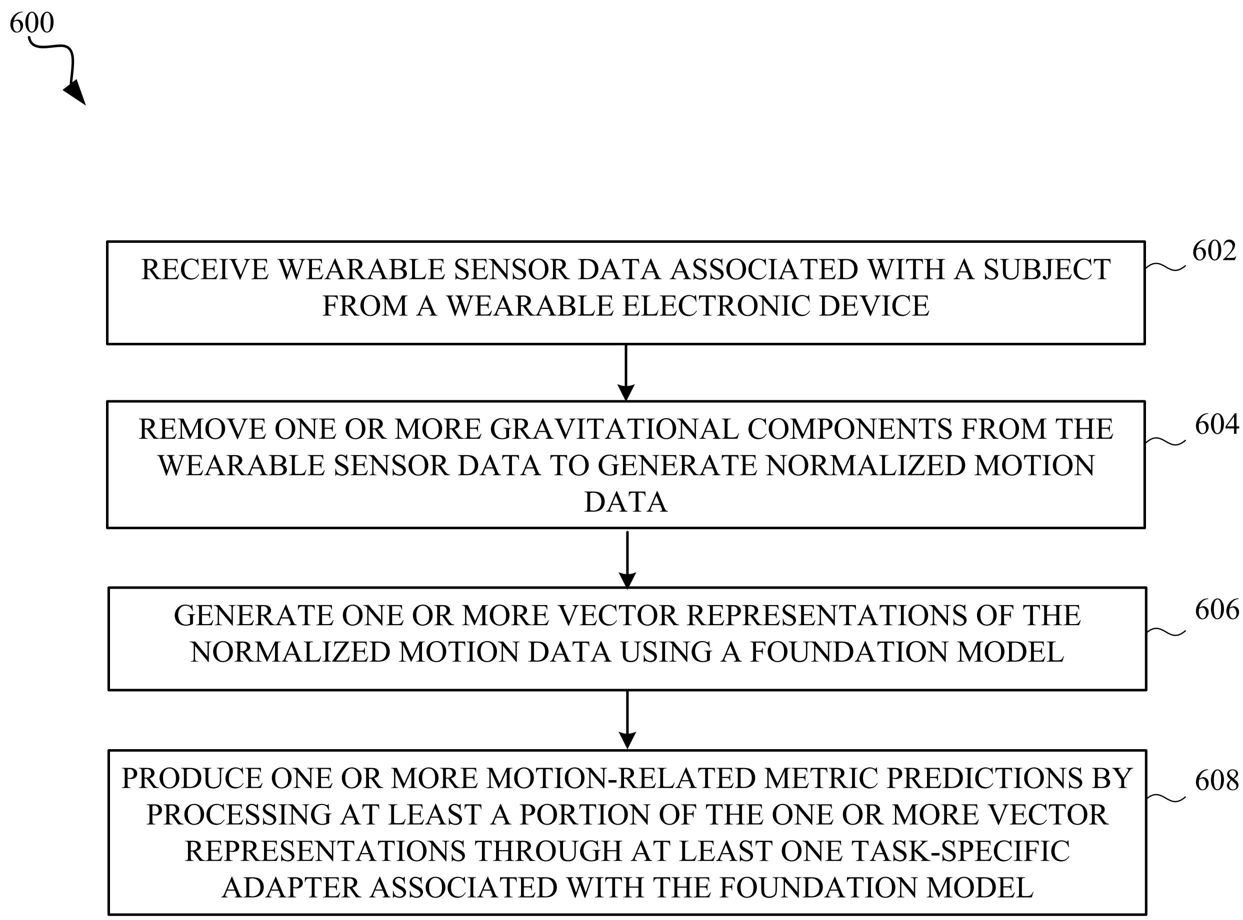
600
RECEIVE WEARABLE SENSOR DATA ASSOCIATED WITH A SUBJECT FROM A WEARABLE ELECTRONIC DEVICE
602
REMOVE ONE OR MORE GRAVITATIONAL COMPONENTS FROM THE WEARABLE SENSOR DATA TO GENERATE NORMALIZED MOTION DATA
604
GENERATE ONE OR MORE VECTOR REPRESENTATIONS OF THE NORMALIZED MOTION DATA USING A FOUNDATION MODEL
606
PRODUCE ONE OR MORE MOTION-RELATED METRIC PREDICTIONS BY PROCESSING AT LEAST A PORTION OF THE ONE OR MORE VECTOR REPRESENTATIONS THROUGH AT LEAST ONE TASK-SPECIFIC ADAPTER ASSOCIATED WITH THE FOUNDATION MODEL
608

