MEJORAS EN EL ALGORITMO DE POSICIONAMIENTO EN INTERIORES MEDIANTE SENSORES RGB-D Y REDES WIFI

Jaime Duque Domingo, jaimeduque@amenofis.com
Carlos Cerrada
UNED, ETSI Informática,
Departamento de Ingeniería de Software y Sistemas Informáticos,
C/Juan del Rosal, 16. 28040 Madrid, Spain

Enrique Valero
Heriot-Watt University, School of Energy, Geoscience, Infrastructure and Society,
Edinburgh EH14 4AS, United Kingdom

Resumen

En este trabajo se presenta una mejora a un algoritmo de posicionamiento en interiores previamente desarrollado por los autores. Dicho algoritmo obtiene la posición de diversas personas que se mueven libremente en un entorno habitado compuesto de distintas habitaciones. El algoritmo tiene en cuenta tanto el posicionamiento mediante visión artificial como el posicionamiento mediante redes WiFi, combinando la información proveniente de ambos sistemas para aumentar la precisión. La mejora propuesta en este artículo hace referencia a una limitación del algoritmo en cuanto a que el número de personas detectadas por los sensores RGB-D tiene que ser igual al número de personas detectadas mediante WiFi. La mejora permite que dicho número no tenga que coincidir, lo cual puede ser habitual en escenarios donde un usuario no sea detectado por uno de los dos métodos. Además, también se simplifica el método de colocación de sensores RGB-D en el escenario.

Palabras clave: posicionamiento, IPS, WPS, RGB-D, Kinect, WiFi

1 INTRODUCCIÓN

En este trabajo se presenta una mejora al algoritmo de posicionamiento en interiores (IPS, del inglés *Indoor Positioning System*) mediante sensores RGB-D y redes WiFi presentado por los autores en [7]. La mejora presentada hace referencia tanto al método de colocación de los sensores en el sistema universal de coordenadas (UCS, del inglés *Universal Coordinate System*), como a una limitación del algoritmo en cuanto al número de personas en seguimiento.

1.1 Sistemas de posicionamiento combinados

Los sistemas combinados ofrecen resultados más precisos, pero normalmente requieren dispositivos especiales. De acuerdo a numerosos autores, la combinación de múltiples tecnologías mejora la eficiencia de los *WiFi Positioning System* (WPS) tradicionales [4, 9, 10]. Los sistemas híbridos combinan WPS con otras tecnologías, como GSM [11], bluetooth [15], o RFID [13, 16]. El trabajo [8] estudia la combinación de WiFi y *Sistemas de Navegación Inercial* (INS) con el fin de obtener la trayectoria de personas. Se utilizan tres sensores: giroscopio, acelerómetro y un sensor de presión atmosférica. Además, se utilizan filtros de partículas y Kalman con el fin de reducir el ruido del posicionamiento.

Algunos sistemas, como la API de geolocalización de *Google Maps* [18], utilizan información de dos tecnologías diferentes (GSM y WiFi) para obtener la posición de usuarios por medio de mapas *fingerprint*. En la línea de utilizar mapas *fingerprint* globales, los autores de [17] plantean un método de actualización continua del mapa mediante *crowdsourced* WiFi, utilizando procesamiento *off-line* de las trayectorias seguidas por los usuarios.

Desde el punto de vista de la visión artificial, los autores de [12] analizan cómo generar un *fingerprint map* con un sensor RGB-D montado en un robot. La utilización de Kinect permite generar un mapa en tiempo real y grabar, en los puntos por donde pasa el robot, las mediciones de intensidad RSSI a distintos puntos WiFi. También en [3], un robot se localiza por medio de tres sistemas diferentes: un telémetro láser, una cámara de profundidad y los valores RSSI. Cada sistema se utiliza independientemente de acuerdo con la zona donde se encuentra el robot. Sin embargo, ninguna de estas

dos soluciones explota las diferentes tecnologías simultáneamente. En el primer trabajo, los sensores RGB-D se utilizan para obtener la posición real, y en el segundo, los autores solo utilizan un sistema a la vez.

