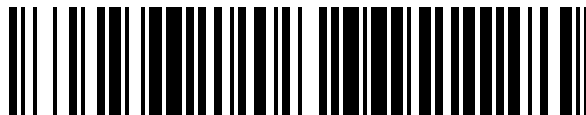


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **1 269 890**

21 Número de solicitud: 202130953

51 Int. Cl.:

**H04W 4/30** (2008.01)

**A61B 5/16** (2006.01)

12

SOLICITUD DE MODELO DE UTILIDAD

U

22 Fecha de presentación:

**21.11.2018**

43 Fecha de publicación de la solicitud:

**09.06.2021**

71 Solicitantes:

**UNIVERSIDAD CARLOS III DE MADRID (85.0%)  
SEI - Parque Científico; Avda. Gregorio Peces  
Barba, 1**

**28919 Leganés (Madrid) ES y  
UNIVERSIDAD POLITÉCNICA DE MADRID (15.0%)**

72 Inventor/es:

**SAN SEGUNDO MANUEL, Rosa;  
SAINZ DE BARANDA, Clara;  
BLANCO RUIZ, Marian;  
LARRABEITI LÓPEZ, David;  
URUEÑA PASCUAL, Manuel;  
ROBLEDO GARCÍA, Jose Carlos;  
PELÁEZ MORENO, Carmen;  
GALLARDO ANTOLÍN, Ascensión;  
MÍNGUEZ SÁNCHEZ, Alba;  
RIESGO ALCAIDE, Teresa;  
LANZA GUTIÉRREZ;  
MARINO ANDRÉS, Rodrigo;  
MIRANDA CALERO, Jose Ángel;  
FELIPE CANABAL, Manuel;  
PORTELA GARCÍA, Marta;  
PÉREZ GARCILÓPEZ, Isabel;  
GARCÍA SOUTO, Jose Antonio;  
LÓPEZ ONGIL, Celia;  
OLÍAS RUIZ, Emilio y  
GARCÍA VALDERAS,**

74 Agente/Representante:  
**UNGRÍA LÓPEZ, Javier**

54 Título: **SISTEMA PARA DETERMINAR UN ESTADO EMOCIONAL DE UN USUARIO**

ES 1 269 890 U

## DESCRIPCIÓN

### SISTEMA PARA DETERMINAR UN ESTADO EMOCIONAL DE UN USUARIO

#### 5 OBJETO DE LA INVENCION

La presente invención se refiere al campo técnico del reconocimiento de emociones a través del procesamiento multimodal de señales fisiológicas y de audio, y más concretamente a la monitorización automática y portátil del estado emocional de un usuario, con la posibilidad de comunicarlo a terceros o establecer, por ejemplo, medidas de seguridad como el envío  
10 de alarmas a una red de contactos o emergencias ante una situación de peligro.

#### ANTECEDENTES DE LA INVENCION

Actualmente, el reconocimiento automático de emociones es un campo en auge y pleno desarrollo que se beneficia de los avances en computación y algoritmos de aprendizaje  
15 automático. Esta tecnología proporciona una interacción hombre-máquina mucho más eficiente y mejora la actual lo que puede aprovecharse en una gran variedad de casos que van desde la detección de estrés, bloqueo o miedo en usuarios, hasta aplicaciones de neuro-marketing, detección de diferentes tipos de violencia, aplicaciones médicas, detección del estado del usuario en aplicaciones e-learning, o detección del estado afectivo para  
20 personas que padecen algún tipo de trastorno (p.ej., autismo).

El estado del arte comprende amplia literatura que relaciona las variaciones fisiológicas medidas en seres humanos con los cambios en sus estados emocionales. En estos trabajos, la detección emocional se lleva a cabo mapeando variables fisiológicas de individuos  
25 expuestos a estímulos externos (vídeos, audios o imágenes) que producen emociones conocidas. Las bases de datos y estudios disponibles, se refieren a soluciones no portables que incluyen centenares de métricas clasificadas, generalmente, haciendo uso de un espacio de clasificación bidimensional "*Arousal – Valence*" (AV), donde el nivel de "*Arousal*" está directamente relacionado con la activación emocional y el de "*Valence*" indica lo  
30 "positiva" o "negativa" que es dicha emoción. Adicionalmente pueden ser incluidas otras dimensiones a este espacio, como por ejemplo dominancia o familiaridad.

Las soluciones que abordan la combinación de variables y optan por un escenario multimodal, implican gran complejidad, alta carga computacional y sensores con peso y

tamaño considerable, cableado, baterías y otros elementos que inevitablemente requieren la inmovilidad del sujeto. En contraste, las alternativas portátiles siguen un enfoque más sencillo basado en variables de una misma naturaleza, lo que condiciona claramente sus aplicaciones.

5

Por otro lado, son muy comunes los dispositivos electrónicos que se encuentran integrados y camuflados en otro tipo de objetos como prendas de vestir. Suelen ser denominados “wearables”, de la expresión en inglés que se refiere al conjunto de dispositivos que comprende pulseras, anillos, gafas, chaquetas o colgantes entre otros, que permiten a un usuario transportar consigo cualquier tipo de dispositivo electrónico de una manera transparente a terceros e incluso a sí mismo, pero en cambio le permiten beneficiarse de ciertas funcionalidades mediante sencillas interacciones con el dispositivo.

Algunas soluciones del estado del arte, referidas a la detección de estados emocionales, recurren a la integración con pulseras u otras prendas de vestir “wearables” e incorporan sensores para detectar algunas emociones básicas por medio de algoritmos inteligentes, que pueden ser combinadas con sensores inerciales o acelerómetros para limpiar las señales fisiológicas y el ruido debido al movimiento del usuario. No obstante, la clasificación de las emociones se realiza exclusivamente en función de las señales fisiológicas, lo que proporciona una robustez limitada que resulta insuficiente para ciertas aplicaciones en las que una falsa alarma es inaceptable, como por ejemplo en casos en los que ciertas emociones hayan sido vinculadas a la solicitud de una ambulancia o cuerpos de seguridad.

Por tanto, se echa en falta en el estado del arte una solución para determinar estados emocionales de los usuarios, con una naturaleza multimodal que garantice una precisión adecuada, pero que se mantenga portátil, adaptable a cada usuario particular y que pueda ser integrada en prendas de vestir de una manera transparente, tanto para el propio usuario como para terceros.

## 30 DESCRIPCIÓN DE LA INVENCION

Con el fin de alcanzar los objetivos y evitar los inconvenientes mencionados anteriormente, la presente invención describe, en un primer aspecto un sistema para determinar un estado

emocional de un usuario, que comprende:

- 5 - un primer dispositivo portátil que dispone de unos medios sensores configurados para la adquisición de un conjunto de señales fisiológicas del usuario, donde el primer dispositivo portátil comprende un primer módulo procesador configurado para determinar, a partir del conjunto de señales fisiológicas, un primer estado emocional parcial con un primer nivel de dicho estado emocional; determinar si el primer estado emocional parcial coincide con un estado emocional objetivo establecido previamente y si el primer nivel del primer estado emocional parcial supera un primer umbral previamente establecido; y transmitir, en caso afirmativo, un mensaje de alarma a un  
10 dispositivo móvil de comunicación del usuario;
- un segundo dispositivo portátil configurado para adquirir una señal de audio al recibir una instrucción de activación desde un dispositivo móvil de comunicación del usuario;
- un segundo módulo procesador configurado para determinar, a partir de la señal de audio, un segundo estado emocional parcial con un segundo nivel (66) de dicho  
15 estado emocional, y para determinar si coincide con el estado emocional objetivo y si supera un segundo umbral previamente establecido; y
- un dispositivo móvil de comunicación que comprende:
  - 20 - un módulo de comunicación inalámbrico configurado para recibir el mensaje de alarma del primer dispositivo portátil y enviar la instrucción de activación al segundo dispositivo portátil; y
  - un módulo analizador configurado para determinar, en caso de que el segundo módulo analizador indique que el segundo nivel ha superado el segundo umbral, la presencia del estado emocional objetivo en el usuario, basado en un análisis de los niveles de los estados emocionales parciales por un algoritmo de  
25 aprendizaje automático.