FIG. 6

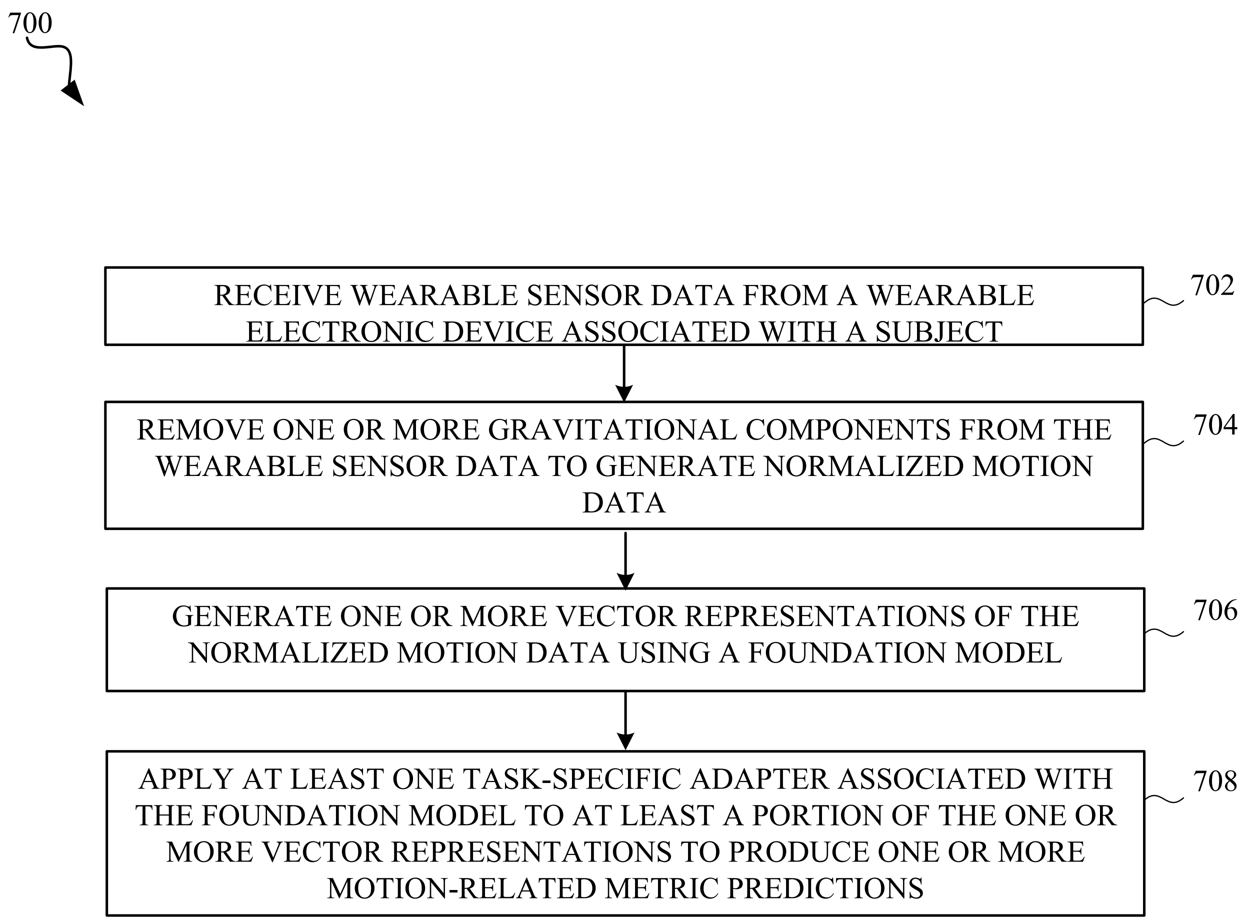
700
RECEIVE WEARABLE SENSOR DATA FROM A WEARABLE ELECTRONIC DEVICE ASSOCIATED WITH A SUBJECT
702
REMOVE ONE OR MORE GRAVITATIONAL COMPONENTS FROM THE WEARABLE SENSOR DATA TO GENERATE NORMALIZED MOTION DATA
704
GENERATE ONE OR MORE VECTOR REPRESENTATIONS OF THE NORMALIZED MOTION DATA USING A FOUNDATION MODEL
706
APPLY AT LEAST ONE TASK-SPECIFIC ADAPTER ASSOCIATED WITH THE FOUNDATION MODEL TO AT LEAST A PORTION OF THE ONE OR MORE VECTOR REPRESENTATIONS TO PRODUCE ONE OR MORE MOTION-RELATED METRIC PREDICTIONS
708

