

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 986 208**

51 Int. Cl.:

**A61C 9/00** (2006.01)

**G06T 5/50** (2006.01)

**G06T 7/33** (2007.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **04.06.2021 PCT/EP2021/065068**

87 Fecha y número de publicación internacional: **09.12.2021 WO21245274**

96 Fecha de presentación y número de la solicitud europea: **04.06.2021 E 21730582 (0)**

97 Fecha y número de publicación de la concesión europea: **17.04.2024 EP 4161437**

54 Título: **Toma de impresión óptica de la arcada dental de un paciente**

30 Prioridad:

**06.06.2020 FR 2005929**

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

**08.11.2024**

73 Titular/es:

**QUERBES, OLIVIER (50.0%)**

**Clairfont**

**31570 Vallesvilles, FR y**

**QUERBES, VÉRONIQUE (50.0%)**

72 Inventor/es:

**QUERBES, OLIVIER y**

**QUERBES, VÉRONIQUE**

74 Agente/Representante:

**CALLE LÓPEZ, Alejandro**

ES 2 986 208 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**

Toma de impresión óptica de la arcada dental de un paciente

5 **[Campo técnico]**

La presente invención se refiere en general a la imagenología tridimensionales (3D) y, en concreto, a técnicas de toma de una impresión óptica de una superficie bucodental como una arcada dental de un paciente, es decir, a técnicas que permiten generar una reconstrucción tridimensional de una arcada dental de un paciente.

10 Los modos de realización aprovechan los avances realizados en arquitecturas de aprendizaje profundo tales como las redes neuronales profundas y las redes de neuronas (o redes neuronales) convolucionales o incluso CNN (del inglés "*Convolutional Neural Networks*") para realizar aplicaciones de imagenología tridimensional a partir de imágenes bidimensionales (2D) adquiridas mediante un dispositivo de adquisición que comprende uno o más sensores de imagen 2D.

15 La invención encuentra usos en diversas aplicaciones de imagenología bucodental, por ejemplo para reconstrucción dental, creación de implantes dentales, cirugía dental con fines estéticos, etc. Estos ejemplos no son limitativos.

20 **[Estado de la técnica anterior]**

La impresión óptica es una técnica de adquisición más eficiente que las impresiones convencionales, físico-químicas, independientemente del material utilizado. Los pacientes prefieren la técnica de impresión digital a las técnicas convencionales, en concreto debido a la ausencia de un reflejo nauseoso inducido. Esta técnica es menos invasiva porque no requiere preparación, ni limpieza o desinfección, y se elimina el riesgo de contaminación infecciosa relacionado con la realización de una impresión tradicional. Es más eficaz y ergonómica, la presencia de una pantalla que permite al facultativo prescindir de la visión directa en la boca. También es más rápida porque la adquisición de datos es relativamente corta y es posible repetir nuevas capturas durante la sesión. Archivar impresiones ópticas es menos engorroso y este tipo de impresión ofrece al facultativo la posibilidad de almacenar indefinidamente la información capturada. Por último, es posible mejorar la transferencia de imágenes digitales entre el consultorio del odontólogo y el laboratorio de fabricación de prótesis, llegado el caso.

35 De una manera general, la impresión óptica es la primera etapa de una cadena digital desmaterializada ("*digital workflow*") que, en inglés, significa "flujo de trabajo digital") que abarca desde la fase inicial hasta la colocación clínica de una prótesis dental, por ejemplo. Realizada usando una cámara asociada con un software, es una etapa esencial del diseño y la fabricación asistidos por ordenador (CAD/CAM), lo que permite limitar las imprecisiones de las prótesis dentales. Para la toma de una impresión, se utiliza una cámara asociada con un software, capaz de producir varios miles de imágenes 3D por segundo.

40 Aunque existen muchos métodos de grabación, hoy en día CAD/CAM utiliza principalmente el método directo por triangulación en luz estructurada, por una parte, y el método indirecto mediante luz estructurada activa y/o focalizada, por otra parte.

45 El principio de triangulación en luz estructurada consiste en colocar un emisor de luz y un receptor (es decir, una cámara), de tal manera que el objeto a grabar (por ejemplo un diente) esté en el vértice del triángulo formado por el emisor, el receptor y el diente. La cámara receptora mide la proyección de un punto de luz sobre el diente. Para obtener una mayor precisión y una adquisición más rápida y de mejor calidad, se proyecta una franja de luces paralelas (y no sólo un rayo de luz). El objeto devuelve los rayos de luz reflejados al receptor. Sin embargo, este método de grabación presenta inconvenientes prácticos relacionados con el posicionamiento del emisor y del receptor en la boca.

50 El método indirecto que utiliza luz activa o estructurada consiste en proyectar una red periódica sobre el objeto y observarlo a través de una segunda red. Se parece al método por triangulación, pero difiere de este desde un punto de vista matemático. En efecto, la información se percibe gracias a la interferencia de dos redes ondulatorias o macroscópicas que se proyectan y observan, respectivamente, sobre el objeto.

55 Además, la adquisición óptica de volúmenes bucodentales como las arcadas dentales se puede realizar con ayuda de un escáner intraoral o IOS (del inglés "*intraoral scanner*"). Más particularmente, la adquisición se puede realizar de dos maneras: imagen por imagen, o por un flujo digital continuo de imágenes. En los dos casos, un mecanismo de reconstrucción 3D permite reconstruir la superficie 3D de la arcada a partir de imágenes 2D adquiridas por el IOS.

60 En el primer caso (procedimiento imagen por imagen), la grabación se realiza mediante una sucesión de tomas de imágenes en primer plano. A continuación, el software coloca las imágenes de un extremo a otro gracias a partes comunes que existen entre ellas. Las imágenes que no son utilizables se rechazan en tiempo real. Al final de la adquisición, el software indica las áreas de carencia para permitir al facultativo tomar su impresión al nivel de esta carencia. El facultativo coloca la cámara encima de la zona a grabar y es con la ayuda de una señal sonora o visual que se indica la toma de la instantánea. Durante la grabación, la cámara debe ser muy estable. Para ello existen

dispositivos de estabilización que son de gran utilidad, en concreto para facultativos inexpertos. Esta toma es más adecuada para impresiones sectoriales, es decir que se refieren sólo a una porción limitada de la arcada dental, porque implican pocas imágenes en comparación con las impresiones extendidas, es decir, que se refieren a toda la arcada dental o a una parte sustancial de ella.

5 El caso del flujo continuo o flujo de vídeo, por su parte, permite grabar los volúmenes bucodentales al pasar sobre las superficies. Se puede grabar una gran cantidad de imágenes. El flujo continuo se puede utilizar para impresiones extendidas, como arcadas dentales completas. Durante la adquisición, el software muestra directamente al facultativo, en tiempo real, un retorno de vídeo que le permite visualizar la posición de la cámara y la construcción del modelo virtual en curso. Si aparece un área de carencia, el facultativo sólo necesita volver a esta última para compensar el agujero negro. Esta toma permite libertad de movimiento, pero se debe respetar una trayectoria rigurosa ("*scan path*", que significa "ruta de escaneo" en inglés).