El módulo analizador de la presente invención, en una de las realizaciones particulares, está además configurado para, una vez determinada la presencia de la emoción objetivo en el usuario, enviar automáticamente un mensaje con información del análisis del módulo  
30 analizador a un sistema remoto, a través de una red de telecomunicación, donde el sistema remoto se selecciona entre: una red de contactos del usuario, un servidor accesible por un servicio de emergencias, y un servidor privado accesible por un tercer usuario autorizado por el usuario.

Los medios sensores del primer dispositivo portátil comprenden, de acuerdo a una de las realizaciones de la invención: unos medios detectores de respuesta galvánica de la piel (GSR), configurados para obtener una señal con información de la conductividad de la piel del usuario; unos medios detectores de volumen sanguíneo del pulso (BVP) configurados para obtener una señal con información del pulso cardíaco del usuario; y unos medios detectores de temperatura configurados para obtener una señal con información de la temperatura de la piel (SKT) del usuario.

La presente invención contempla la posibilidad de que el módulo analizador sea una unidad computacional configurada para determinar el estado emocional del usuario de acuerdo a unos datos de entrenamiento proporcionado previamente.

Opcionalmente, la presente invención comprende un módulo controlador configurado para un entrenamiento inicial del sistema, donde el módulo controlador comprende una primera base de datos de señales fisiológicas cuantificadas según estímulos audiovisuales asociados previamente a estados emocionales específicos y una segunda base de datos de señales de audio asociadas previamente a los estados emocionales específicos.

Una realización particular de la presente invención contempla incorporar en el segundo dispositivo portátil un micrófono configurado para adquirir la señal de audio. Adicionalmente, se contempla la posibilidad de que el segundo módulo procesador incorpore un detector de actividad vocal configurado para determinar la presencia de silencios en la señal de audio adquirida y contabilizar dichos silencios.

Opcionalmente, al menos uno de los dispositivos portátiles de la presente invención comprende un pulsador configurado para, tras una pulsación del usuario, transmitir un mensaje de auxilio al dispositivo móvil de comunicación; y donde el dispositivo móvil de comunicación está además configurado para, en respuesta a la recepción del mensaje de auxilio, reenviar el mensaje de auxilio a un grupo de contactos del usuario previamente establecido.

En una de las realizaciones de la invención, se contempla que el sistema comprenda una pulsera, con una primera carcasa que aloja en su interior el primer dispositivo portátil; y un colgante, con una segunda carcasa que aloja en su interior el segundo dispositivo portátil;

donde el dispositivo móvil de comunicación es un teléfono móvil de tipo smartphone que integra el segundo módulo procesador. Alternativamente, otra realización de la invención contempla que el sistema comprenda una pulsera, con una primera carcasa que aloja en su interior el primer dispositivo portátil; un colgante, con una segunda carcasa que aloja en su interior el segundo dispositivo portátil y el segundo módulo procesador; donde el dispositivo móvil de comunicación es un teléfono móvil de tipo smartphone.

Un segundo aspecto de la invención se refiere a un método para determinar la presencia de un estado emocional en un usuario, que comprende los pasos de: adquirir, mediante unos medios sensores dispuestos en un primer dispositivo portátil un conjunto de señales fisiológicas; determinar, por un primer módulo procesador, un primer estado emocional parcial a partir del conjunto de señales fisiológicas, con un primer nivel de dicho primer estado emocional parcial; determinar si el primer estado emocional parcial coincide con un estado emocional objetivo establecido previamente y si el primer nivel del primer estado emocional parcial supera un primer umbral previamente establecido; en caso afirmativo, transmitir un mensaje de alarma a un dispositivo móvil de comunicación del usuario, mediante un módulo de comunicación inalámbrico; como resultado de la recepción del mensaje de alarma, adquirir una señal de audio, mediante un segundo dispositivo portátil en comunicación con el dispositivo móvil de comunicación; determinar, en un segundo módulo procesador un segundo estado emocional parcial, a partir de la señal de audio, con un segundo nivel de dicho segundo estado emocional parcial; determinar si el segundo estado emocional parcial coincide con el estado emocional objetivo establecido previamente y si el segundo nivel de dicho segundo estado emocional parcial supera un segundo umbral previamente establecido; en caso afirmativo, determinar, en un módulo analizador, la presencia del estado emocional objetivo del usuario, basado en un análisis de los niveles de los estados emocionales parciales por un algoritmo de aprendizaje automático.

Adicionalmente, en una de las realizaciones de la invención, se contempla la posibilidad de que determinar la presencia del estado emocional objetivo del usuario comprenda los pasos de: determinar un nivel emocional total basado en los niveles de los estados emocionales parciales; comparar el nivel emocional total determinado con un tercer umbral previamente establecido; y determinar que el estado emocional está presente en caso de que dicho nivel emocional total supere dicho tercer umbral.

De manera opcional, en una realización de la invención, una vez determinada la presencia de la emoción objetivo en el usuario, se envía automáticamente un mensaje con información del análisis del módulo analizador a un sistema remoto, a través de una red de telecomunicación (7), donde el sistema remoto se selecciona entre: una red de contactos del usuario, un servidor accesible por un servicio de emergencias, y un servidor privado accesible por un tercer usuario autorizado por el usuario.

Se contempla la posibilidad de que el conjunto de señales fisiológicas comprenda tres señales fisiológicas y, adquirir por los medios sensores el conjunto de señales fisiológicas, comprenda adquirir una señal con información de la conductividad de la piel (GSR) del usuario, adquirir una señal con información del pulso cardíaco del usuario (BVP), y adquirir una señal con información de la temperatura de la piel (SKT) del usuario.

En una de las realizaciones de la presente invención, determinar un estado emocional parcial con un nivel de dicho estado emocional parcial, comprende mapear un punto en un espacio de tres dimensiones que representa todos los estados emocionales, basado en unos valores numéricos de las tres variables de “Placer/Valencia” – “Excitación” – “Dominancia”, asignados para el conjunto de señales fisiológicas adquirido o para ciertas características de la señal de audio adquirida.

Adicionalmente, se contempla la posibilidad de añadir una etapa previa de entrenamiento que comprende: alimentar una primera base de datos con señales fisiológicas cuantificadas según estímulos audiovisuales asociados previamente a estados emocionales específicos; alimentar una segunda base de datos con señales de audio con unas características espectrales y/o prosódicas asociadas previamente a los estados emocionales específicos; registrar una desviación, respecto de la primera y segundas bases de datos, de las señales fisiológicas y las características espectrales y prosódicas de la señal de audio proporcionadas por el usuario; y adaptar el primer módulo procesador, el segundo módulo procesador y el módulo analizador a las desviaciones registradas para el usuario.

La invención propuesta se basa en tecnologías inalámbricas para ofrecer una solución distribuida en varios dispositivos conectados con el teléfono móvil del usuario, preferentemente una pulsera y un colgante. Así, el usuario puede transportar la invención sin que le resulte ninguna molestia y sin que sea percibida por terceros. Además, al estar

totalmente integrada en distintos wearables, resulta totalmente transparente para un eventual agresor, lo que reduce las posibilidades de que sean detectados y anulados por dicho agresor.

5 La presente invención monitoriza, registrar y recoge eventos derivados de la detección de emociones del usuario, como es el caso del pánico y estrés motivados por ejemplo, por un ataque sexual o violento. De forma posterior a la detección de dichos eventos, ventajosamente pueden desencadenarse de manera automática un conjunto de alarmas hacia una red de contactos “guardianes”, previamente configurada a través de una  
10 aplicación móvil, o hacia servicios de emergencias/seguridad.