FIG. 7

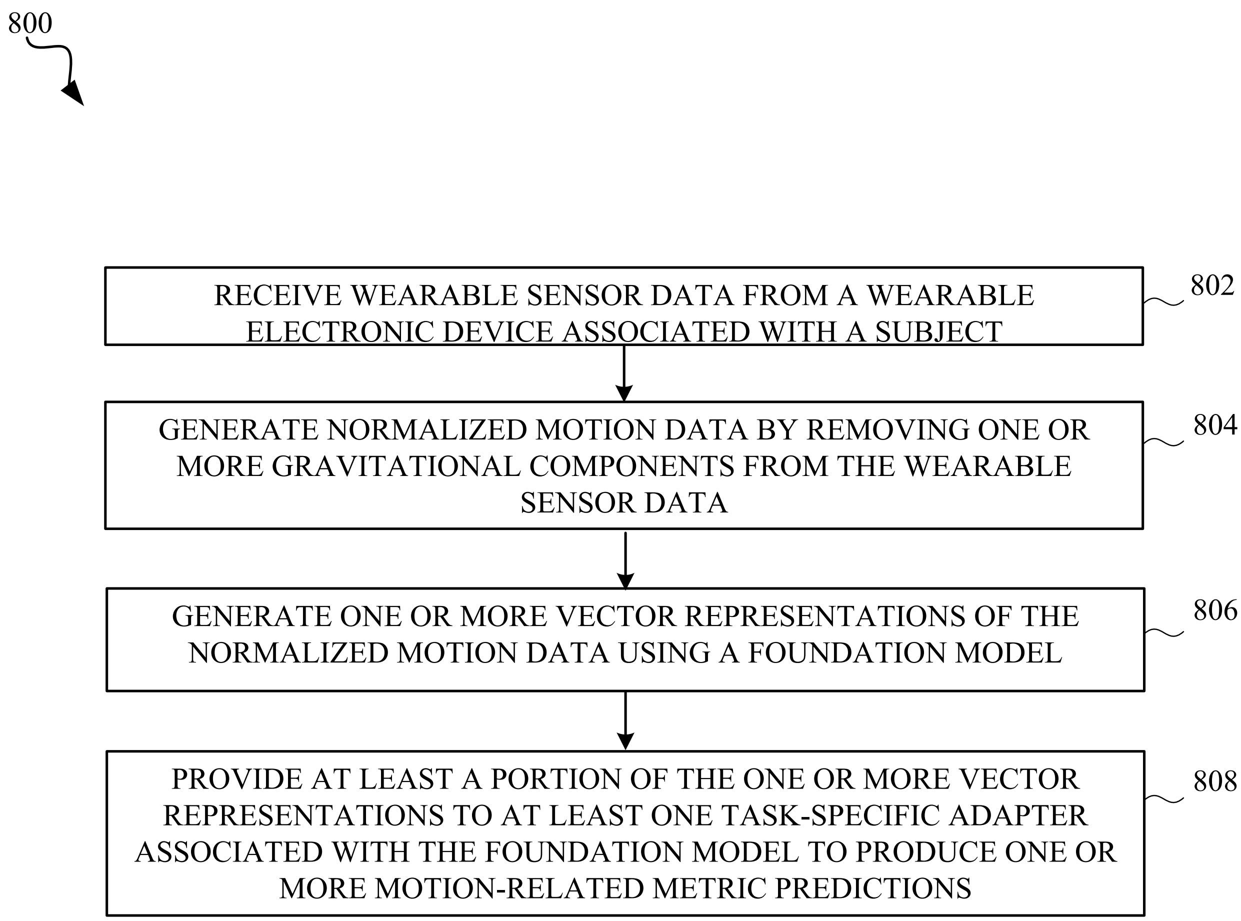
800
RECEIVE WEARABLE SENSOR DATA FROM A WEARABLE ELECTRONIC DEVICE ASSOCIATED WITH A SUBJECT
802
GENERATE NORMALIZED MOTION DATA BY REMOVING ONE OR MORE GRAVITATIONAL COMPONENTS FROM THE WEARABLE SENSOR DATA
804
GENERATE ONE OR MORE VECTOR REPRESENTATIONS OF THE NORMALIZED MOTION DATA USING A FOUNDATION MODEL
806
PROVIDE AT LEAST A PORTION OF THE ONE OR MORE VECTOR REPRESENTATIONS TO AT LEAST ONE TASK-SPECIFIC ADAPTER ASSOCIATED WITH THE FOUNDATION MODEL TO PRODUCE ONE OR MORE MOTION-RELATED METRIC PREDICTIONS
808