15 No obstante, utilizar un IOS no es necesariamente fácil, en particular porque el movimiento del facultativo está limitado por el "*tracking*", es decir, la ruta de escaneo. Respetar esta ruta es una necesidad para tener una arcada coherente. Esto implica que el movimiento debe seguir una trayectoria continua e ininterrumpida. Por ejemplo, si se parte del molar posterior de la derecha, no se puede ir a grabar el molar posterior a la izquierda de un salto. Al contrario, se debe llegar a este por una ruta continua, es decir, mover la cámara a lo largo de toda la arcada y de forma no discontinua: premolares derechos, a continuación canino derecho, incisivos, canino izquierdo, premolares izquierdos y finalmente molares izquierdos. Este es el principal fallo de una adquisición por parte de un IOS. Los facultativos deben entrenarse para realizar rutas continuas. Si el movimiento se tambalea hacia un lado o va demasiado rápido, el facultativo pierde el "*tracking*" y debe regresar para encontrar un área de la arcada ya reconstruida, esperar a que la cámara la reconozca, y a continuación comenzar desde allí en la dirección inicial.

25 Existen, además de los IOS, otros dispositivos de toma de imágenes bidimensionales, que podrían usarse ventajosamente para combinar con los IOS.

30 Entre estos dispositivos, se pueden contar teléfonos móviles, tabletas, gafas de realidad aumentada y cámaras digitales. Estos dispositivos, aunque menos precisos que los IOS, tienen la ventaja de un amplio campo de visión, que a menudo abarca por sí solo el campo de una arcada dental. Ciertamente, no son directamente adecuados para reconstruir tejido dental debido a su especularidad y falta de textura. Se presentará un algoritmo que permite transformar cada una de sus imágenes bidimensionales en una reconstrucción tridimensional burda pero densa (es decir, con una gran cantidad de puntos 3D).

35 La invención parte de la observación de que, combinando el amplio campo de un dispositivo de este tipo, al que llamaremos dispositivo global en lo sucesivo, con la precisión de un IOS, al que se llamará dispositivo local en lo sucesivo, se podría agilizar el proceso clínico de adquisición de una impresión dental extensa.

40 El documento US2013330684 divulga un sistema que comprende una cámara intraoral, así como un conjunto de otros sensores/emisores en el cuerpo de la cámara y/o en un anexo de hardware al que se puede fijar la cámara. Entre estos otros sensores, se proporciona una cámara 2D para tomar fotografías del rostro del paciente y eventualmente realizar una reconstrucción 3D del mismo. Es posible un ajuste 3D entre el modelo 3D de la arcada dental del paciente reconstruido por la cámara intraoral y el modelo 3D del rostro obtenido a partir de fotografías tomadas por la cámara 2D.

45 El documento US2008305458A divulga un método que comprende la generación de una imagen intraoral tridimensional de la corona de un diente o las coronas de una pluralidad de dientes seleccionados, la generación de una imagen tridimensional de la raíz del diente o de las raíces correspondientes de los dientes seleccionados correspondientes, y la composición de las imágenes tridimensionales de la(s) corona(s) del diente y la(s) raíz(es) del diente para producir una imagen compuesta. La imagen de las raíces de los dientes se produce mediante imagenología de rayos X o imagenología por resonancia magnética (IRM).

50 El documento US2018/168781 describe un sistema de realidad aumentada para odontólogos. El sistema comprende un escáner, una pantalla de realidad aumentada y un escáner. Un objetivo del sistema es en concreto la visualización en realidad aumentada de información en una vista en tiempo real tomada por el facultativo. D1 describe la visualización de una imagen 2D sobre una reconstrucción 3D ya establecida.

60 Por último, el documento FR3032282A1 divulga la generación, a través de gafas de realidad aumentada, de una reconstrucción 3D a partir de dos o más imágenes. Esta reconstrucción 3D puede especificarse localmente gracias a la información proporcionada por un escáner intraoral (IOS) más preciso, con fusión de la reconstrucción 3D y esta información mediante ajuste 3D.

**[Descripción de la invención]**

65 La invención propone un procedimiento y un dispositivo capaces de paliar los inconvenientes de la técnica anterior mencionados anteriormente que resultan de la necesidad de garantizar el "*tracking*" al tomar una impresión extendida

de la arcada dental de un paciente, es decir, la generación de una reconstrucción tridimensional de la arcada dental de un paciente.

En efecto, la invención aborda estos inconvenientes proponiendo una solución al problema del "*tracking*", con un primer aspecto de la invención que se refiere a un procedimiento de generación de una reconstrucción tridimensional de una arcada dental de un paciente, llamada impresión óptica extendida, que comprende las siguientes etapas implementadas bajo el control de un facultativo:

- la obtención de al menos una imagen bidimensional global de la arcada dental adquirida con un primer sensor de imagen que tiene un campo de visión relativamente amplio y una precisión relativamente baja, y la generación, mediante primeros medios de procesamiento de datos digitales, de una reconstrucción tridimensional global, relativamente burda, de la arcada dental a partir de dicha imagen bidimensional;
- la adquisición de una pluralidad de imágenes bidimensionales locales de sectores asociados respectivos de la arcada dental con un segundo sensor de imagen que tiene un campo de visión relativamente más restringido y una precisión relativamente más alta que el primer sensor de imagen, y la generación, mediante segundos medios de procesamiento de datos digitales, de reconstrucciones tridimensionales locales respectivas, relativamente más precisas que la reconstrucción tridimensional global, de dichos sectores asociados de la arcada dental;
- la fusión, mediante un algoritmo de fusión implementado por terceros medios de procesamiento de datos digitales, de la reconstrucción tridimensional global y de las reconstrucciones tridimensionales locales en una única y misma reconstrucción tridimensional total en la que las reconstrucciones tridimensionales locales se sustituyen respectivamente por porciones de la reconstrucción tridimensional global que corresponden a los sectores asociados de la arcada dental, desempeñando la reconstrucción global un papel de guía espacial absoluta, comprendiendo dicha fusión un ajuste de las reconstrucciones tridimensionales locales en la reconstrucción tridimensional global, estando adaptado dicho ajuste para minimizar una distancia entre las nubes de puntos correspondientes a cada uno de los sectores asociados de la arcada dental y la nube de puntos correspondientes de la reconstrucción tridimensional global.

Combinando el amplio campo del primer sensor de imagen (que puede ser un smartphone o similar, y que en los sucesivos se denominará aparato o dispositivo global), con la precisión del segundo sensor de imagen (que puede ser un IOS o similar, y que en lo sucesivo se denominará aparato o dispositivo local), se mejora y simplifica el proceso clínico de toma de una impresión óptica extendida de una arcada dental de un paciente. En efecto, el facultativo puede, gracias al aparato global, realizar una reconstrucción global tridimensional en una sola imagen, a continuación con el IOS reconstruir sólo las zonas (o sectores) de la arcada dental que requieren mayor precisión, produciendo una o más reconstrucciones locales precisas. Un algoritmo adicional se encarga de la fusión de la reconstrucción global con cada una de las reconstrucciones locales, para producir una única reconstrucción tridimensional mixta que corresponda a la impresión óptica deseada.