De acuerdo a todo lo anterior, las características de la presente invención implican multitud de ventajas para el usuario. La incorporación de la señal de audio en el proceso de decisión, clasificación y determinación del estado emocional del usuario, proporciona un nuevo  
15 enfoque, puesto que, gracias a ella, y al algoritmo de aprendizaje que se aplica sobre la misma, se aporta un valor añadido al sistema propuesto, proporcionando una mayor robustez o precisión a la inferencia emocional y, además, identificando posibles sonidos externos al usuario (silencios, portazos, disparos, etc.). Por tanto, la presente invención proporciona al usuario la capacidad de reaccionar y avisar rápidamente ante posibles  
20 agresiones, por ejemplo de carácter sexual.

### **BREVE DESCRIPCIÓN DE LAS FIGURAS**

Para completar la descripción de la invención y con objeto de ayudar a una mejor comprensión de sus características, de acuerdo con un ejemplo preferente de realización de  
25 la misma, se acompaña un conjunto de dibujos en donde, con carácter ilustrativo y no limitativo, se han representado las siguientes figuras:

- La **figura 1** representa esquemáticamente una realización del sistema completo de la invención.  
30

- La **figura 2** representa una realización de unos de los dispositivos portátiles con forma de colgante.



- La **figura 3** representa una realización de unos de los dispositivos portátiles con forma de pulsera.

5 - La **figura 4** muestra en un diagrama de bloques la naturaleza multimodal de la presente invención.

- La **figura 5** muestra mediante un diagrama de bloques el tratamiento de las señales fisiológicas en una de las realizaciones de la invención.

10 - La **figura 6** muestra mediante un diagrama de bloques el tratamiento de la señal de audio en una de las realizaciones de la invención.

- La **figura 7** representa un espacio de tres dimensiones utilizado por la presente invención para repartir el espacio entre ocho cuadrantes emocionales.

15 - La **figura 8** representa en un diagrama de bloques la fase de entrenamiento y el método de configuración inicial personalizada del sistema de la presente invención.

20 - La **figura 9** representa esquemáticamente una realización del sistema completo de la invención.

### **DESCRIPCIÓN DE UNA REALIZACIÓN PREFERENTE DE LA INVENCION**

La presente invención divulga un método y un sistema distribuido para detectar el estado emocional de un usuario, lo que puede corresponder a situaciones de riesgo o agresiones  
25 inminentes o a otras emociones susceptibles de ser utilizadas para fines médicos, deportivos, etc. Utiliza para ello una efectiva integración multimodal de señales fisiológicas y físicas, preferentemente audio externo y voz (aunque también podrían usarse otras señales audiovisuales), mediante sensores integrables en una solución “wearable” portable y camuflable en prendas de vestir y/o complementos, que es capaz de alertar a un círculo de  
30 confianza del usuario o a las fuerzas de seguridad.

Una de las realizaciones de la invención, especialmente ventajosa en la detección de una emoción objetivo asociada a estados de pánico o bloqueos referidos a situaciones de violencia o agresión sexual, se encuentra representada en las **figuras 1 y 9**, donde el

sistema se compone de tres dispositivos principales: dos dispositivos portátiles camuflados en prendas de vestir o “wearables”, que en este caso son una pulsera **1** y un colgante **2**, y un dispositivo móvil de comunicación, que en este caso es un teléfono móvil **3** de tipo smartphone en el que se ejecuta una aplicación **4** diseñada específicamente.

5

La pulsera adquiere y monitoriza señales fisiológicas, las cuales son capturadas a través de sensores biométricos **5** en pequeños periodos de tiempo, y aplica algoritmos de aprendizaje automático a fin de disponer de un primer nivel de alerta en caso de detección positiva de la emoción objetivo, que en esta realización es la detección de una emoción de pánico o bloqueo ante una posible agresión. Dicha alerta se envía mediante un módulo de comunicación inalámbrico de corto alcance **6**, por ejemplo Bluetooth, al teléfono móvil del usuario, el cual evalúa el contenido y envía, a través de la red de comunicación inalámbrica, una orden al colgante para activar la siguiente capa de detección del sistema. El colgante comienza entonces a adquirir audio del entorno del usuario, lo comprime y lo envía al teléfono móvil, el cual aplica algoritmos de aprendizaje automático para detectar indicios de riesgo, como por ejemplo cierto nivel de estrés, utilizando dicho audio adquirido por el micrófono dispuesto en el colgante del usuario. Ambas informaciones, fisiológicas y físicas, se integran en el teléfono móvil para proporcionar un análisis más robusto y preciso. Finalmente, en caso de detectarse un positivo (un estado emocional que supera un umbral de riesgo preestablecido que determina que existe violencia o indicios de ésta), se avisa a una red de contactos **11** de confianza, previamente establecida por el usuario durante la configuración inicial de la aplicación software **4** o directamente a un servicio de emergencias **9**, por medio de una red de telecomunicación **7**. La red de telecomunicación utilizada puede ser cualquier red apta para telefonía móvil (GPRS/3G/4G) o estar basada directamente en una conexión WiFi a Internet, que envía notificaciones push **10** a la red de contactos elegidas a través de un servidor **8** dedicado.

Para un mejor funcionamiento de la invención, es recomendable personalizar para cada usuario los umbrales de riesgo y los niveles a considerar como positivos mediante un proceso de entrenamiento o adaptación previo a su puesta en marcha, ya que cada usuario puede tener una reacción corporal muy diferente ante una determinada emoción. Así, un sistema personalizado permitirá identificar las emociones de una forma más robusta que uno genérico.

La **figura 2** representa una realización de uno de los dispositivos portátiles en los que se distribuye la presente invención. En este caso, el dispositivo portátil está implementado en un colgante **2**, aunque en otras realizaciones también se contemplan wearables como pendientes, diademas, piercings o broches. El colgante comprende una carcasa que aloja en su interior los componentes electrónicos necesarios para la adquisición de una señal física, en este caso un micrófono **20** para captar una señal de audio, para las comunicaciones inalámbricas con el teléfono móvil, batería y microprocesador. La carcasa cuenta con una microperforación **21**, en su cara exterior, coincidente con el micrófono alojado en el interior de la carcasa para facilitar la recepción del audio. La cara exterior del colgante presenta un botón **22** de pánico accionable por el usuario manualmente, que envía inmediatamente, por medio siempre del teléfono móvil, un mensaje de auxilio a la red de contactos. El botón de pánico se encuentra camuflado en el diseño del colgante. En una de las realizaciones de la invención, la cara trasera del colgante dispone de un pequeño agujero que permite acceder a los componentes electrónicos del interior de la carcasa mediante un objeto alargado de tipo puntero o un alfiler. Este acceso se limita a un botón de reset para reiniciar el dispositivo.

En una de las realizaciones de la presente invención, el colgante comprende adicionalmente una cámara para la adquisición de la señal física. En este caso, además de adquirir una señal de audio como opción preferente, añade funcionalidades e información adicional al sistema adquiriendo imágenes y vídeo.

La **figura 3** representa una realización de otro de los dispositivos portátiles en los que se distribuye la presente invención. En este caso, el dispositivo portátil está implementado en una pulsera **1**, aunque en otras realizaciones también se contemplan wearables como brazaletes o tobilleras. La pulsera tiene una carcasa que aloja en su interior un microprocesador y los componentes electrónicos necesarios para la adquisición de las señales fisiológicas del usuario. En esta realización se incluye un primer sensor **31** detector de respuesta galvánica de la piel, que preferiblemente se dispone fuera de la carcasa para facilitar el contacto con la piel del usuario de un par de electrodos, dicho primer sensor está configurado para obtener una señal con información de la conductividad de la piel (GSR) del usuario; un segundo sensor **32** detector de volumen sanguíneo del pulso (BVP)

configurados para obtener una señal con información del pulso cardíaco del usuario; y un tercer sensor **33** detector de temperatura configurado para obtener una señal con información de la temperatura de la piel (SKT) del usuario. Los sensores se disponen en la cara interna de la pulsera para que, al ser colocada en la muñeca del usuario, se mantengan en contacto con su piel.