En [16], los autores proponen combinar la identificación y el posicionamiento basados en RFID con el sensor Kinect para obtener la posición precisa de una persona dentro de un entorno. En este caso, un lector RFID se encuentra fijo en la habitación. Cada usuario lleva su propia etiqueta RFID, que permite identificarle, mientras que el sensor Kinect obtiene los esqueletos de las personas de cara a mejorar la precisión de localización. Cada esqueleto se compone de las coordenadas de las diferentes articulaciones de una persona: cuello, hombros, codos, rodillas, etc.

1.2 Funcionamiento del algoritmo previo

Como se ha comentado anteriormente, en este artículo se presenta una versión mejorada de un algoritmo de alta precisión en interiores [7], que a su vez generalizaba el algoritmo presentado en [6]. En dicho trabajo, se busca identificar y estimar la posición de personas en escenarios complejos habitados, donde se mueven simultáneamente. Ambas tecnologías, WPS y cámaras de profundidad, están conectadas a un servidor central como se muestra en la Figura 1. El WPS hace uso de varios puntos de acceso (APs) ubicados alrededor del entorno (ej., *routers*), que están al alcance de los teléfonos móviles de los usuarios, que hacen uso de la aplicación de posicionamiento. Los teléfonos miden el nivel de las señales recibidas y envían los datos a un servidor web. Simultáneamente, cada sensor RGB-D registra una parte del escenario, mediante una cámara y un sistema de proyección infrarroja, obteniendo mapas de profundidad. Dichos mapas se procesan para identificar los esqueletos de las personas y afinar los resultados obtenidos por el WPS, mediante mapas *fingerprint*, proporcionando valores más precisos para el posicionamiento. Para aumentar la precisión del algoritmo se tienen en cuenta las posiciones por donde se han movido las personas en distintos instantes de tiempo o *time stamps*. Finalmente, el servidor central obtiene las posiciones de las personas y las devuelve a los teléfonos móviles.

Los esqueletos se obtienen mediante las técnicas presentadas en [2, 14], donde los autores proponen algoritmos para predecir con rapidez y precisión las posiciones 3D de las articulaciones del cuerpo a partir de imágenes en profundidad.

El algoritmo presenta un error de posicionamiento inferior a 0.6 metros para 20 usuarios, si bien tiene una limitación en cuanto al número de personas detectadas por ambos sistemas de posicionamiento. El número de personas detectadas por los sensores RGB-D tiene que ser igual al número de personas detectadas mediante el posicionamiento WiFi. En este trabajo se amplía el algoritmo para permitir un número variable de personas detectadas por un sistema u otro. Además, se presenta una simplificación de los sensores Kinect en el escenario UCS.

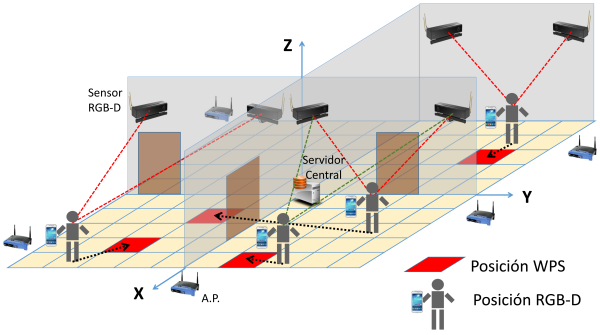

Figura 1: Escenario del sistema

En el siguiente apartado, Sección 2, se presenta la mejora del algoritmo para permitir un distinto número de personas detectadas por los dos sistemas. A continuación, en la Sección 3, se presenta el método de colocación de los sensores RGB-D en un entorno complejo de múltiples habitaciones. Finalmente, en la Sección 4, se muestran las conclusiones obtenidas y futuros trabajos planteados.

2 ALGORITMO DE EMPAREJAMIENTO EXTENDIDO

Una de las limitaciones del algoritmo presentado en [7] es que el número de usuarios detectados a través de las cámaras RGB-D debe ser igual al detectado utilizando WPS. En esta sección, se presenta la modificación del algoritmo para ser válido para cualquier situación.