Si el facultativo desea reconstruir con precisión zonas locales disociadas, es decir sin superposición alguna entre ellas, el proceso de fusión se lo permite: en efecto, la reconstrucción global juega el papel de guía espacial absoluta para la colocación respectiva de cada una de las reconstrucciones locales, las cuales se encuentran en la reconstrucción global por ajuste en el mismo marco de referencia espacial.

Por otra parte, el problema de la pérdida de "*tracking*" se resuelve: en efecto, si se parte del principio de que, según modos de implementación de la invención que se explicarán, cada imagen procedente del IOS puede producir una densa reconstrucción 3D, entonces cada una de estas imágenes puede conducir a una reconstrucción local, que se puede realinear sobre la reconstrucción global, y que luego se suma a todas las demás reconstrucciones locales realizadas previamente pero sin necesidad de superposición entre las zonas correspondientes de la arcada dental para asegurar la continuidad, y sin necesidad de seguir una ruta de trazado impuesta. Según los modos de implementación de la invención, la continuidad de la impresión está asegurada por la reconstrucción global subyacente, cualquiera que sea la relación, e incluso en ausencia de relación, entre las reconstrucciones locales de sectores de la arcada dental respectivamente asociados con estas reconstrucciones locales.

El experto en la materia apreciará que tampoco es necesario realizar reconstrucciones globales y locales al mismo tiempo, ni en el mismo lugar. Así, la reconstrucción global puede ser generada por el propio paciente por medio de un aparato global como su teléfono inteligente y transmitida de forma remota (por correo electrónico o por red celular 3G-LTE, 4G o 5G, etc.) a su facultativo para obtener una primera opinión (diagnóstico clínico, estética, etc.). Si es necesario, el facultativo puede ofrecer, a continuación, a su paciente una cita en el consultorio para completar la reconstrucción global gracias a otro aparato local.

Se recuerda que, en el contexto de las técnicas de aprendizaje profundo utilizadas en el campo de la inteligencia artificial (o IA, del inglés "*Artificial Intelligence*"), los modelos algorítmicos de aprendizaje supervisado requieren conocimientos específicos del problema estudiado, mientras que los modelos algorítmicos de aprendizaje no supervisado no los requieren. El aprendizaje de la red neuronal convolucional según los modos de realización de la invención se denomina "semisupervisado" en el sentido de que la CNN está diseñada y entrenada para una aplicación particular.

Ventajosamente, por otra parte, la invención puede implementar estereoscopia de múltiples vistas para generar reconstrucciones 3D globales o locales más precisas a partir, cada una, de una pluralidad de imágenes 2D del mismo objeto tomadas desde diferentes ángulos de visión respectivos mediante un dispositivo binocular de adquisición de imágenes. Pero también puede solucionarlo funcionando, llegado el caso, con un único sensor de imagen 2D, es decir con un dispositivo de adquisición de imágenes 2D (o generador de imágenes 2D) que es monocular y que sirve para adquirir una única imagen 2D para cada reconstrucción 3D a generar.

El procedimiento según el primer aspecto de la invención puede comprender, además, las siguientes características tomadas individualmente o en combinación:

- el primer sensor de imagen puede ser un sensor de imagen de un aparato digital portátil del paciente, como un teléfono móvil, y la imagen bidimensional global puede ser adquirida por el propio paciente con ayuda de su aparato digital portátil y ser transmitida desde dicho aparato digital portátil al facultativo para la generación de la reconstrucción tridimensional global;
- el procedimiento puede comprender la adquisición de la imagen bidimensional global por del facultativo con ayuda del primer sensor de imagen;
- el primer sensor de imagen puede estar comprendido en unas gafas de realidad aumentada que comprenden además los primeros medios de procesamiento digitales para la generación de la reconstrucción tridimensional global;
- el segundo sensor de imagen puede ser un escáner intraoral, que comprende los segundos medios de procesamiento de datos digitales para la generación de las reconstrucciones tridimensionales locales;
- la reconstrucción tridimensional global y/o las reconstrucciones tridimensionales locales pueden efectuarse generando para cada una un mapa de profundidad a partir de una única imagen 2D adquirida por el primer sensor o por el segundo sensor de imagen, respectivamente, sobre la base de una arquitectura de aprendizaje profundo previamente entrenada para tal fin.

Además, un segundo aspecto de la invención se refiere a un programa informático que comprende instrucciones que, cuando el programa informático se carga en la memoria de un ordenador y es ejecutado por un procesador de dicho ordenador, causan la implementación por el ordenador de todas las etapas de la fase de predicción del procedimiento según el primer aspecto.

Un tercer aspecto de la invención se refiere a un dispositivo de generación de una reconstrucción tridimensional de una arcada dental de un paciente, llamada impresión óptica extendida, que comprende:

- un primer sensor de imagen que tiene un campo de visión relativamente amplio y una precisión relativamente baja para adquirir al menos una imagen bidimensional global de la arcada dental, y primeros medios de procesamiento de datos digitales para generar una reconstrucción tridimensional global, relativamente burda, de la arcada dental a partir de dicha imagen bidimensional;
- un segundo sensor de imagen que tiene un campo de visión relativamente más restringido y una precisión relativamente más alta que el primer sensor de imágenes, para adquirir una pluralidad de imágenes bidimensionales locales de sectores asociados respectivos de la arcada dental, y segundos medios de procesamiento de datos digitales para generar reconstrucciones tridimensionales locales respectivas, relativamente más precisas que la reconstrucción tridimensional global, de dichos sectores asociados de la arcada dental; y,
- terceros medios de procesamiento de datos digitales configurados para implementar un algoritmo de fusión adaptado para fusionar la reconstrucción tridimensional global y las reconstrucciones tridimensionales locales en una única y misma reconstrucción tridimensional total en la que las reconstrucciones tridimensionales locales se sustituyen respectivamente por porciones de la reconstrucción tridimensional global que corresponden a los sectores asociados de la arcada dental, desempeñando la reconstrucción global un papel de guía espacial absoluta, comprendiendo el algoritmo de fusión un ajuste de las reconstrucciones tridimensionales locales en la reconstrucción tridimensional global, estando adaptado dicho ajuste para minimizar una distancia entre las nubes de puntos correspondientes a cada uno de los sectores asociados de la arcada dental y la nube de puntos correspondientes de la reconstrucción tridimensional global.

El dispositivo según el tercer aspecto de la invención puede comprender, además, las siguientes características tomadas individualmente o en combinación:

- el dispositivo puede comprender gafas de realidad aumentada que comprenden el primer sensor de imagen y, además, los primeros medios de procesamiento digitales para la generación de la reconstrucción tridimensional global;
- el dispositivo puede comprender un escáner intraoral que comprende el segundo sensor de imagen y, además, los segundos medios de procesamiento de datos digitales para la generación de las reconstrucciones tridimensionales locales
- los primeros medios de procesamiento digitales y/o los segundos medios de procesamiento digitales pueden estar configurados para generar la reconstrucción tridimensional global y las reconstrucciones tridimensionales locales, respectivamente, generando para cada uno un mapa de profundidad a partir de una única imagen 2D adquirida

por el primer sensor o por el segundo sensor de imagen, respectivamente, sobre la base de una arquitectura de aprendizaje profundo previamente entrenada para tal fin.