El interior de la carcasa también aloja un módulo de comunicaciones inalámbrica de corto alcance, preferiblemente Bluetooth, una batería y microprocesador **36**. En una de las realizaciones, la carcasa además cuenta con una perforación **34**, en su cara exterior, coincidente con un botón de reset alojado en el interior de la carcasa y accesible con un objeto alargado de tipo puntero o un alfiler para reiniciar el dispositivo. La cara exterior de la pulsera presenta un botón **35** de pánico accionable por el usuario manualmente, que envía inmediatamente, por medio siempre del teléfono móvil, un mensaje de auxilio a la red de contactos o el destinatario que se haya configurado previamente.

Con referencia a la **figura 4**, el sistema multimodal de reconocimiento de emociones, en una realización preferente de la presente invención, se alimenta de las siguientes variables fisiológicas: conductividad de la piel **40** (“*Galvanic Skin Response*”, GSR), volumen sanguíneo del pulso **41** (“*Blood Volume Pulse*”, BVP), y temperatura **42** (“*Skin Temperature*”, SKT). Y por otro lado, se alimenta de una variable física, que en este caso es el audio **43** e incluye la voz del usuario junto con el sonido del entorno.

Una vez adquiridas las señales, se procede a la extracción **44** de información de manera individual de todas y cada una de las variables. Posteriormente, todo el conjunto de datos o información sensible es fusionado **45** a fin de determinar **46** una única categoría de estado emocional para el usuario. Aunque el dispositivo de comunicación está configurado para recibir tanto la información procesada por el primer dispositivo portátil como por el segundo dispositivo portátil, el método de la presente invención se lleva a cabo en dos etapas, donde la primera etapa (procesado de las señales fisiológicas) actúa de llave para la segunda etapa (procesado del audio), de manera que sin una primera alarma, detectada exclusivamente a partir de las señales fisiológicas, no se establecen el resto de comunicaciones ni de procesos en el segundo dispositivo portátil y, sólo una vez que se ha

producido una segunda alarma tras el procesado del audio, se activa el procesado multimodal en un módulo analizador **47** del dispositivo de comunicación, que fusiona **45** los datos de ambas clasificaciones, determinando **46** la emoción que está sintiendo el usuario.

- 5 Por tanto, existen dos bloques principales y diferenciados, el primero encargado del tratamiento de las señales fisiológicas y el segundo del audio. Dichos bloques son totalmente independientes, tanto en funcionalidad como en operación, aunque la conjunción de ambos proporciona una fusión multimodal que beneficia de manera notable a la inferencia emocional.

10

La **figura 5** representa el primero de dichos bloques, concretamente el bloque encargado del tratamiento de señales fisiológicas, donde un módulo procesador **50** que comprende un microprocesador integrado en la pulsera se encarga de todo el procesado. Las señales fisiológicas obtenidas por los sensores biométricos de la pulsera son procesadas en varias  
15 fases, en primer lugar, las señales en bruto **51** sufren una eliminación del ruido **52**. A continuación, se realiza un proceso de normalización **53** de las señales y extracción **54** de características, como por ejemplo la media, desviación típica, media de los valores absolutos de la primera diferencia, media de los valores absolutos de la primera diferencia de la señal normalizada, media de los valores absolutos de la segunda diferencia, media de  
20 la primera diferencia de la señal suavizada, etc. Después de extraer las características de cada una de las señales fisiológicas por separado, se procede a fusionar **55** las características para finalmente clasificar **56** los resultados y obtener una señal de salida con un nivel emocional **57**.

- 25 El clasificador **56** aplica un algoritmo clásico de aprendizaje automático de bajo coste computacional para clasificar la emoción percibida por el sistema, por ejemplo siguiendo el método de los K-vecinos más cercanos, o abreviado KNN (del inglés "*K-Nearest-Neighbors*"). Este algoritmo necesita de unos datos de entrenamiento o puntos del espacio, los cuales se han obtenido previamente durante la configuración del sistema, la cual se  
30 desarrolla de manera offline y se detalla más adelante.

El nivel emocional **57** determinado a la salida del módulo procesador **50**, presenta un nivel de confianza que indica, en tanto por ciento, la probabilidad de que la combinación de características extraída de las señales fisiológicas se corresponda con una emoción en concreto. El nivel de confianza de cada emoción es, por tanto, la métrica que cuantifica la presencia de dicha emoción en el usuario.

Dado que el estado emocional del usuario se representa en un continuo en el que las emociones se solapan, varias emociones pueden darse a la vez con diferentes niveles de confianza. En el caso de definir una emoción objetivo particularmente interesante para una aplicación concreta, como es el caso de esta realización preferente para los estados de pánico, donde la emoción objetivo que se correspondería con el cuadrante -p+a-d de la **figura 7**, etiquetado como “miedo”, la señal de salida con un nivel emocional **59** comprende información de si el nivel de confianza de la emoción objetivo es superior o no a un umbral determinado previamente. El umbral se sitúa en el valor a partir del cual se considera que la emoción objetivo es predominante frente al resto de emociones. Este umbral de detección es el punto a partir del cual la emoción de interés es predominante (en tiempo), con respecto a las restantes emociones detectadas en el mismo periodo.

En caso de superar el umbral establecido para el nivel de confianza de la emoción objetivo, la pulsera establece una comunicación con el teléfono del usuario, para transmitir el nivel de emoción objetivo detectado en un primer mensaje de alarma, el cual provoca que el teléfono del usuario inmediatamente envíe un mensaje de activación al colgante para que comience a grabar audio.

La **figura 6** representa el segundo de dichos bloques para llevar a cabo la extracción de la categoría o tipo de emoción, concretamente el bloque encargado del tratamiento de la señal física. Las señales físicas, en este caso el audio, comienzan a adquirirse tras recibir en el colgante una orden enviada desde el teléfono móvil, el cual envía dicha orden solo tras recibir una indicación, primer mensaje de alarma, desde la pulsera de que el análisis de las señales fisiológicas ha clasificado la emoción objetivo con una confianza superior al umbral predeterminado. En ese caso, el micrófono del colgante es activado y comienza a grabar una señal de audio **60**. Esta señal de audio es procesada localmente por el microprocesador del colgante para comprimirla y transmitirla **61** al teléfono móvil por medio de la red de

comunicación inalámbrica Bluetooth. Una vez recibida en el módulo procesador 62, que en esta realización está integrado en el teléfono móvil, la señal comprimida se descomprime **63** y se procede a la extracción **64** de características. Después de extraer las características de la señal de audio, se procede a clasificarla **65** y obtener una señal de salida con una clasificación de la emoción objetivo y un nivel de confianza **66** obtenidos independientemente del análisis de la señales fisiológicas.

En una realización alternativa, el módulo procesador **62** está integrado en el microprocesador del segundo dispositivo portátil, el colgante por ejemplo, de manera que la señal de salida con una clasificación de la emoción objetivo y un nivel de confianza **66** es obtenida en el propio colgante antes de transmitir nada al teléfono móvil. Una vez se determina la presencia de la emoción objetivo es cuando se activa la transmisión al teléfono móvil para que su módulo analizador **47** active el procesado multimodal explicado anteriormente, que fusiona **45** los datos de ambas clasificaciones, determinando **46** la presencia de la emoción objetivo en el usuario, en función del análisis mediante un algoritmo de aprendizaje automático de las señales de salida de los bloques de las **figuras 5 y 6**.

Para la clasificación y cálculo del nivel de confianza de una emoción, el módulo procesador extrae diversas características espectrales y prosódicas de la señal de audio que, posteriormente serán clasificadas **65** aplicando un algoritmo clásico de aprendizaje automático de bajo coste computacional, para clasificar y calcular el nivel de confianza **66** de la emoción percibida por el sistema y confirmar o rechazar así, la emoción objetivo detectada por el primer dispositivo portátil. En función de si la emoción objetivo detectada por este segundo dispositivo, el colgante, supera un cierto umbral de confianza preestablecido, que no tiene que ser el mismo umbral establecido para el primer dispositivo, se procede a enviar el mensaje de auxilio a la red de contactos configurada por el usuario, o de acuerdo a otras realizaciones, a un servicio de emergencias, un centro médico o cualquier otro agente que sea considerado apropiado para actuar ante la detección de una cierta emoción.