El algoritmo desarrollado para resolver este nuevo problema tiene en cuenta las siguientes consideraciones:

- Sea n_1 el número de usuarios detectado a través de WPS, y n_2 el número de usuarios detectados por las cámaras RGB-D.
- Sea m el número de *time stamps* registrados en el experimento. Ambos sistemas de posicionamiento proporcionan sus correspondientes medidas de posición para determinados usuarios en cada *time stamp* (para n_1 usuarios en WPS y n_2 usuarios en RGB-D).

- Sea $P_i(t)$ la posición del usuario i provista por el WPS en el *time stamp* t, y sea $P_j^{'}(t)$ la posición del esqueleto del usuario j provista por los sensores RGB-D en el *time stamp* t. Al igual que en el algoritmo anterior, hay que tener en cuenta que la numeración de usuarios de la WPS no es la misma que la de los sensores RGB-D, lo que significa que inicialmente $P_i(t)$ no está asociado a $P_i^{'}(t)$ sino a otro $P_j^{'}(t)$. El problema, también en este algoritmo, consiste en determinar para todos los *time stamp* considerados, cuál es el índice i asociado a cada índice j.

- Sea s_{ij} el elemento binario de la *Matriz de Emparejamiento,* S, en la fila i, columna j. Dicho elemento enlaza la trayectoria del usuario i, provista por WPS, y la trayectoria del usuario j, provista por el sistema de sensores RGB-D.

- La solución del algoritmo es encontrar la matriz S, válida para el conjunto completo de datos del experimento, tal que:

$$\text{si } n_1 > n_2, \begin{cases} \forall j = 1, ..., n_2 : \sum_{x=1}^{n_1} s_{xj} = 1 \\ \forall i = 1, ..., n_1 : \sum_{y=1}^{n_2} s_{iy} \leq 1 \\ \sum_{x=1}^{n_1} \sum_{y=1}^{n_2} s_{xy} = n_2 \end{cases} \quad (1)$$

$$\text{si } n_1 = n_2 = n, \begin{cases} \forall j = 1, ..., n_2 : \sum_{x=1}^{n_1} s_{xj} = 1 \\ \forall i = 1, ..., n_1 : \sum_{y=1}^{n_2} s_{iy} = 1 \\ \sum_{x=1}^{n_1} \sum_{y=1}^{n_2} s_{xy} = n \end{cases} \quad (2)$$

$$\text{si } n_2 > n_1, \begin{cases} \forall j = 1, ..., n_2 : \sum_{x=1}^{n_1} s_{xj} \leq 1 \\ \forall i = 1, ..., n_1 : \sum_{y=1}^{n_2} s_{iy} = 1 \\ \sum_{x=1}^{n_1} \sum_{y=1}^{n_2} s_{xy} = n_1 \end{cases} \quad (3)$$

El caso $n_1 = n_2 = n$ corresponde al algoritmo presentado en [7], donde el número de usuarios detectados por cada uno de los dos sistemas es igual. Gráficamente, se puede ver en la Figura 2 una matriz S, donde $n_1 > n_2$. En la imagen se puede observar cómo los cálculos coinciden con las consideraciones indicadas.

Al igual que en el algoritmo anterior, para encontrar la matriz S, todos los m *time stamps* considerados deben ser tenidos en cuenta, de forma que el problema es una optimización de la combinación entre las trayectorias WPS y las de los esqueletos para todos los usuarios. Al igual que en el algoritmo anterior, la mejor combinación es la que consigue que la suma de distancias entre todas las trayectorias de usuarios sea mínima, como se muestra en la Expresión 4, adaptada a este nuevo algoritmo extendido.

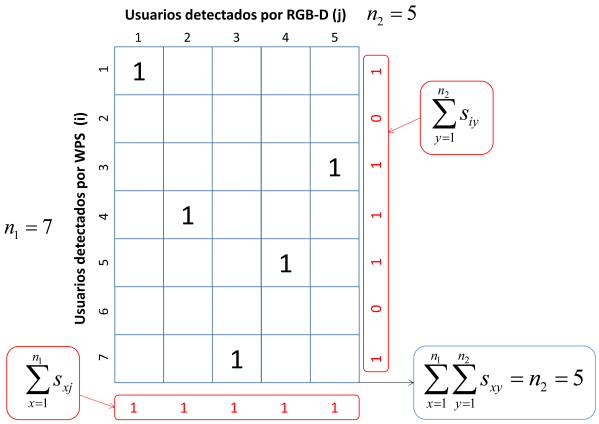

Figura 2: Ejemplo de matriz S para un distinto número de usuarios WPS y RGB-D.