5 Por último, un cuarto y último aspecto de la invención se refiere a un aparato de toma de una impresión óptica que comprende un ordenador que comprende los terceros medios de procesamiento de datos digitales del dispositivo según el tercer aspecto anterior, así como medios adecuados para la implementación del procedimiento según el primer aspecto.

10 **[Descripción de los dibujos]**

Otras características y ventajas de la invención aparecerán también a la lectura de la descripción que va a seguir. Esta es puramente ilustrativa y debe leerse respecto a los dibujos adjuntos, en los que:

- 15 [Figura 1] La figura 1 es una representación del rostro de un paciente humano cuya arcada dental puede ser objeto de una impresión óptica 3D gracias al procedimiento según la invención;
- [Figura 2] La figura 2 es una imagen 2D de una imagen 2D de la arcada dental global del paciente de la figura 1, obtenida con un aparato de adquisición de imágenes de campo amplio y baja precisión;
- 20 [Figura 3] La figura 3 es una impresión óptica 3D poco precisa obtenida mediante reconstrucción tridimensional a partir de la imagen 2D de la figura 2, según modos de implementación de la invención;
- [Figura 4] La figura 4 es una serie de imágenes 3D correspondientes a reconstrucciones 3D locales precisas de sectores determinados de la arcada dental del paciente de la figura 1, generadas a partir de una pluralidad de imágenes 2D asociadas respectivas, cada una tomada por un aparato de adquisición de imágenes de campo estrecho y alta precisión;
- 25 [Figura 5] La figura 5 es una impresión óptica 3D obtenida mediante la implementación del procedimiento según la invención, mediante ajuste y fusión de las reconstrucciones 3D locales precisas de la figura 4 con la reconstrucción 3D global poco precisa de la figura 3;
- 30 [Figura 6] La figura 6 es un diagrama funcional que ilustra la implementación del procedimiento según el primer aspecto de la invención, por medio de un dispositivo de toma de una impresión óptica según el tercer aspecto de la invención;
- 35 [Figura 7] La figura 7 es un diagrama funcional que ilustra el entrenamiento de una de las CNN que se utilizan mediante el procedimiento de la figura 6; y,
- 40 [Figura 8] La figura 8 es un diagrama de bloques que ilustra una arquitectura CNN que puede utilizarse en modos de realización de la invención.

**[Descripción de modos de realización]**

45 En la descripción de modos de realización que sigue y en las figuras de los dibujos adjuntos, elementos iguales o similares llevan las mismas referencias numéricas en los dibujos. A menos que esto proporcione información adicional útil para la claridad de la presentación en su conjunto, elementos que aparecen varias veces en diferentes figuras a las que se hará referencia sucesivamente, no se describirán en detalle cada vez, para no sobrecargar innecesariamente la descripción.

50 Se describirán modos de implementación de la invención en el contexto no limitativo de la adquisición de una impresión dental óptica extendida, por un odontólogo, por ejemplo.

Antes de comenzar la descripción de modos de realización detallados, parece útil, además, especificar la definición de ciertas expresiones o términos que se emplearán en ellos. A no ser que se acuerde de otra manera, estas definiciones se aplican sin perjuicio de otras definiciones que el experto en la materia pueda encontrar en determinadas obras de literatura especializada.

60 Por "arcada dental de un paciente", se entiende los 32 dientes que forman la dentadura de un ser humano, es decir todos los tejidos dentales estáticos incluida la encía, los dientes y material protésico eventual (coronas, anillos, "scan bodies", "inlay-core", "inlay-onlay", etc.). Llegado el caso, según el contexto, en ocasiones se distinguirá entre la arcada superior que corresponde a la mitad de los dientes que están implantados en la mandíbula superior del individuo, por una parte, y la arcada inferior que corresponde a la otra mitad de los dientes que están implantados en la mandíbula inferior de dicho individuo, por otra parte.

65 Por impresión dental "óptica" de un paciente, se entiende una impresión de los dientes del paciente en forma de imagen, en este caso una imagen tridimensional (3D), a diferencia de una impresión física convencional obtenida por

moldeo de las arcadas dentales superior e inferior. Además, una impresión óptica "extendida" de una arcada dental es una representación en forma de imagen 3D de toda la arcada, o de una parte sustancial de la arcada, es decir, una parte de la arcada correspondiente a al menos una semiarcada, es decir, la media arcada vista desde el lado izquierdo o derecho del paciente. Dicha arcada extendida se opone a una impresión de un solo sector de dichas arcadas, limitada a un diente o sólo a unos pocos dientes.

Una "imagen", o "vista", o incluso "escaneo" ("scan" en inglés), está constituido por un conjunto de puntos de la escena tridimensional real. Para una imagen 2D adquirida por un dispositivo de adquisición de imágenes o un dispositivo generador de imágenes (por ejemplo, un sensor CCD o un sensor CMOS), los puntos en cuestión son los puntos de la escena real proyectados en el plano focal del sensor 2D que sirve para adquirir la imagen 2D y están definidos por los píxeles de la imagen 2D. Para una imagen 3D reconstruida (también llamada "reconstrucción 3D", este término designa el producto o resultado del procesamiento de reconstrucción 3D), los puntos en cuestión son una nube de puntos 3D obtenida mediante una transformación de un "mapa de profundidad" (véase la definición a continuación). Una nube de puntos de este tipo define un esqueleto de la escena tridimensional. Y una malla 3D de esta nube de puntos, por ejemplo una malla de puntos 3D triangulados, puede definir una envoltura de la misma.

Un " $n$ -plete" de imágenes, donde  $n$  es un número entero mayor o igual a la unidad, es un conjunto de  $n$  imágenes tomadas simultáneamente por el o los sensores de un dispositivo de adquisición de imágenes (o dispositivo de imagenología). En general, el dispositivo de imagenología tiene dos sensores, que permiten generar un doblete (es decir, una 2-plete) de imágenes de la misma escena adquiridas respectivamente simultáneamente por cada uno de estos dos sensores. Pero puede tener tres, cuatro, cinco, etc., sensores que permiten adquirir un triplete (es decir, un 3-plete), un cuadrplete (es decir, un 4-plete), un quintplete (es decir, un 5-plete), etc., de imágenes de la misma escena adquiridas respectivamente simultáneamente por cada uno de dichos sensores. Sin embargo, el dispositivo también puede comprender solamente un único sensor asociado a una disposición óptica basada en espejo(s) y eventualmente prisma(s), que permite adquirir sucesivamente varias imágenes de la misma escena vista desde diferentes ángulos.

Un dispositivo de imagenología "monocular" es un dispositivo de adquisición que comprende solamente un único sensor de imagen.

El "ajuste" (en inglés "*registration*") consiste en determinar la relación espacial entre dos representaciones (imagen 2D o superficie 3D) de un mismo objeto de manera que se superpongan las representaciones de un mismo punto físico.