En una realización, previo a la propia clasificación, además se utiliza un detector de actividad de voz ("*Voice Activity Detection*", VAD) para eliminar y contabilizar posibles

silencios dentro del audio grabado. De esta forma, no solamente se puede llevar a cabo la detección de la voz del usuario, sino que también se pueden detectar posibles sonidos del entorno relevantes (silencios, portazos, disparos, etc.).

5 Las salidas de los dos bloques representados en las figuras 5 y 6, que extraen parcialmente la categoría o tipo de emoción del usuario al procesar por separado la combinación de señales fisiológicas y la señal de audio respectivamente, se fusionan para determinar cuál es el estado emocional del usuario dentro de unos cuadrantes emocionales definidos previamente y, específicamente, para determinar si la emoción objetivo está presente o no  
 10 en el usuario. La **figura 7** muestra un espacio de tres dimensiones utilizado por la presente invención para repartir el espacio entre ocho cuadrantes emocionales. Esta distribución se basa en un modelo “*PAD-Space*” con tres ejes de coordenadas que representan todas los estados emocionales en función de los valores numéricos asignados para las variables de Placer/Valencia – Excitación - Dominancia (“*Pleasure/Valence – Arousal – Dominance*”, en  
 15 *inglés*), donde cada uno de dichos estados emocionales se representa en un vértice del cubo representado en dicho espacio tridimensional. Así, se asigna un vértice del cubo a cada uno de los siguientes ocho estados emocionales: alegría **71** (+p+a+d), gratitud **72** (+p+a-d), sumisión **73** (+p-a-d), angustia **74** (-p-a-d), alivio **75** (+p-a+d), desprecio **76** (-p-a+d), miedo **77** (-p+a-d), e ira **78** (-p+a+d). Por ejemplo, la variable de “Placer/Valencia” mide lo agradable que el usuario percibe un cierto estímulo. Así que, la “ira” **78** o el “miedo”  
 20 **77**, al ser emociones catalogadas como no placenteras, se ubican en el extremo negativo (-p). En contraste, la “alegría” **71**, al ser una emoción catalogada como placentera, se sitúa en el extremo del placer (+p).

25 La situación emocional del usuario queda por tanto representada en uno de los ocho cuadrantes emocionales definidos por estas tres dimensiones, en función de los resultados obtenidos para cada una de las tres variables, Placer/Valencia – Excitación – Dominancia, al analizar, por un lado en el clasificador **58**, la combinación de las características extraídas de todas las señales fisiológicas y, por otro lado, en el módulo clasificador **64**, las  
 30 características extraídas de la señal de audio. Finalmente, como se mostraba en la figura 4, la combinación **45** de la información proporcionada por las señales fisiológicas y la señal de audio resulta en una única caracterización final **46** del estado emocional del usuario, que comprende la emoción objetivo y un nivel de confianza.



Una vez la presente invención ha detectado automáticamente el cuadrante emocional en el que se sitúa el usuario, puede configurarse para enviar esta información a una localización/usuario remoto por medio de una red de telecomunicación inalámbrica, donde se tomarán las acciones preventivas o resolutivas necesarias que se consideren en cada una de las aplicaciones de la invención. La información puede incluir también una información de geolocalización del usuario, que se envía automáticamente al determinarse que el nivel de la emoción objetivo ha superado un umbral preestablecido.

La **figura 8** representa la fase de entrenamiento y método de configuración inicial personalizada del sistema de la presente invención, la cual proporciona al algoritmo clásico de aprendizaje automático del módulo analizador los vectores de características y las etiquetas necesarias para realizar el entrenamiento. Para ello, se dispone un módulo controlador con algoritmos de aprendizaje automático y acceso a dos bases de datos independientes, una de señales fisiológicas **81** y otra de señales de audio **82**. Dicha configuración inicial se divide en dos procesos diferentes en base a las señales a capturar, que comparten el acondicionamiento de señal y extracción de características de las **figuras 5 y 6**.

En el caso de las señales fisiológicas, el usuario es estimulado mediante contenidos audiovisuales, previamente etiquetados con un cuadrante emocional específico. Las variaciones fisiológicas producidas son registradas y almacenadas **83** en la base de datos de señales fisiológicas **81** y, al final del proceso, el módulo controlador obtiene un modelo predictivo **84** entrenado para el usuario particular. De esta manera, la combinación de las características seleccionadas de las tres variables fisiológicas utilizadas en esta realización preferente, que obviamente podrían ser otras en realizaciones alternativas de la invención, queda caracterizada numéricamente y la influencia de sus variaciones se puede trasladar directamente a cada uno de los tres ejes del espacio PAD utilizado para representar todos los estados emocionales, que pueden ser mapeados en función de los valores numéricos asignados para las variables de Placer/Valencia – Excitación – Dominancia.

El caso de las señales de audio es análogo al descrito para las señales fisiológicas, con la diferencia de que el proceso de entrenamiento se realiza con grabaciones de voz locutadas por el propio usuario, de textos previamente etiquetados con un cuadrante emocional

específico. Así, las variaciones en las características espectrales y prosódicas de la voz del usuario/usuario, son registradas y almacenadas **85** en la base de datos de señales de audio **82**. Al final del proceso, el módulo controlador obtiene un modelo predictivo **86** entrenado para el usuario particular. De esta manera, la combinación de las características seleccionadas de la señal de audio queda caracterizada numéricamente y la influencia de sus variaciones puede trasladarse directamente a cada uno de los tres ejes del espacio PAD, utilizado para representar todos los estados emocionales, que pueden ser mapeados en función de los valores numéricos asignados para las variables de Placer/Valencia – Excitación – Dominancia.

10

La presente invención no debe verse limitada a la forma de realización aquí descrita. Otras configuraciones pueden ser realizadas por los expertos en la materia a la vista de la presente descripción. En consecuencia, el ámbito de la invención queda definido por las siguientes reivindicaciones.

15

## REIVINDICACIONES

1. Sistema para determinar un estado emocional de un usuario, caracterizado por que comprende:

5 - un primer dispositivo portátil (1) que dispone de unos medios sensores (5) configurados para la adquisición de un conjunto de señales fisiológicas (51) del usuario, donde el primer dispositivo portátil comprende:

un primer módulo procesador (50) configurado para:

10 determinar, a partir del conjunto de señales fisiológicas, un primer estado emocional parcial con un primer nivel (57) de dicho estado emocional;

determinar si el primer estado emocional parcial coincide con un estado emocional objetivo establecido previamente a través de un algoritmo de aprendizaje automático y dinámico y si el primer nivel del primer estado emocional parcial supera un primer umbral previamente establecido; y

15 transmitir, en caso afirmativo, un mensaje de alarma a un dispositivo móvil de comunicación del usuario junto con las señales fisiológicas;

- un segundo dispositivo portátil (2) configurado para adquirir una señal de audio (60) al recibir una instrucción de activación desde un dispositivo móvil de comunicación del usuario;

20 - un segundo módulo procesador (62) configurado para determinar, a partir de la señal de audio, un segundo estado emocional parcial con un segundo nivel (66) de dicho estado emocional, y para determinar si coincide con el estado emocional objetivo y si supera un segundo umbral previamente establecido; y

- un dispositivo móvil de comunicación (3) que comprende:

25 - un módulo de comunicación inalámbrico (6) configurado para recibir el mensaje de alarma del primer dispositivo portátil y enviar la instrucción de activación al segundo dispositivo portátil; y

30 - un módulo analizador (47) configurado para determinar, en caso de que el segundo módulo analizador indique que el segundo nivel ha superado el segundo umbral, la presencia del estado emocional objetivo en el usuario, basado en un análisis de los niveles de los estados emocionales parciales (57, 66) por el algoritmo de aprendizaje automático y dinámico.

2. Sistema de acuerdo a la reivindicación 1 donde el módulo analizador está además  
35 configurado para, una vez determinada la presencia de la emoción objetivo en el usuario,

enviar automáticamente un mensaje con información del análisis del módulo analizador a un sistema remoto, a través de una red de telecomunicación (7), donde el sistema remoto se selecciona entre: una red de contactos del usuario, un servidor accesible por un servicio de emergencias, y un servidor privado accesible por un tercer usuario autorizado por el usuario.

