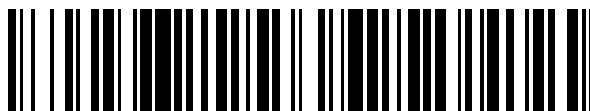


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 758 734**

51 Int. Cl.:

**G06T 7/80** (2007.01)

**H04N 17/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.05.2009** **E 09450093 (1)**

97 Fecha y número de publicación de la concesión europea: **04.09.2019** **EP 2249580**

54 Título: **Procedimiento para calibrar la imagen de una cámara**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**06.05.2020**

73 Titular/es:  
**KAPSCH TRAFFICOM AG (100.0%)**  
**Am Europlatz 2**  
**1120 Wien , AT**

72 Inventor/es:  
**KAHLER, PETER;**  
**JANNER, CHRISTIAN y**  
**LÖHNDORF, MAIKE**

74 Agente/Representante:  
**ELZABURU, S.L.P**

**ES 2 758 734 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

**DESCRIPCIÓN**

Procedimiento para calibrar la imagen de una cámara

5 La presente invención se refiere a un procedimiento para calibrar la imagen de una cámara, en particular una cámara de vigilancia del tráfico, con respecto a su espacio visual, para lo cual se determinan los parámetros de proyección de cámara de una proyección del espacio visual sobre la imagen.

10 Las proyecciones de cámara pueden modelarse sobre clases muy diferentes, por ejemplo, como modelo de cámara oscura, modelo de "lente gruesa" o de "lente delgada", modelo de ojo de pez, etc., con o sin considerar defectos y distorsiones de lentes. Es común a todos los modelos que pueda describirse la proyección por medio de los denominados parámetros de proyección que forman directamente, por ejemplo en modelos simplificados, los coeficientes de una matriz de proyección geométrica.

15 El uso de una transformación lineal directa (DLT) para determinar parámetros de proyección de cámara es conocido desde hace tiempo, véase, por ejemplo, en Abdel-Aziz, Y.I., y Karara, H.M: "Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close-range Photogrammetry", Proceedings of the Symposium on Close-Range Photogrammetry, páginas 1-18, 1971, Falls Church, Virginia, American Society of Photogrammetry. Dos procedimientos de calibración basados en la transformación lineal directa (DLT) conocida se describen en Hatze, H: "High-precision three-dimensional photogrammetric calibration and object space reconstruction using a modified DLT-approach", Journal of Biomechanics, Pergamon Press, Nueva York, NY, US, Bd. 21, N° 7, 1 de enero de 1988, páginas 533-538, ISSN: 0021-9290.

20 Para la aplicación de la DLT a la determinación de los parámetros de proyección de cámara es necesario el conocimiento de al menos seis puntos de confinamiento espacial en el espacio visual, los denominados "puntos mundiales" y su asociación a los puntos de imagen en la imagen de cámara. Para ello, según el actual estado de la técnica, seis marcas deben disponerse distribuidas en todo el espacio visual de la cámara, acotarse y seguidamente asociarse a los puntos de imagen correctos en la imagen de cámara.

25 En cámaras de vigilancia del tráfico para túneles de carretera o autopistas, que pueden tener un gran campo de visión de hasta varios centenares de metros, esto significa un elevado coste de tiempo porque las marcas individuales deben disponerse, por ejemplo, a mano sobre la carretera y acotarse una con respecto a otra por medio de telémetros de láser o cintas métricas. En general, para ello es necesario también un bloqueo del túnel o del tramo de carretera, lo cual es extremadamente dificultoso: precisamente en cámaras de túnel es necesaria una nueva calibración frecuente de las cámaras de vigilancia porque estas se pueden desajustar durante un lavado regular del túnel o por colisiones de vehículos pesados que chocan con este. Asimismo, en cada operación de reparación o mantenimiento de una cámara, esta debe calibrarse de nuevo.

La invención tiene el objetivo de crear un procedimiento para calibrar imágenes de cámara, que pueda realizarse de forma más sencilla y rápida que los procedimientos conocidos hasta ahora. Este objetivo se logra con un procedimiento del tipo citado al principio que se caracteriza por las etapas de la reivindicación 1.

35 Con ayuda del procedimiento según la invención, ya no es necesario colocar y medir a mano las al menos seis marcas en el espacio visual necesarias para determinar los parámetros de proyección de cámara. En lugar de ello, se utiliza un soporte móvil, por ejemplo un vehículo, con marcas medidas previamente, que se mueve a través del campo de visión. En este caso, ni siquiera es necesario conocer la extensión del movimiento en el mundo real, de modo que, por ejemplo, pueden seleccionarse en gran medida de forma discrecional la velocidad y la dirección de marcha del soporte.