$$\min \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} s_{ij} \cdot \sum_{t=1}^{m} d_E(P_i(t), P_j^{'}(t)) \quad (4)$$

donde d_E representa la distancia euclídea entre un esqueleto y una celda WPS. Esta se mide como la distancia entre las coordenadas del cuello del esqueleto y el centroide de la celda WPS. El algoritmo extendido se resuelve utilizando el *Algoritmo Aditivo de Balas* [1].

Al haber un distinto número de usuarios detectados por WPS y RGB-D, algún usuario no será enlazado. Esto podría ocurrir, por ejemplo, si un usuario no lleva consigo un móvil pero es detectado por las cámaras. Como el sistema busca la combinación óptima, se minimizan los posibles falsos positivos.

3 COLOCACIÓN DE LOS SENSORES RGB-D

Los sensores RGB-D utilizados han sido sensores Kinect v2 [19], sensores que permiten obtener mapas de profundidad y extraer los esqueletos de las personas que detectan, incluyendo en cada esqueleto las coordenadas de las articulaciones de cada persona dentro del mapa de profundidad. Al utilizar múltiples sensores RGB-D, es necesario colocarlos en distintas posiciones del escenario y transformar las coordenadas que devuelven en co-

ordenadas de un mismo UCS. El método de colocación, presentado por los autores en [5], ha sido simplificado para un escenario como el mostrado en la Figura 3, donde se pueden ver con círculos los sensores RGB-D así como las trayectorias de tres usuarios seguidas tanto por WPS como por los sensores Kinect.

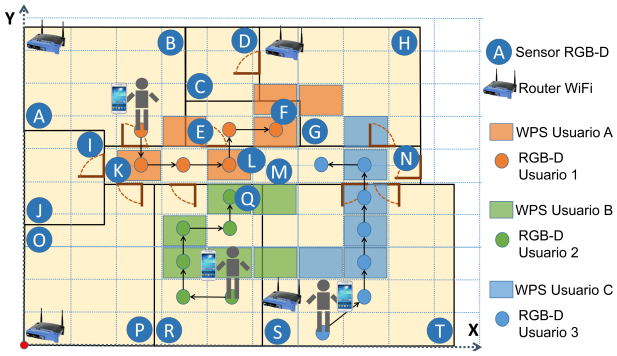

Figura 3: Plano del escenario con trayectorias de usuarios y sensores RGB-D

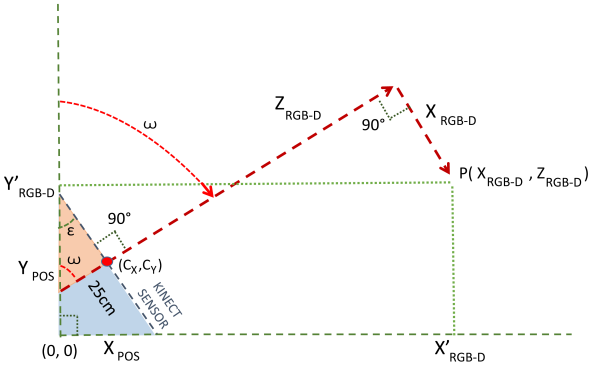

Figura 4: Simplificación de la transformación mediante matriz de rotación y traslación

El método de posicionamiento de los sensores dentro del UCS se realiza mediante el uso de una matriz de rotación seguida de una traslación. La matriz de rotación permite girar las coordenadas obtenidas por el sensor RGB-D respecto al UCS. Para ello es necesario rotar el ángulo ω en el sentido de las agujas del reloj (ver Figura 4). El ángulo ω puede ser calculado a partir de ε, obtenido en la Ecuación 5. Ambos ángulos pertenecen a un triángulo, con otro ángulo de 90°, por lo que el valor de ω es el mostrado en la Ecuación 6.

$$\varepsilon = \arcsin\left(\frac{X_{POS}}{25}\right) \tag{5}$$

$$\omega = 90^\circ - \varepsilon = 90^\circ - arcsin\left(\frac{X_{POS}}{25}\right) \tag{6}$$