5

**3.** Sistema de acuerdo a cualquiera de las reivindicaciones anteriores donde los medios sensores del primer dispositivo portátil comprenden:

- unos medios detectores (31) de respuesta galvánica de la piel, configurados para obtener una señal con información de la conductividad de la piel (GSR) del usuario;

10 - unos medios detectores (32) de volumen sanguíneo del pulso (BVP) configurados para obtener una señal con información del pulso cardíaco del usuario;

- unos medios detectores (33) de temperatura configurados para obtener una señal con información de la temperatura de la piel (SKT) del usuario.

15 **4.** Sistema de acuerdo a cualquiera de las reivindicaciones anteriores donde el módulo analizador es una unidad computacional configurada para determinar el estado emocional del usuario de acuerdo a unos datos de entrenamiento proporcionado previamente.

20 **5.** Sistema de acuerdo a cualquiera de las reivindicaciones anteriores que además comprende un módulo controlador configurado para un entrenamiento inicial del sistema, donde el módulo controlador comprende una primera base de datos (81) de señales fisiológicas cuantificadas según estímulos audiovisuales asociados previamente a estados emocionales específicos y una segunda base de datos (82) de señales de audio asociadas previamente a los estados emocionales específicos.

25

**6.** Sistema de acuerdo a cualquiera de las reivindicaciones anteriores donde el segundo dispositivo portátil comprende además un micrófono (20) configurado para adquirir la señal de audio.

7. Sistema de acuerdo a la reivindicación 6, donde el segundo módulo procesador además comprende un detector de actividad vocal configurado para determinar la presencia de silencios en la señal de audio adquirida y contabilizar dichos silencios.

5 **8.** Sistema de acuerdo a cualquiera de las reivindicaciones anteriores donde, al menos uno de los dispositivos portátiles comprende un pulsador (22) configurado para, tras una pulsación del usuario, transmitir un mensaje de auxilio al dispositivo móvil de comunicación; y donde el dispositivo móvil de comunicación está además configurado para, en respuesta a la recepción del mensaje de auxilio, reenviar el mensaje de auxilio a un grupo de contactos  
10 del usuario previamente establecido.

**9.** Sistema de acuerdo a cualquiera de las reivindicaciones anteriores que además comprende:

- 15 - una pulsera con una primera carcasa que aloja en su interior el primer dispositivo portátil;
- un colgante con una segunda carcasa que aloja en su interior el segundo dispositivo portátil;

donde el dispositivo móvil de comunicación es un teléfono móvil de tipo smartphone que integra el segundo módulo procesador.

20

**10.** Sistema de acuerdo a cualquiera de las reivindicaciones 1-8 que además comprende:

- una pulsera con una primera carcasa que aloja en su interior el primer dispositivo portátil;
- 25 - un colgante con una segunda carcasa que aloja en su interior el segundo dispositivo portátil y el segundo módulo procesador;

donde el dispositivo móvil de comunicación es un teléfono móvil de tipo smartphone.