40 Los parámetros de proyección de cámara así determinados hacen posible seguidamente una reasignación del sistema de coordenadas de imagen de la imagen de cámara a un plano arbitrariamente seleccionado en el sistema de coordenadas mundiales del espacio de cámara, por ejemplo a un plano de carretera, con lo que se logra el objetivo de la calibración. A partir de la imagen de cámara calibrada de esta manera, pueden derivarse posteriormente otras informaciones, por ejemplo la velocidad y la dirección de marcha de un vehículo sobre una carretera contenida en la imagen. En el procedimiento propuesto se realiza el cálculo citado en un modelo de matriz con ayuda del método de la transformación lineal directa ("DLT"), en particular, de preferencia por solución de la ecuación DLT

$$D(s,t) \cdot m = p$$

50 con respecto a m, s y t, con

D(s,t), matriz DLT de todos los puntos mundiales y de imagen, en la que los segundos puntos mundiales se expresan por los primeros puntos mundiales desplazados con el decalaje (s, t, 0),

m, vector de los parámetros de proyección de cámara, y

p, vector de todos los puntos de imagen.

Esta forma de realización es especialmente adecuada para cámaras de vigilancia del tráfico que observan un plano de carretera, sobre el que se desplaza el soporte con un decalaje bidimensional, de modo que la tercera coordenada del decalaje pueda suponerse igual a cero.

- 5 Según una forma de realización especialmente ventajosa de la invención, para resolver la ecuación DLT se resuelve la tarea de minimización

$$\|D(s,t) \cdot m - p\| \rightarrow \min$$

10 por medio del procedimiento de Gauss-Newton. El procedimiento de Gauss-Newton se conoce en matemáticas, de modo que es necesario entrar aquí más en detalle. El uso del procedimiento de Gauss-Newton permite una solución rápida de la ecuación DLT con suficiente exactitud.

15 Para la forma de funcionamiento básica del procedimiento, no se requiere necesariamente que todas las marcas se detecten en cada imagen, siendo suficiente que se detecten al menos cuatro marcas en una imagen y al menos tres marcas en otra imagen. A partir de estos 7 puntos de imagen se pueden elaborar 14 ecuaciones, con las que se pueden determinar los 11 parámetros de proyección de cámara de un modelo de proyección de cámara oscura sencillo y hasta 3 coordenadas del decalaje.

20 En cualquier caso, es especialmente favorable que, como soporte, se utilice un vehículo y como decalaje se utilice un trayecto de circulación del vehículo. Por tanto, el procedimiento según la invención es adecuado especialmente para calibrar las imágenes de cámaras de vigilancia del tráfico, en particular cámaras de túneles y autopistas. El procedimiento de la invención puede realizarse durante los tiempos de funcionamiento del túnel o de la carretera y no requiere ninguna medida de acotamiento o bloqueo en el tramo de carretera.

Preferiblemente, se seleccionan como marcas puntos destacados del vehículo o se instalan, en particular, medios luminiscentes sobre un soporte trasero y un soporte de techo o de espejo retrovisor del vehículo, de modo que los vehículos convencionales puedan reequiparse para ello de manera correspondiente.

25 Es especialmente ventajoso que, según otra característica de la invención, se capturen más de dos imágenes para detectar otros puntos de imagen de las marcas, que se incorporan en el cálculo de los parámetros de proyección de cámara, con lo que puede elevarse la precisión del procedimiento.

Otra forma de realización del procedimiento de la invención se caracteriza por que se ejecuta repetidamente y se promedian los respectivos parámetros de proyección de cámara obtenidos, con lo que puede mejorarse aún más la precisión.

30 La invención se explica seguidamente con más detalle con ayuda de un ejemplo de realización representado en los dibujos adjuntos. En los dibujos, muestran:

la figura 1, el procedimiento de la invención con ayuda de dos posiciones de un vehículo que sirve como soporte de marca sobre una carretera, y

35 las figuras 2a y 2b, las imágenes de la cámara de vigilancia del tráfico de la figura 1 que pertenecen a las dos posiciones del vehículo.

La figura 1 muestra una cámara 1, en cuyo espacio visual se encuentra una carretera 2, de la que la cámara 1 toma al menos dos imágenes 3<sub>I</sub>, 3<sub>II</sub> (figuras 2a, 2b) en un intervalo temporal. Por ejemplo, la cámara 1 es una videocámara con una tasa de repetición de imágenes de, por ejemplo, 25 fps (cuadros por segundo).