La matriz de rotación general en dos dimensiones viene determinada por la Expresión 7. El ángulo de giro que considera es en contra de las agujas del reloj.

$$R(\mu) = \begin{bmatrix} \cos\mu & -\sin\mu \\ \sin\mu & \cos\mu \end{bmatrix} \tag{7}$$

El vector de traslación respecto al origen de coordenadas de la sala viene determinado por $T(C_X, C_Y)$. Aplicando primero la rotación, según las agujas del reloj, y después la traslación, se obtienen las coordenadas del sensor respecto a la sala según la Ecuación 8.

$$\begin{bmatrix} X'_{RGB-D} \\ Y'_{RGB-D} \end{bmatrix} = \begin{bmatrix} \cos(-\omega) & -\sin(-\omega) \\ \sin(-\omega) & \cos(-\omega) \end{bmatrix} \begin{bmatrix} X_{RGB-D} \\ Z_{RGB-D} \end{bmatrix} + \begin{bmatrix} C_X \\ C_Y \end{bmatrix} \tag{8}$$

Utilizando la Ecuación 9, se pueden obtener las expresiones finales que transforman las coordenadas obtenidas por el sensor en coordenadas del UCS (ver Ecuación 10). Como los sensores devuelven los valores en la misma unidad que se utiliza en el UCS, es decir metros, no es necesario realizar un escalado.

$$\begin{cases} X_{UCS} = X_{CORNER} + X'_{RGB-D} \\ Y_{UCS} = Y_{CORNER} + Y'_{RGB-D} \end{cases} \tag{9}$$

$$\begin{bmatrix} X_{UCS} \\ Y_{UCS} \end{bmatrix} = \begin{bmatrix} \cos(-\omega) & -\sin(-\omega) \\ \sin(-\omega) & \cos(-\omega) \end{bmatrix} \begin{bmatrix} X_{RGB-D} \\ Z_{RGB-D} \end{bmatrix} + \begin{bmatrix} C_X \\ C_Y \end{bmatrix} + \begin{bmatrix} X_{CORNER} \\ Y_{CORNER} \end{bmatrix} \tag{10}$$

4 CONCLUSIONES

Se ha presentado en este artículo una mejora al algoritmo de posicionamiento previamente desarrollado por los autores, el cual ofrece un error de posicionamiento inferior a 0.6 metros si hay 20 personas en el escenario, e inferior a 0.3 metros si hay hasta 10 usuarios. El algoritmo se ha mejorado en dos sentidos, tanto a nivel de colocación de los sensores en el UCS, como a nivel de la limitación existente por la que el número de personas detectadas por cada sistema de posicionamiento debía

ser igual. La utilización de múltiples cámaras reduce los problemas de oclusión, pero cabe destacar que el algoritmo extendido funciona aunque alguna persona no sea detectada por los sensores RGB-D.

Aunque el emparejamiento de rutas presentado en este trabajo integra trayectorias de personas obtenidas mediante WPS y RGB-D, se plantea como trabajo futuro el estudio del algoritmo con rutas obtenidas mediante otros sistemas de posicionamiento, como pueden ser el GPS o las señales de las torres de telefonía móvil.

Agradecimientos

Este trabajo ha sido desarrollado gracias a la ayuda del proyecto de investigación DPI2016-77677-P del MICINN. También pertenece a las actividades llevadas a cabo en el marco de la red de investigación CAM RoboCity2030 S2013/MIT-2748, de la Comunidad de Madrid.

English summary

IMPROVEMENT OF THE INDOOR POSITIONING SYSTEM BASED ON RGB-D SENSORS AND WIFI NETWORKS

Abstract

This paper presents an improvement of an indoor positioning algorithm previously developed by the authors. This algorithm obtains the position of several people who move freely in an inhabited environment composed of different rooms. The algorithm takes into account computer vision positioning and WiFi positioning, combining the information from both systems to increase accuracy. The improvement proposed in this article refers to a limitation of the algorithm by which the number of people detected by RGB-D sensors has to be equal to the number of people detected by WiFi. The improvement allows working with different number of people in each system, which can be common in scenarios where a user is not detected by one of the methods. Besides, the improvement also simplifies the method of placement of RGB-D sensors in the scenario.