El "cálculo de colocación" es la estimación de la posición y orientación de la escena fotografiada en relación con el generador de imágenes (sensor de imagen). Este es uno de los problemas fundamentales de la visión por ordenador, llamada a menudo "*Perspective-n-Points*" (PnP). Este problema consiste en estimar la colocación (2-plete [ $R_i$ ;  $t_j$ ] formada a partir de la matriz de rotación  $R_i$  y el vector de traducción  $t_j$ ) de la cámara en relación con un objeto en la escena, lo que equivale a encontrar la colocación que permite reducir el error de reproyección entre un punto en el espacio y su equivalente 2D en la imagen. Un enfoque más reciente, llamado ePNP (del inglés "*Efficient Perspective-n-Point*"), supone que la cámara está calibrada y toma la opción de superar los problemas de calibración normalizando los puntos 2D multiplicándolos por la inversa de la matriz intrínseca. Este enfoque añade a esto el hecho de configurar la colocación de la cámara pasando por 4 puntos de control, asegurando que la transformación estimada sea rígida. Proceder de este modo permite acortar los tiempos de cálculo.

Un "mapa de profundidad" ("*Depth map*" en inglés) asociado a una imagen 2D, es una forma de representación 2D de información 3D reconstruida, correspondiente a la porción de la escena 3D reproyectada en la imagen 2D. En la práctica, es un conjunto de valores, codificados en forma de niveles (o tonos) de gris, respectivamente asociados con cada píxel  $p_i$  de la imagen 2D: cuanto mayor es la distancia entre el punto de la escena tridimensional y el plano de la imagen 2D, más oscuro es el píxel.

Una "red de neuronas (o neuronal) convolucional" o "red de neuronas (o neuronal) de convolución" o incluso CNN (del inglés "*Convolutional Neural Networks*"), es un tipo de red neuronal artificial acíclica ("*feed-forward*"), que consiste en una pila multicapa de perceptrones, cuyo objetivo es preprocesar pequeñas cantidades de información. Una CNN se compone de dos tipos de neuronas artificiales, dispuestas en "estratos" o "capas" que procesan sucesivamente la información:

- las neuronas de procesamiento, que procesan una porción limitada de la imagen (llamada "campo receptivo") a través de una función de convolución; y,
- las neuronas de puesta en común (total o parcial) de las salidas, llamadas neuronas de "*pooling*" (que significa "agrupación" o "puesta en común", en inglés), que permiten comprimir la información reduciendo el tamaño de la imagen intermedia (a menudo mediante submuestreo).

Todas las salidas de una capa de procesamiento permiten reconstruir una imagen intermedia, que sirve de base para la siguiente capa. Se puede aplicar un procesamiento correctivo no lineal y puntual entre cada capa para mejorar la relevancia del resultado. Actualmente, las CNN tienen amplias aplicaciones en el campo del reconocimiento de imágenes.

La **figura 1** muestra el rostro de un paciente, más particularmente una mujer joven 10 sonriendo. Como se muestra en el detalle 15 de la figura, que corresponde a la zona de la boca del paciente, la boca se abre cuando el paciente sonríe. Esto revela al menos parte de su dentadura. La zona de la boca comprende dientes, en este caso la arcada dental superior 11, de la encía 12, el labio superior 13 y el labio inferior 14. Más generalmente, una arcada dental del paciente 10 (o ambas arcadas, superior e inferior) es(son) visible(s) cuando el paciente abre la boca. Esta arcada dental puede ser objeto de una toma de impresión óptica 3D gracias al procedimiento según la invención.

La **figura 2** muestra una imagen 2D de toda la arcada dental del paciente en la figura 1 (arcada superior y arcada inferior incluidas), obtenida con un aparato de adquisición de imágenes de campo amplio y baja precisión (o "aparato global" en el contexto de la presente descripción).

La **figura 3** es una impresión óptica 3D poco precisa obtenida mediante reconstrucción tridimensional a partir de la imagen 2D de la figura 2, según modos de implementación de la invención que se describirán más adelante.

La **figura 4** es una serie de superficies 3D correspondientes a reconstrucciones 3D locales precisas de sectores determinados de la arcada dental del paciente de la figura 1, generadas a partir de una pluralidad de imágenes 2D asociadas respectivas, cada una tomada por un aparato de adquisición de imágenes, llamado aparato local, que tiene un campo relativamente más estrecho pero que es relativamente más preciso que el aparato global que ha servido para adquirir la imagen 2D de la figura 2. Por ejemplo, la reconstrucción 41 corresponde al sector del canino superior derecho del paciente; la reconstrucción 42 corresponde al sector del canino superior izquierdo del paciente; y la reconstrucción 43 corresponde al sector de los dos incisivos inferiores del paciente.

Por último, la **figura 5** muestra una impresión óptica 3D extendida de la arcada del paciente, que también se llama reconstrucción 3D especificada 50, que se obtuvo mediante la implementación del procedimiento según la invención. Más particularmente, la mejora de la precisión de la reconstrucción 3D especificada 50 en comparación con la reconstrucción 3D global 30, proviene del ajuste y fusión de las reconstrucciones 3D locales precisas 41, 42 y 43 de la figura 4 con la reconstrucción 3D global poco precisa 30 de la figura 3. En la reconstrucción 3D especificada de la figura 5, las reconstrucciones 3D locales 41, 42 y 43 han sido sustituidas por las partes correspondientes de la reconstrucción 3D global 30 de la figura 3. En el ejemplo, las reconstrucciones 3D locales 41 y 42 se colocan en la arcada superior, en su lugar al nivel del canino superior derecho y al nivel del canino superior izquierdo, respectivamente, mientras que la reconstrucción 3D local 43 se coloca en la arcada inferior, en su lugar al nivel de los dos incisivos inferiores.

El experto en la materia apreciará que, si se adquieren imágenes 2D a partir de las cuales se generan reconstrucciones 3D locales correspondientes cada una a una determinada porción de la arcada dental del paciente de manera que, juntas, dichas reconstrucciones 3D locales abarcan la totalidad de dicha arcada, entonces su fusión con la reconstrucción 3D global 30 de la figura 3 permite generar una reconstrucción 3D extendida de la arcada (es decir una reconstrucción 3D que abarca una parte sustancial de la arcada, por ejemplo una semiarcada o incluso la arcada completa) y posiblemente, por supuesto, tanto la arcada superior como la arcada inferior, que presenta la misma precisión que las reconstrucciones locales. Y, sin embargo, es posible que las reconstrucciones 3D locales, como las reconstrucciones 41, 42 y 43 de la figura 4, se hayan adquirido sin ningún orden en particular y sin superposición parcial entre ellas, es decir, en realidad sin ninguna relación particular entre ellas excepto que están, obviamente, todas basado en imágenes 2D de la arcada dental de un mismo paciente que se busca reconstruir en 3D con la mayor precisión posible. Dicho de otra manera, su adquisición con un aparato local como un IOS no necesita respetar las severas limitaciones que están ligadas a la necesidad de respetar una ruta de "*tracking*", que se han mencionado en la introducción de la presente descripción.

La reconstrucción 3D especificada 50 de la figura 5 se puede visualizar en una pantalla, en gafas de realidad virtual o en gafas de realidad aumentada; también puede almacenarse en una memoria masiva o en cualquier otro medio de almacenamiento adecuado; también se puede imprimir en 3D, etc., dependiendo de las necesidades de la aplicación.