40 Un sistema de coordenadas mundiales (x, y, z) está asociado a los objetos reales (puntos mundiales) en el espacio visual de la cámara 1, y un sistema de coordenadas de imagen u, v está asociado a los puntos de imagen (píxeles) de las imágenes de cámara bidimensionales 3<sub>I</sub>, 3<sub>II</sub>.

Los objetos en el espacio visual pueden ser en sí de cualquier tipo; aplicaciones a modo de ejemplo son el uso de una cámara 1 para vigilar vías de agua, lugares públicos, zonas peatonales, aeropuertos, pistas de rodadura, estacionamientos, edificios, áreas de entrada y espera, etc.

45 Como conoce el experto, la reproducción proyectiva geométrica de puntos mundiales w<sub>i</sub> con las coordenadas (x<sub>i</sub>, y<sub>i</sub>, z<sub>i</sub>), por ejemplo los puntos mundiales mostrados w<sub>1</sub>, w<sub>2</sub>, ...w<sub>8</sub>, sobre puntos de imagen p<sub>i</sub> con las coordenadas (u<sub>i</sub>, v<sub>i</sub>), por ejemplo, los puntos de imagen mostrados p<sub>1</sub>, p<sub>2</sub>, ...p<sub>8</sub>, por ejemplo como modelo de cámara oscura en perspectiva puede modelarse como:

$$\begin{pmatrix} \sigma u_i \\ \sigma v_i \\ \sigma \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & 1 \end{pmatrix} \cdot \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} \quad (1)$$

o abreviadamente

$$p_i = M \cdot w_i \quad (2)$$

- 5 La matriz M se designa también matriz de proyección de cámara del modelo de cámara seleccionado y sus coeficientes  $m_{kl}$  se designan también parámetros de proyección de cámara. Si se utilizan modelos de cámara distintos del modelo de cámara oscura aquí descrito en detalle, por ejemplo, modelos de cámara de "lente gruesa", "lente delgada", de ojo de pez u otros modelos, con o sin consideración de defectos o distorsiones de lentes, entonces la proyección de cámara puede modelarse también por más y/u otros parámetros de proyección de cámara que no forman necesariamente solo una matriz de este tipo.
- 10 Cuando se conocen la posición exacta de la cámara 1 y todos sus parámetros internos como la distancia focal, el tamaño de sensor, etc., se podrían construir directamente los parámetros de la proyección de cámara, aquí la matriz M. El objetivo de la calibración de cámara es determinar sin este conocimiento los parámetros de proyección de cámara, aquí la matriz M. Por tanto, en el presente ejemplo, pueden obtenerse los 11 coeficientes  $m_{11}, m_{12}, \dots, m_{33}$ .
- 15 En el supuesto de que conozcan al menos seis asociaciones de puntos mundiales  $w_i$  a puntos de imagen  $p_i$ , los 11 parámetros de proyección de cámara  $m_{11}, m_{12}, \dots, m_{33}$  de la matriz M pueden obtenerse como parámetros de una transformación lineal directa (DLT) según la siguiente ecuación DLT (aquí establecida para ocho asociaciones; la finalidad de esto se pondrá claramente de manifiesto más tarde):

$$\begin{pmatrix} x_1 & y_1 & z_1 & 1 & 0 & 0 & 0 & 0 & -u_1 x_1 & -u_1 y_1 & -u_1 z_1 \\ 0 & 0 & 0 & 0 & x_1 & y_1 & z_1 & 1 & -v_1 x_1 & -v_1 y_1 & -v_1 z_1 \\ x_2 & y_2 & z_2 & 1 & 0 & 0 & 0 & 0 & -u_2 x_2 & -u_2 y_2 & -u_2 z_2 \\ 0 & 0 & 0 & 0 & x_2 & y_2 & z_2 & 1 & -v_2 x_2 & -v_2 y_2 & -v_2 z_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_8 & y_8 & z_8 & 1 & 0 & 0 & 0 & 0 & -u_8 x_8 & -u_8 y_8 & -u_8 z_8 \\ 0 & 0 & 0 & 0 & x_8 & y_8 & z_8 & 1 & -v_8 x_8 & -v_8 y_8 & -v_8 z_8 \end{pmatrix} \cdot \begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{21} \\ \vdots \\ m_{33} \end{pmatrix} = \begin{pmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ \vdots \\ u_8 \\ v_8 \end{pmatrix} \quad (3)$$