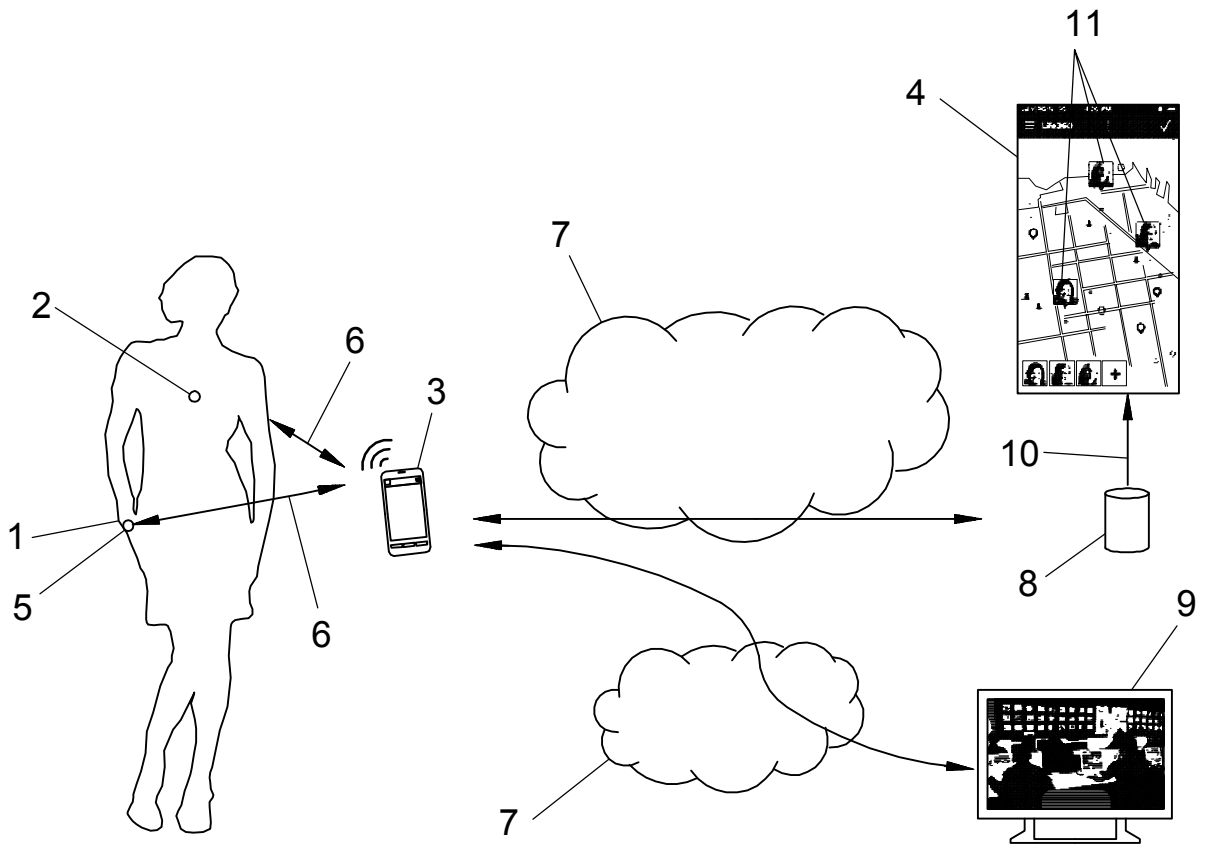


FIG. 1

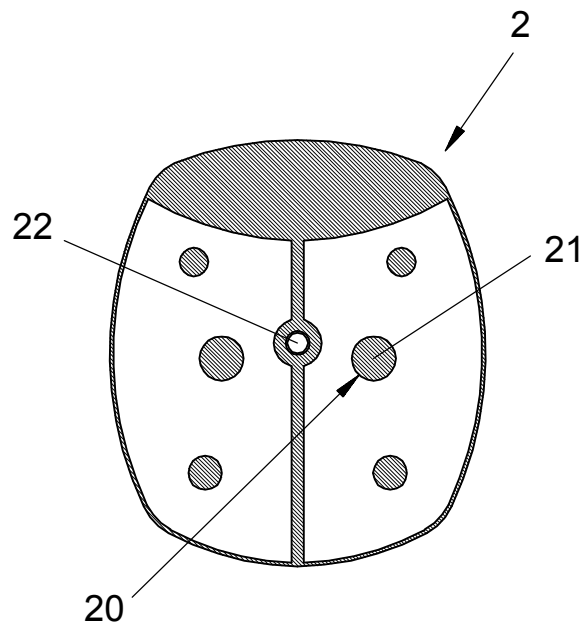


FIG. 2

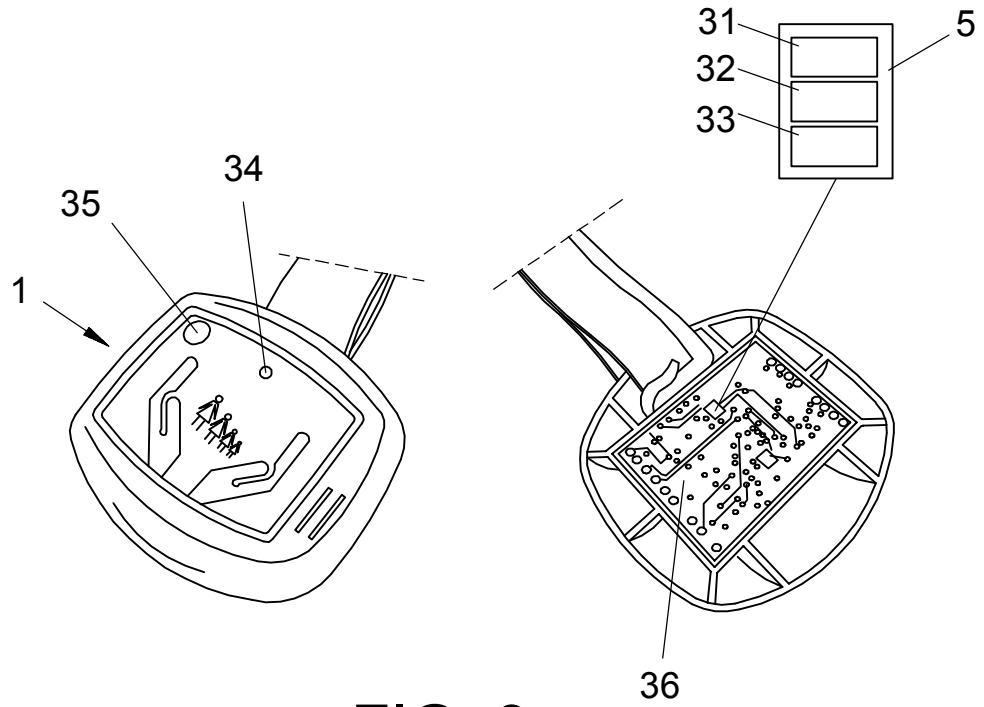


FIG. 3

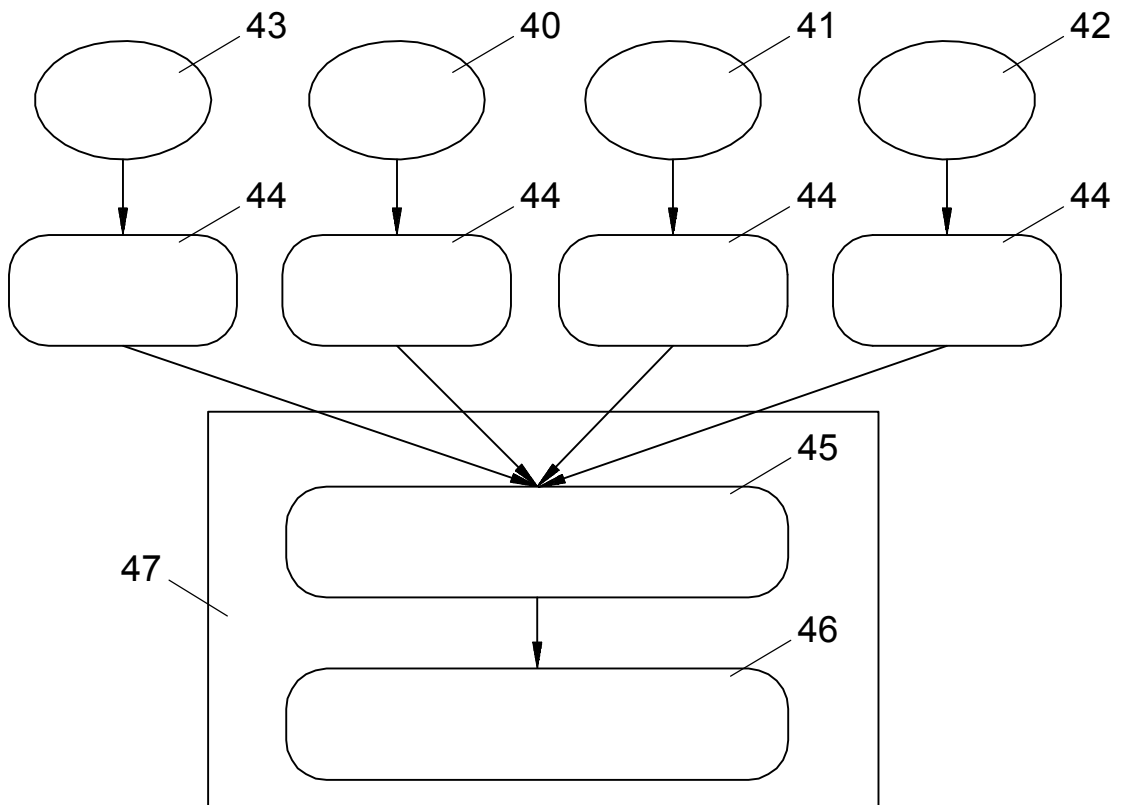


FIG. 4

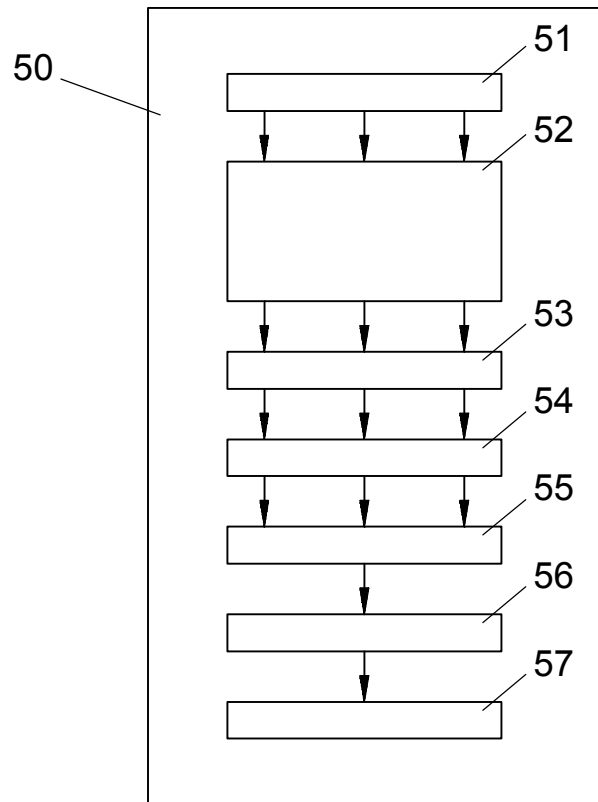