Con referencia al diagrama funcional de la **figura 6**, a continuación se describirán tanto un ejemplo de implementación del procedimiento según el primer aspecto de la invención como los medios funcionales de un sistema de toma de una impresión óptica según el tercer aspecto de la invención. En la figura, las etapas del procedimiento están representadas por flechas gruesas. Los demás elementos representados, aparte de las imágenes 2D y aparte de las superficies 3D correspondientes a las reconstrucciones 3D obtenidas a partir de dichas imágenes 2D, son medios del sistema de toma de impresiones. Se comenzará por presentar estos últimos.

El sistema de toma de impresiones comprende un primer aparato 61 de captura de imágenes (o aparato de adquisición de imágenes, o incluso generador de imágenes) denominado aparato global en la presente descripción. Por construcción, este aparato global 61 tiene un campo de visión relativamente amplio pero una precisión relativamente baja. Puede ser una cámara fotográfica o una cámara, que puede integrarse en un dispositivo portátil personal del paciente o facultativo. Este tipo de dispositivo portátil comprende, sin que la lista sea exhaustiva: un teléfono móvil, una tableta digital, gafas de realidad aumentada, una cámara intraoral bidimensional, una cámara fotográfica, etc. Posteriormente nos referiremos a este tipo de aparato con la expresión "aparato global" porque nos permite obtener

una vista global de la arcada.

En un modo de realización, el aparato global 61 se utiliza en modo vídeo, adquiriendo secuencias temporales de imágenes 2D. Con este modo de funcionamiento, un algoritmo tipo SLAM (del inglés "*Simultaneous Localization And Mapping*") se puede aplicar secuencialmente a la sucesión (o serie) de imágenes, para producir una reconstrucción 3D incremental a medida que llegan nuevas imágenes.

En otro modo de realización, el aparato global 61 se utiliza en modo foto, adquiriendo sólo una imagen a la vez. Para producir una reconstrucción 3D densa a partir de esta única imagen adquirida, es posible recurrir a la predicción de mapas de profundidad mediante una sola imagen por medio de la red de aprendizaje profundo 361. Para más detalles sobre este procedimiento, el experto en la materia puede consultar, por ejemplo, el artículo de Laina et al "*Deeper Depth Prediction with Fully Convolutional Residual Networks*", IEEE International Conference on 3D Vision (3DV), 2016.

Para entrenar este tipo de red de aprendizaje, es necesario disponer de una base de datos de entrenamiento que comprenda pares {imagen 2D; mapa de profundidad asociado}. Más adelante se describirá un procedimiento que permite generar una base de datos de este tipo, con referencia al diagrama funcional en la **figura 7**.

Como se muestra en la figura 6, el sistema para la implementación del procedimiento, comprende además un segundo aparato 62 de captura de imágenes, llamado aparato local en la presente descripción. Este aparato local 62 tiene, él también por construcción, un campo de visión relativamente más estrecho pero una precisión relativamente mayor, que el campo de visión y la precisión, respectivamente, del aparato global 61. Puede ser una cámara intraoral destinada a la reconstrucción 3D de superficies intraorales a partir de imágenes 2D pasivas, por ejemplo, un escáner intraoral (IOS).

Una cámara intraoral 62 de este tipo está adaptada y destinada a la reconstrucción 3D de superficies intraorales a partir de imágenes 2D pasivas, proporcionadas por uno o más sensores de imagen. Este tipo de dispositivo produce, en tiempo real, reconstrucciones 3D superficiales de tejidos intraorales. La reconstrucción puede ser texturizada en el caso de que al menos uno de los sensores sea un sensor de imagen en colores, por ejemplo codificado en formato RGB.

En el caso de una cámara intraoral con varios sensores pasivos, la reconstrucción se puede realizar gracias a algoritmos convencionales de estereoscopia de múltiples vistas. Se describe un ejemplo de este tipo de algoritmo, por ejemplo, en el artículo de A. Geiger, J. Ziegler y C. Stiller, "*StereoScan: Dense 3d reconstruction in real-time*" 2011 IEEE Intelligent Vehicles Symposium (IV), Baden-Baden, 2011, págs. 963-968.

En el caso de una cámara intraoral con un solo sensor, es posible consultar el artículo de R. A. Newcombe, S. J. Lovegrove y A. J. Davison, "*DTAM: Dense tracking and mapping in real-time*", 2011 International Conference on Computer Vision, Barcelona, 2011, páginas. 2320-2327, o incluso el artículo de D. Eigen, C. Puhrsch y R. Fergus, "*Depth Map Prediction from a Single Image using a Multi-Scale Deep Network*", NIPS'14, Proceedings of the 27th International Conference on Neural Information Processing Systems - Volumen 2, diciembre de 2014 páginas 2366-2374.

En el caso de varios sensores con una fuente de luz estructurada, la tecnología depende de esta fuente de luz estructurada. El experto en la materia puede consultar las notas técnicas (guías de usuario) de los fabricantes de este tipo de cámaras (lista).

Aunque esto no es imprescindible, cada uno de los aparatos de captura de imágenes 61 y 62 puede ser ventajosamente un generador de imágenes monocular, es decir que comprende únicamente un único sensor de imagen. En realidad, un generador de imágenes monocular cuesta menos que los generadores de imágenes binoculares (o, más generalmente, aparatos estereoscópicos de múltiples vistas) adaptados para generar reconstrucciones 3D mediante estereoscopia de múltiples vistas. Más adelante se verá cómo se pueden generar reconstrucciones tridimensionales, según modos de implementación, a partir de una única imagen bidimensional de la escena a reconstruir. Por supuesto, el uso de generadores de imágenes binoculares o estereoscópicos de múltiples vistas, sin embargo, también es posible en el contexto de implementaciones de la invención.

El generador de imágenes de los aparatos 61 y 62 es, por ejemplo, un sensor pasivo. Puede ser un sensor monocular en tecnología CMOS o en tecnología CCD. Una escena estática puede ser fotografiada mediante un dispositivo de captura de imágenes monocular, en blanco y negro o en color. Sin embargo, el experto en la materia apreciará que, para permitir obtener reconstrucciones 3D texturizadas, cada uno de los aparatos 61 y 62 y en cualquier caso al menos el aparato local 62 que proporciona reconstrucciones locales más precisas, comprende al menos un sensor de imagen en color, por ejemplo en formato RGB. Además, cada uno de los aparatos 61 y 62 puede adquirir imágenes 2D en modo fotografía (una única imagen 2D a la vez) o en modo vídeo (varias imágenes 2D a lo largo del tiempo y en cualquier frecuencia).

En los modos de implementación del procedimiento, el aparato global 61 está adaptado y destinado a producir una

reconstrucción 3D global 30 de los tejidos dentales más rápidamente (debido a su campo de visión ampliado) pero eventualmente menos preciso que el aparato local. El aparato local 62 puede usarse entonces para especificar toda o parte de la reconstrucción global 30, produciendo una o más reconstrucciones locales más precisas como las reconstrucciones 41-43, que se puede fusionar con la reconstrucción global 30, en la localización adecuada. Se describirán más adelante, en un primer momento, modos de implementación de la reconstrucción 3D mediante el aparato global 61 que son previsible para producir una reconstrucción 3D global. A continuación se detallará un modo de implementación para fusionar la reconstrucción global 30 y las reconstrucciones locales 41-43.

De antemano, se continúa con la presentación de los medios que constituyen el sistema de impresión óptica extendida que se muestra en la figura 6.