o abreviadamente

$$D \cdot m = p \quad (4)$$

- 20 En una primera etapa, un soporte móvil 4 en el espacio visual de la cámara 1, en el ejemplo mostrado un vehículo, se provee ahora de al menos cuatro marcas o primeros puntos mundiales  $w_1, w_2, w_3, w_4$  visibles por la cámara 1. Los puntos mundiales  $w_1, w_2, w_3, w_4$  se seleccionan de manera que abarquen un volumen y sea conocida su posición en el sistema de coordenadas mundiales  $x, y, z$ .
- 25 En una primera posición I del soporte 4, la cámara 1 toma una primera imagen  $3_I$ . En la imagen  $3_I$ , inmediatamente o con posterioridad, manualmente o con ayuda de procedimientos de reconocimiento de imagen, se detectan los puntos de imagen  $p_1, p_2, p_3, p_4$  de las cuatro marcas o puntos mundiales  $w_1, w_2, w_3, w_4$ .
- Seguidamente, el soporte 4 se desplaza con un decalaje  $v(s, t, 0)$  en el espacio visual de la cámara 1, por ejemplo el vehículo se mueve avanzando en un trayecto de circulación. En esta segunda posición II, se captura de nuevo una
- 30 imagen  $3_{II}$  y se detectan en ella los nuevos puntos de imagen  $p_5, p_6, p_7, p_8$  de las marcas o puntos mundiales  $w_1, w_2, w_3, w_4$  originales desplazados con respecto a los nuevos puntos mundiales  $w_5, w_6, w_7, w_8$ .
- Por tanto, para resolver la ecuación (3), los nuevos o segundos puntos mundiales  $w_5, w_6, w_7, w_8$  pueden expresarse por el decalaje  $v(s, t, 0)$  de los viejos o primeros cuatro puntos mundiales  $w_1, w_2, w_3, w_4$  como:

$$(x_5, y_5, z_5) = (x_1 + s, y_1 + t, z_1) \quad (5)$$

$$(x_6, y_6, z_6) = (x_2 + s, y_2 + t, z_2) \quad (6)$$

$$(x_7, y_7, z_7) = (x_3 + s, y_3 + t, z_3) \quad (7)$$

$$(x_8, y_8, z_8) = (x_4 + s, y_4 + t, z_4) \quad (8)$$

5 Como puede verse, se asumió en el ejemplo mostrado un decalaje  $v$  paralelo al plano  $x/y$  del sistema de coordenadas mundiales seleccionado, en el supuesto de que el soporte 4 sea un vehículo, que se mueve a distancia  $z$  constante del plano de la carretera 2, y se haya fijado correspondientemente el sistema de coordenadas mundiales  $x, y, z$ . Sin embargo, en caso de que se desee, podría utilizarse también un decalaje  $v$  tridimensional.

Por tanto, en la ecuación (3) o (4), en las últimas ocho filas de la matriz  $D$ , los puntos mundiales  $w_5, w_6, w_7, w_8$  pueden expresarse por los puntos mundiales  $w_1, w_2, w_3, w_4$  y los parámetros de decalaje  $s, t$ , y nosotros llamamos a esto matriz modificada  $D(s, t)$ .

10 La ecuación (3) o (4) posee ahora 13 incógnitas: 11 parámetros de proyección de cámara  $m_{kl}$  de la matriz de proyección de cámara  $M$  y 2 parámetros de decalaje  $s, t$ . La ecuación DLT (3) o (4) se convierte así en la nueva ecuación DLT

$$D(s, t) \cdot m = p \quad (9)$$

La ecuación (9) puede resolverse seguidamente como tarea de minimización de la forma:

$$\|D(s, t)m - p\| \rightarrow \min \quad (10)$$

15 En la tarea de minimización se busca  $m_{11}, m_{12}, \dots, m_{33}$  y  $s$  y  $t$ , de modo que se haga mínima la norma  $\|D(s, t)m - p\|$ .

Para ello es adecuado básicamente cada procedimiento conocido en la técnica; la tarea de minimización (10) se resuelve preferentemente con ayuda del procedimiento de Gauss-Newton.