Keywords: indoor positioning, IPS, WPS, RGB-D, Kinect, WiFi.

Referencias

[1] Balas, E., (1965) An additive algorithm for solving linear programs with zero-one variables, *Operations Research*, 13(4):517–546.

[2] Barmpoutis, A., (2013) Tensor body: Real-time reconstruction of the human body and avatar synthesis from rgb-d, *Cybernetics, IEEE Transactions on*, 43(5):1347–1356.

[3] Biswas, J., et Veloso, M., (2014) Multi-sensor mobile robot localization for diverse environments, *RoboCup 2013: Robot World Cup XVII*, pages 468–479.

[4] Deng, Z., Yu, Y., Yuan, X., Wan, N. et Yang, L., (2013) Situation and development tendency of indoor positioning, *China Communications*, 10(3):42–55.

[5] Duque Domingo, J., Cerrada, C., Cerrada, J.A. et Valero, E., (2017) Locating multiple camera sensors and wireless access points for a generalized indoor positioning system, *ICWMC 2017*, page 59.

[6] Duque Domingo, J., Cerrada, C., Valero, E., et Cerrada, J.A., (2016) Indoor positioning system using depth maps and wireless networks, *Journal of Sensors*.

[7] Duque Domingo, J., Cerrada, C., Valero, E., et Cerrada, J.A., (2017) An improved indoor positioning system using rgb-d cameras and wireless networks for use in complex environments, *Sensors*, 17(10):2391.

[8] Evennou, F., et Marx, F., (2006) Advanced integration of WiFi and inertial navigation systems for indoor mobile positioning, *Eurasip journal on applied signal processing*, 2006:164–164.

[9] Harle, R., (2013) A survey of indoor inertial positioning systems for pedestrians, *IEEE Communications Surveys and Tutorials*, 15(3):1281–1293.

[10] He, S., Chan, S., et Gary, H., (2016) WiFi fingerprint-based indoor positioning: Recent advances and comparisons, *IEEE Communications Surveys & Tutorials*, 18(1):466–490.

[11] Machaj, J., et Brida, P. (2016) Impact of optimization algorithms on hybrid indoor positioning based on GSM and WiFi signals, *Concurrency and Computation: Practice and Experience*.

[12] Mirowski, P., Palaniappan, R., et Ho, T.K., (2012) Depth camera slam on a low-cost WiFi mapping robot, *Technologies for Practical Robot Applications (TePRA), IEEE International Conference on*, pages 1–6. IEEE.

[13] Papapostolou, A., et Chaouchi, H., (2012) Integrating rfid and wlan for indoor positioning and IP movement detection, *Wireless Networks*, 18(7):861–879.

[14] Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., Cook, M., et Moore, R., (2013) Real-time human pose recognition in parts from single depth images, *Communications of the ACM*, 56(1):116–124.

[15] Su, H.K., Liao, Z.X., Lin, C.H., et Lin, T.M., (2015) A hybrid indoor-position mechanism based on bluetooth and WiFi communications for smart mobile devices, *2015 International Symposium on Bioelectronics and Bioinformatics (ISBB)*, pages 188–191.

[16] Wang, C.S., et Chen, C.L., (2014) RFID-based and kinect-based indoor positioning system, *Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronic Systems (VITAE), 2014 4th International Conference on*, pages 1–4. IEEE.

[17] Wilk, P., Karciarz, J., et Swiatek, J., (2015) Indoor radio map maintenance by automatic annotation of crowdsourced WiFi fingerprints, *Indoor Positioning and Indoor Navigation (IPIN), 2015 International Conference on*, pages 1–8. IEEE, 2015.

[18] Google Maps Geolocation API, *https://developers.google.com/maps/ documentation/geolocation (último acceso: 31 de Mayo de 2018)*.

[19] Kinect v2, *https://developer.microsoft.com/ en-us/windows/kinect (último acceso: 15 de Abril de 2018)*.

© 2018 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution CC-BY-NC 3.0 license (http://creativecommons.org/licenses/by-nc/3.0/).