El sistema comprende una unidad de cálculo 260, por ejemplo, un microprocesador de un ordenador ("*Central Processing Unit*") o un grupo de procesadores, por ejemplo dentro de una tarjeta gráfica (GPU, para "*Graphical Processing Unit*"). La unidad de cálculo está configurada para ejecutar algoritmos de procesamiento de datos digitales codificados en software. En particular, en modos de implementación, la unidad de cálculo 260 está configurada para ejecutar un primer algoritmo 261 y un segundo algoritmo 262 de reconstrucción 3D, así como un algoritmo 263 de fusión de superficies 3D, al que se volverá más adelante.

Por otra parte, el sistema comprende además un sistema 360 de almacenamiento de información digital, por ejemplo un disco duro con tecnología SSD (del inglés "*solid-state drive*") u otro. El sistema de almacenamiento 360 está adaptado para almacenar permanentemente la información digital que define las redes de neuronas convolucionales 361 y 362 (o CNN). Estas redes de neuronas 361 y 362 son utilizadas por los algoritmos 261 y 262, respectivamente, para generar reconstrucciones 3D a partir, cada vez, de al menos una imagen 2D. Las redes de neuronas 361 y 362 están entrenadas para este fin, durante las respectivas fases de aprendizaje a las que también se volverá más adelante.

El experto en la materia apreciará que el sistema de almacenamiento 360 puede incluirse físicamente directamente en el sistema de toma de una impresión óptica, o estar alejado del mismo, por ejemplo en un ordenador o en un servidor de cálculo remoto. En este último caso, el sistema puede acceder al servidor de cálculo por una red de comunicación ad-hoc, por ejemplo por una intranet o Internet, por ejemplo a través de una red Ethernet o una red de comunicaciones por radio inalámbrica como una red 3G-LTE, 4G o 5G, sin que esto afecte a la implementación del procedimiento.

En un modo de realización ilustrado por la **figura 8**, la arquitectura de aprendizaje profundo 361 es una red de neuronas convolucional 80 que puede tener una estructura completamente convencional. Este tipo de CNN está disponible en bibliotecas conocidas por el experto en la materia y de libre acceso. Como entrada, se proporciona una imagen 25 bidimensional en forma de matriz de píxeles. Tiene dos dimensiones para una imagen en escala de grises. Llegado el caso, el color está representado por una tercera dimensión, de profundidad 3 para representar los colores fundamentales [Rojo, Verde, Azul].

La figura 8 muestra un ejemplo particular de CNN, que en realidad es una FCN (del inglés "*Fully Convolutional Network*") inspirado por el artículo de J. Long, E. Shelhamer y T. Darrell, "Fully convolutional networks for semantic segmentation", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, págs. 3431-3440. Esta FCN consta de dos partes muy diferenciadas, según una arquitectura de codificación/decodificación.

La primera parte de la FCN de codificación es la parte convolucional propiamente dicha. Ella comprende, la "capa de procesamiento convolucional" 81, que tiene una sucesión de filtros, o "núcleos de convolución", aplicados en estratos. La capa de procesamiento convolucional 81 funciona como un extractor de características de las imágenes 2D admitidas en la entrada de la CNN. En el ejemplo, la imagen de entrada 25 pasa a través de la sucesión de núcleos de convolución, creando cada vez una nueva imagen llamada mapa de convolución. Cada núcleo de convolución tiene dos capas de convolución 811 y 812, y una capa 813 de reducción de la resolución de la imagen mediante una operación de puesta en común también denominada operación de máximo local ("*maxpooling*", en inglés).

La salida de la parte convolucional 81 se proporciona a continuación como entrada a una última capa de convolución 820 que captura todo el campo de acción visual de la capa anterior y, que imita, de este modo, una capa completamente conectada (capa "*fully connected*" en inglés).

Por último, una capa de deconvolución 830 final produce como salida un mapa de profundidad que, gracias a la implementación de los modos de realización, es un mapa de profundidad a escala.

Cabe señalar que el algoritmo 362 que tiene la función de generar las reconstrucciones 3D locales también puede tener la estructura de la red de neuronas convolucional 80 de la figura 8, o cualquier otra estructura de aprendizaje profundo conocida.

La **figura 7** es un diagrama funcional que ilustra el aprendizaje de una de las redes de neuronas convolucionales 361 y 362 que se utilizan mediante el procedimiento de la figura 6, para generar la reconstrucción 3D global y las

reconstrucciones 3D locales, respectivamente.

Se recuerda que no existe ningún requisito previo en cuanto al número de sensores presentes en el generador de imágenes global 61 (ni en el generador de imágenes local 62, además), este número puede ser igual a la unidad. Dicho de otra manera, los modos de implementación de la invención permiten aplicaciones en las que el aparato global 61 del sistema de adquisición de impresiones sólo consta de un único sensor.

Para el entrenamiento de la red de neuronas 631, se utiliza un aparato de adquisición de imágenes o vídeos 2D comparable al aparato global 61, en el sentido de que tiene características intrínsecas idénticas o similares. Este aparato es, por ejemplo, un teléfono móvil que consta de uno o más sensores pasivos, al menos uno de los cuales es un sensor de imagen en color (si se desea obtener reconstrucciones 3D texturizadas), y cuyo campo de visión es lo suficientemente amplio como para incluirlo en una sola imagen de toda la arcada dental, aunque esto implique que su precisión 2D sea menor.

Considerando un grupo de varias personas, para las cuales previamente se ha obtenido una reconstrucción 3D precisa y texturizada de la arcada dental de cada una de ellas, por ejemplo, utilizando una cámara intraoral destinada a la reconstrucción 3D de la superficie intraoral a partir de imágenes 2D pasivas como se ha descrito anteriormente, con la que se sigue la ruta de "tracking" de acuerdo con el modo de uso de dicho IOS para tomar una impresión óptica de una arcada dental según la técnica anterior. Para cada una de estas personas, con el aparato global mencionado anteriormente también se toman imágenes aproximadas en 2D de la arcada dental desde diferentes ángulos de visión. En cada una de estas imágenes, se anotan manualmente puntos característicos 2D. Estos mismos puntos se anotan manualmente en la reconstrucción 3D de la arcada dental de la persona. Para cada imagen, se obtiene así un conjunto de correspondencias punto 2D/punto 3D, entre las imágenes 2D adquiridas por el aparato global y la superficie 3D obtenida previamente por el IOS.

Gracias a estas correspondencias, y a las propiedades intrínsecas del sensor, se calcula la colocación relativa del modelo 3D con respecto a la imagen 2D en cuestión, por ejemplo con el algoritmo ePNP mencionado anteriormente. Cabe señalar que si no se conocen las propiedades intrínsecas del sensor del aparato global, es posible tener en cuenta propiedades intrínsecas aleatorias, siempre que se mantengan constantes durante todo el proceso de aprendizaje y a continuación cuando se utilice la CNN en la fase de predicción, para la predicción de reconstrucciones 3D globales)

Una vez estimada la colocación, se puede calcular un mapa de profundidad del modelo 3D en la imagen 2D. Se obtiene por lo tanto, para cada persona, pares {imagen 2D; mapa de profundidad asociado} para el aprendizaje de la red de aprendizaje profundo 361. En la figura 7, estos pares llevan las referencias 71, 72, 73, etc. Se almacenan en una base de datos de aprendizaje 70.