20 Como resultado, se determinan de esta manera – codeterminando simultáneamente el decalaje  $v$  – los parámetros de proyección de cámara  $m_{kl}$  de la matriz de proyección de cámara de la ecuación (1) de modelo de cámara oscura. Por tanto, como consecuencia adicional, a cada punto mundial cualquiera se le puede asociar un punto de imagen, o viceversa, a cada punto de imagen se le puede asociar un punto mundial en un plano seleccionado del espacio visual, por ejemplo, un punto sobre la carretera 2. Por tanto, pueden acotarse, por ejemplo, el lugar o la velocidad de los vehículos 4 sobre la carretera 2 a partir de las imágenes de cámara  $3_I, 3_{II}$ .

25 Dado que la ecuación DLT (9) o la tarea de minimización (10) solo tiene 13 incógnitas (siempre que se asuma un decalaje bidimensional  $v$ ) o 14 incógnitas (cuando se asume un decalaje tridimensional  $v$ ), son en sí suficientes para la solución siete asociaciones de punto mundial a punto de imagen  $w_i \leftrightarrow p_i$ . Por tanto, en el caso más sencillo, en una imagen se detectan cuatro puntos de imagen y en otra imagen se detectan también solo tres puntos de imagen; cualquier incremento adicional del número de las asociaciones detectadas  $w_i \leftrightarrow p_i$ , por ejemplo, aprovechando otras imágenes de cámara para más de dos posiciones  $I, II$  del soporte 4 y/o empleando más de cuatro marcas sobre el soporte 4, puede aprovecharse para mejorar la precisión de la solución. Asimismo, es posible repetir el procedimiento para estabilizar por promediado los parámetros de proyección de cámara determinados en las distintas pasadas.

35 La invención no está limitada a la forma de realización representada, sino que comprende todas las variantes y modificaciones que están en el ámbito de las reivindicaciones anexas.

**REIVINDICACIONES**

1. Procedimiento para calibrar la imagen (3<sub>i</sub>, 3<sub>ii</sub>) de una cámara 1, en particular una cámara de vigilancia del tráfico, con respecto a su espacio visual, para lo cual se determinan los parámetros de proyección de cámara (m<sub>kl</sub>) de una proyección de cámara del espacio visual sobre la imagen, caracterizado por las etapas de:

5 habilitar un soporte móvil (4) con al menos cuatro marcas visibles por la cámara (1) que confinan un volumen y forman unos primeros puntos mundiales conocidos (w<sub>i</sub>) en el espacio visual,

tomar una primera imagen (3<sub>i</sub>) del espacio visual,

desplazar el soporte (4) en el espacio visual con un decalaje desconocido (v) para obtener unos segundos puntos mundiales (w<sub>i</sub>) dependientes de este en el espacio visual,

10 tomar una segunda imagen (3<sub>ii</sub>) del espacio visual, detectar las marcas en la primera imagen (3<sub>i</sub>) como primeros puntos de imagen (p<sub>i</sub>) y en la segunda imagen (3<sub>ii</sub>) como segundos puntos de imagen (p<sub>i</sub>), y

determinar los parámetros de proyección de cámara (m<sub>kl</sub>), codeterminando el decalaje a partir de los primeros puntos mundiales conocidos, los segundos puntos mundiales dependientes del decalaje, los primeros puntos de imagen detectados y los segundos puntos de imagen detectados con ayuda del método de la transformación lineal directa ("DLT").

15 2. Procedimiento según la reivindicación 1, caracterizado por que la determinación se realiza resolviendo la ecuación DLT:

$$D(s,t) \cdot m = p$$

con respecto a m, s y t, significando

20 D(s, t), matriz DLT de todos los puntos mundiales y de imagen, en la que los segundos puntos mundiales se expresan por los primeros puntos mundiales desplazados con el decalaje (s, t, 0),

m, vector de los parámetros de proyección de cámara, y

p, vector de todos los puntos de imagen.

25 3. Procedimiento según la reivindicación 2, caracterizado por que, para resolver la ecuación DLT, se resuelve la tarea de minimización

$$\|D(s,t) \cdot m - p\| \rightarrow \min$$

por medio del procedimiento de Gauss-Newton.

4. Procedimiento según una de las reivindicaciones 1 a 3, caracterizado por que se detectan al menos cuatro marcas en una imagen y al menos tres marcas en otra imagen.