FIG. 8

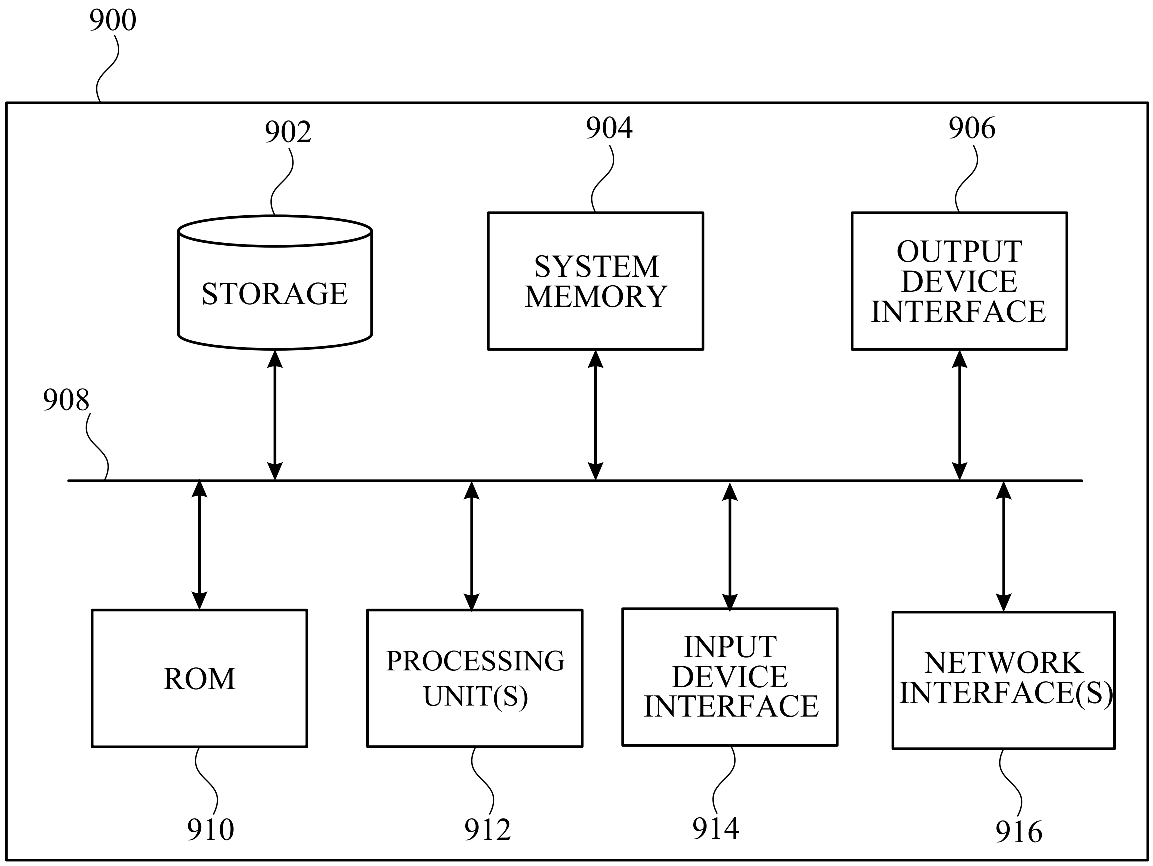
900
902
STORAGE
904
SYSTEM MEMORY
906
OUTPUT DEVICE INTERFACE
908
ROM
PROCESSING UNIT(S)
INPUT DEVICE INTERFACE
NETWORK INTERFACE(S)
910
912
914
916

FIG. 9