FIG. 5

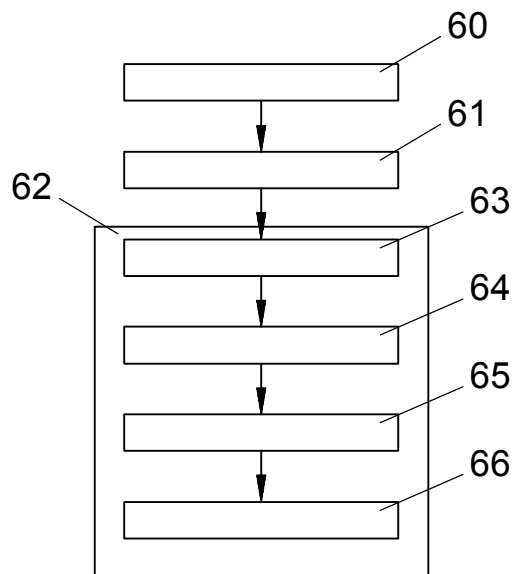


FIG. 6



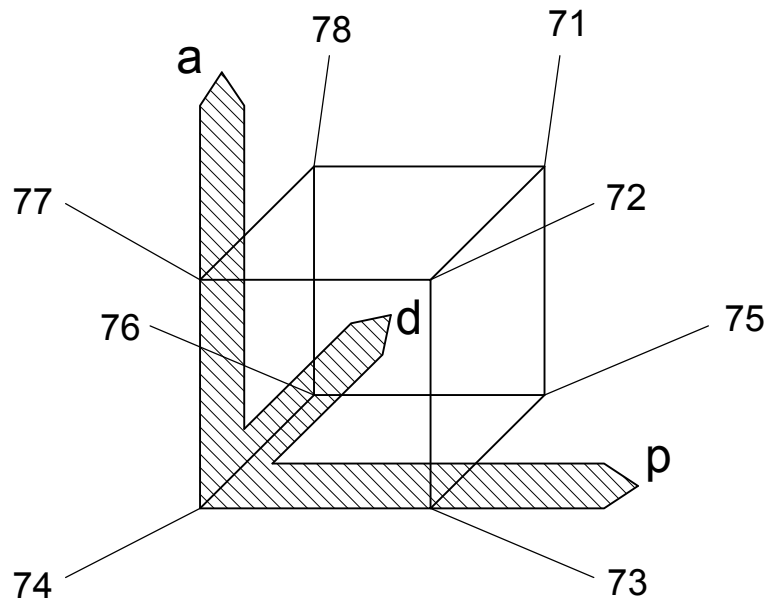


FIG. 7

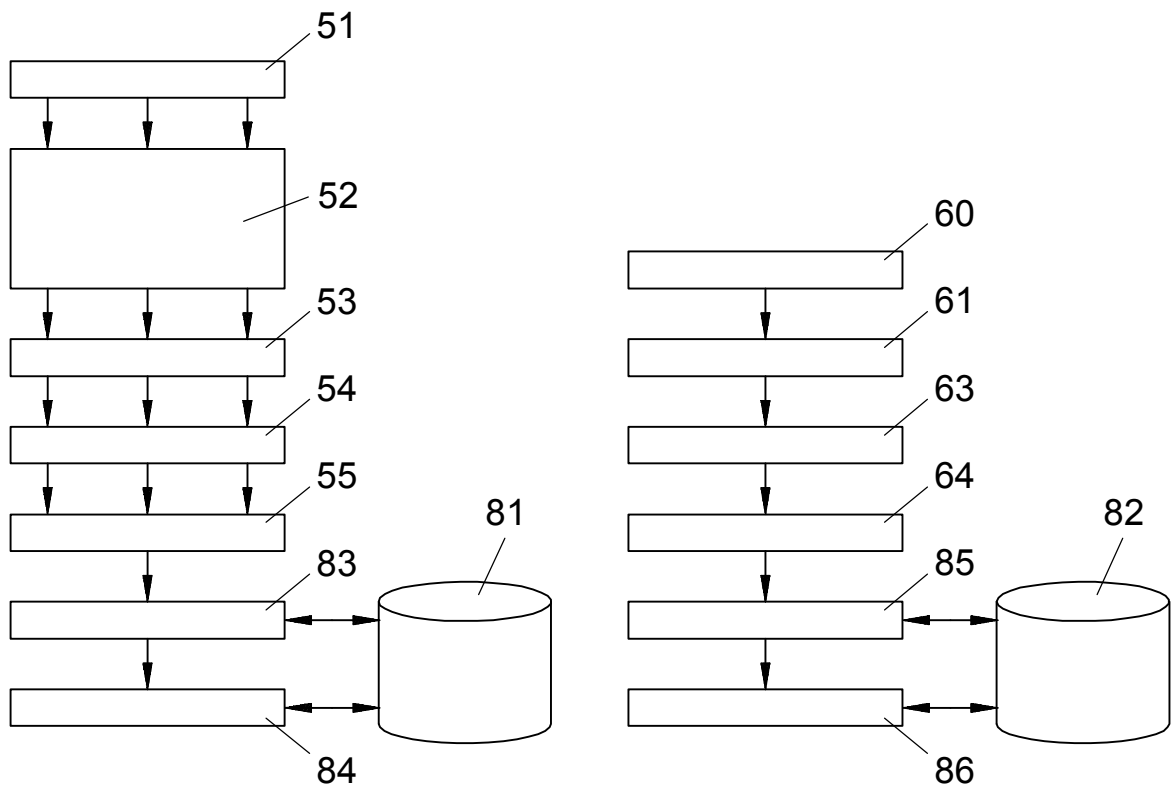


FIG. 8

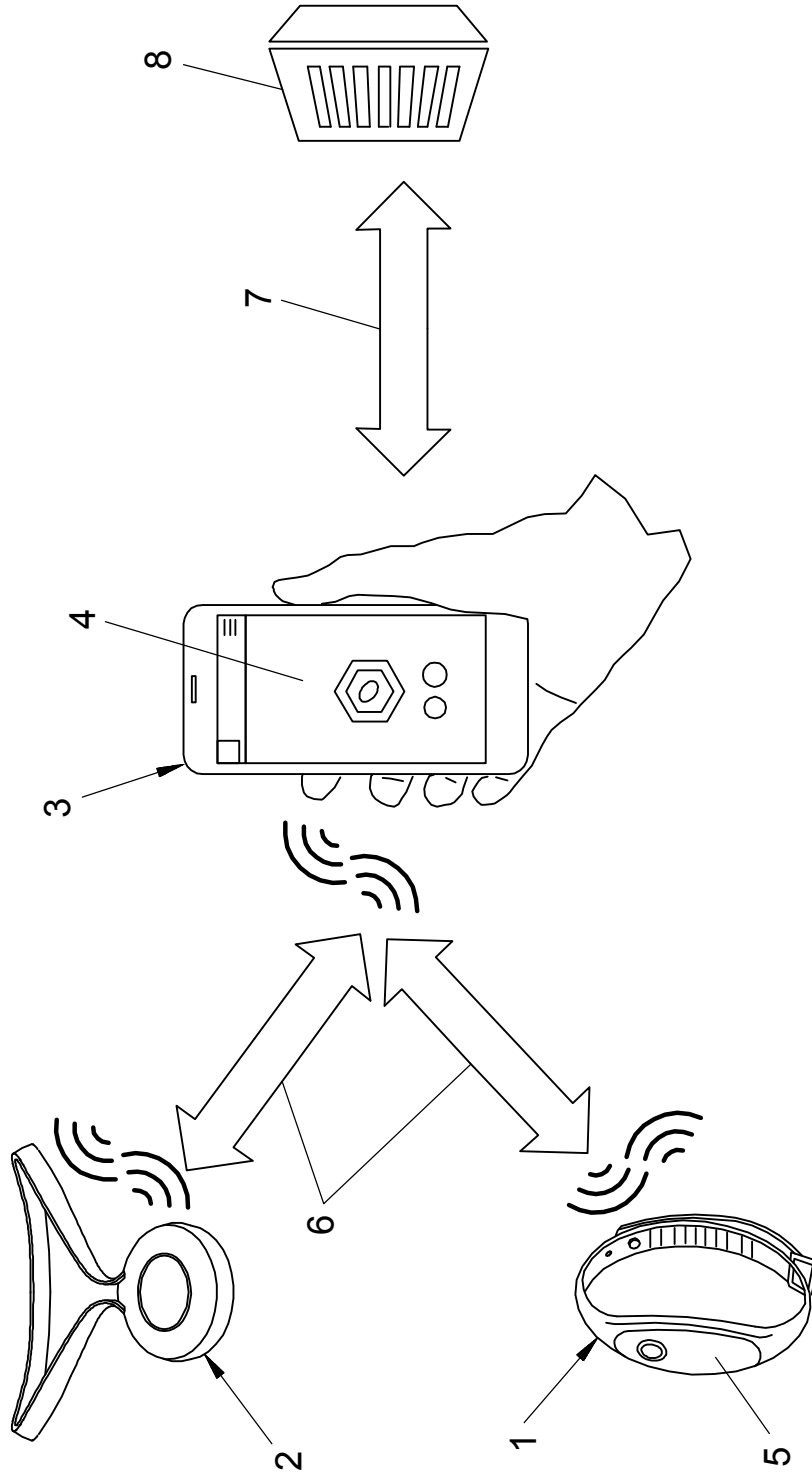


FIG. 9