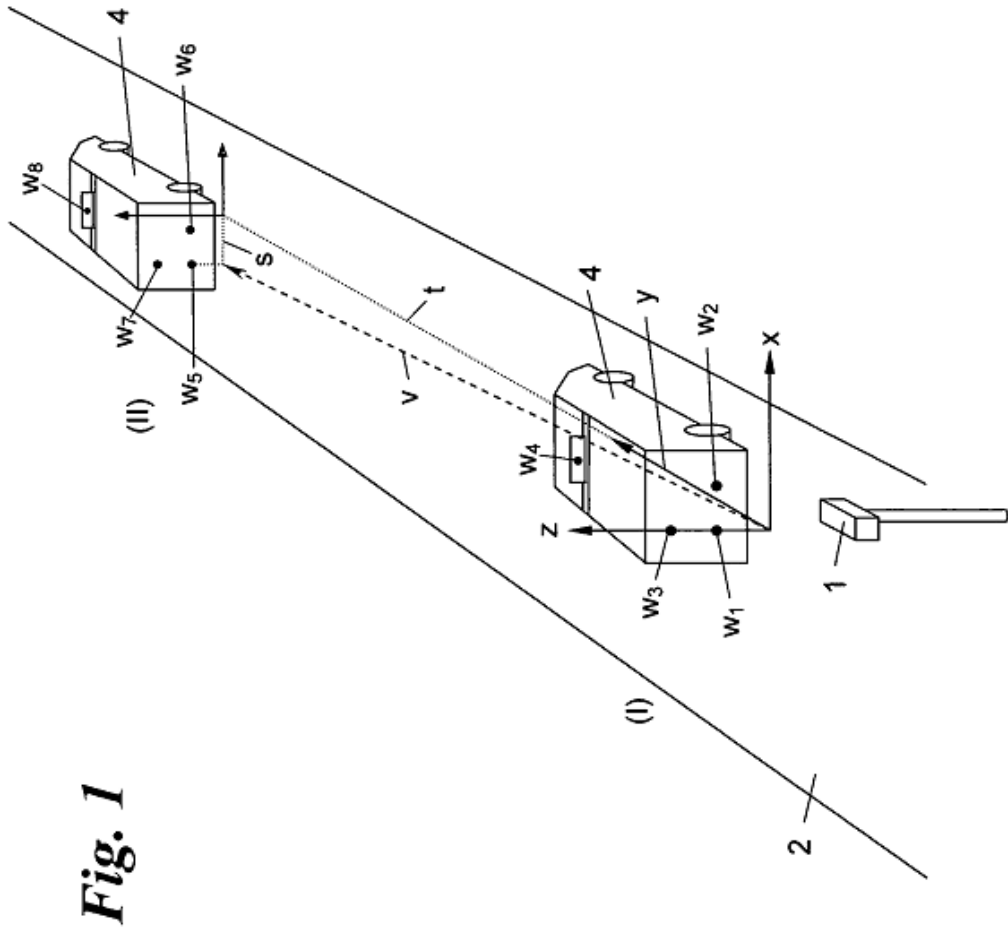
30 5. Procedimiento según una de las reivindicaciones 1 a 4, caracterizado por que se capturan más de dos imágenes (3<sub>i</sub>, 3<sub>ii</sub>) para detectar otros puntos de imagen (p<sub>i</sub>) de las marcas (w<sub>i</sub>), que se incorporan en la determinación de los parámetros de proyección de cámara (m<sub>kl</sub>).

6. Procedimiento según una de las reivindicaciones 1 a 5, caracterizado por que como soporte (4) se utiliza un vehículo y como decalaje (v) se utiliza un trayecto de circulación del vehículo (4).

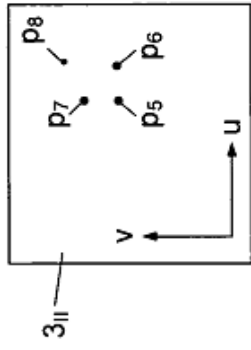
35 7. Procedimiento según la reivindicación 6, caracterizado por que como marcas se utilizan puntos destacados del vehículo (4).

8. Procedimiento según la reivindicación 6, caracterizado por que, como marcas (w<sub>i</sub>) se instalan medios luminiscentes sobre un soporte trasero y un soporte de techo o de espejo retrovisor del vehículo (4).

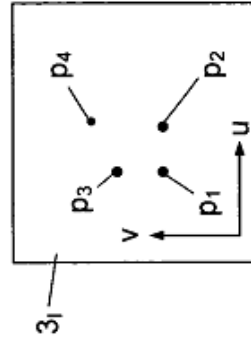
40 9. Procedimiento según una de las reivindicaciones 1 a 8, caracterizado por que este se repite y se promedian los respectivos parámetros de proyección de cámara determinados.



**Fig. 1**



**Fig. 2b**



**Fig. 2a**