La solución propuesta para el aprendizaje de las estructuras de aprendizaje profundo 261 y 262 es una solución para de aprendizaje de la red de neuronas convolucional 80 de manera semisupervisada. La noción de semisupervisión se refiere al hecho de que se ha establecido un procedimiento separado para generar datos de aprendizaje: hay, por lo tanto, una transferencia "manual" de datos de un algoritmo a otro, es decir, el proceso de generación de los pares de aprendizaje almacenados 71, 72, 73, etc., por una parte, proceso de aprendizaje propiamente dicho, por otra parte. Como apreciará el experto en la materia, sin embargo, la generación de datos de aprendizaje puede seguir siendo automática. Por oposición, la no supervisión se refiere al hecho de que la red encontraría sola, mediante el análisis de los datos de entrada, reglas de inferencia que permiten producir mapas de profundidad, lo cual es difícilmente concebible. Y la supervisión total se refiere a dar datos de aprendizaje no procedentes de un proceso automático sino de un proceso completamente manual, por ejemplo, creando mapas de profundidad "a mano" mediante mediciones físicas, que tampoco es el caso aquí.

Como se muestra en la figura 7, el aprendizaje semisupervisado según los modos de realización se realiza a partir de los pares de aprendizaje almacenados en la base de datos de aprendizaje 70. Este aprendizaje se puede implementar mediante una calculadora adecuada, por ejemplo, un procesador gráfico (o GPU, del inglés "Graphical Processing Unit") capaz de realizar grandes cantidades de cálculo.

El aprendizaje de la estructura de aprendizaje profundo 361 se efectúa de una vez por todas, por el fabricante del sistema de toma de impresión óptica 3D. Una vez entrenada, esta red es capaz de predecir un mapa de profundidad para cualquier imagen 2D tomada por el aparato global 61 y que consta de una porción dentada. A continuación se carga en el sistema de almacenamiento de datos 360 de cada aparato vendido, para ser utilizado en la fase de predicción. A continuación puede ser objeto a actualizaciones, llegado el caso.

El experto en la materia apreciará la relevancia de la predicción mediante la red de aprendizaje profundo 361 no se limita solo a las personas implicadas en la fase de aprendizaje. Ella se limita, en cambio, a imágenes 2D que constan de porciones dentadas. Además, la red de aprendizaje profundo 361 solo puede predecir el mapa de profundidad para la porción dentada (es decir, más bien, todos los tejidos dentales estáticos y visibles en la superficie, incluyendo, por lo tanto, la encía, los dientes y material protésico eventual). También se limita al tipo de aparato implicado en el aprendizaje, debido a sus propiedades intrínsecas. Por ejemplo, si el aprendizaje se realizó con gafas de realidad

aumentada de una determinada marca y de una determinada generación (por ejemplo *Google Glass Enterprise™ Edición 2*), la predicción solo será válida para esta marca y para esta generación de gafas (suponiendo que la variabilidad de las propiedades intrínsecas entre todas estas gafas sea baja con respecto a la precisión de la predicción objetivo), pero no para todos los aparatos de esta marca ni para todos los aparatos de esta generación, respectivamente.

En la fase de predicción, el mapa de profundidad predicho se puede convertir ventajosamente en una nube de puntos 3D texturizada, por simple proyección inversa. Esta nube de puntos 3D puede a su vez ser procesada mediante un algoritmo de malla triangular (véase por ejemplo el artículo de Kazhdan et al "Screened Poisson Surface reconstruction", ACM Transactions on Graphics (TOG), Volumen 32 (3), 1 de junio de 2013) para producir un modelo triangulado texturizado.

Para eliminar la ambigüedad vinculada al factor de escala, se puede utilizar la técnica descrita en otra solicitud de patente de los mismos inventores y mismos solicitantes, presentada el mismo día ante la misma oficina de Propiedad Industrial, y con el título "*Génération de modèles 3D à l'échelle à partir d'images 2D produites par un dispositif d'imagerie monoculaire*". Si no, el experto en la materia apreciará que esta ambigüedad no es un bloqueo y puede resolverse en una etapa posterior, durante el ajuste con la reconstrucción local, que se explicará más adelante.

En los párrafos anteriores se describe cómo es posible obtener reconstrucciones 3D en tiempo real desde ambos tipos de aparato, a saber, el aparato global 61 y el aparato local 62.

Se hace referencia nuevamente a la figura 6, esta vez para describir la implementación del procedimiento de toma de una impresión óptica extendida de la arcada dental de un paciente. Las etapas de este procedimiento se implementan bajo el control de un facultativo, es decir, un dentista o cirujano-dentista, y/o su asistente, llegado el caso.

En un primer momento, el procedimiento comprende etapas que permiten al facultativo obtener la reconstrucción 3D global 30. Para esto, en 160, con el aparato global 61 se captura una imagen 2D poco precisa 20 de la arcada dental del paciente. Luego, en 161, se genera la reconstrucción 3D global 30, a partir de la imagen 20, por el algoritmo de reconstrucción 261. Estos medios de procesamiento de datos digitales 261 están adaptados para generar una reconstrucción tridimensional global 30, relativamente burda, de la arcada dental a partir de la imagen bidimensional 20.

En el ejemplo de implementación representado en la figura, el algoritmo 261 es ejecutado por la unidad de cálculo 260 del sistema de toma de impresión 3D extendida. Como ya se dijo, no obstante, esto no es obligatorio. En efecto, cuando el aparato global 61 es un teléfono inteligente, ya sea el teléfono inteligente del paciente o el del facultativo, el algoritmo 261 puede ejecutarse mediante un programa de aplicación instalado en el teléfono inteligente, que por lo tanto es capaz, él mismo, de generar la reconstrucción 3D global poco precisa 30. Esta se transmite a continuación al sistema, mediante cualquiera de los medios de comunicación ya presentados anteriormente, para que el facultativo disponga de ella. El algoritmo de reconstrucción 261 también se puede ejecutar en un servidor de cálculo, como también ya se explicó en lo anterior, y ser transmitido al sistema, que a continuación es capaz de comunicarse con dicho servidor para este fin. Por último, cuando el aparato global 61 es, por ejemplo, un par de gafas de realidad aumentada, el algoritmo 261 se puede ejecutar en dichas gafas directamente, de modo que la reconstrucción 3D global 30 la obtenga el propio facultativo, en su consultorio. A continuación se hace accesible a la unidad de cálculo 260, a la que se pueden conectar las gafas de realidad aumentada, por ejemplo mediante una conexión USB (del inglés "*Universal Serial Bus*"), o mediante Bluetooth™ o mediante una red inalámbrica de corto alcance como una red Wi-Fi o NFC (del inglés "*Near Field Communication*"), por ejemplo.

A continuación, en un segundo momento, el procedimiento comprende la adquisición de una pluralidad de imágenes 2D locales de sectores asociados respectivos de la arcada dental con el aparato local 62. Se recuerda que el aparato local 62 tiene un campo de visión relativamente más restringido y una precisión relativamente más alta que el aparato global 61. Además, el procedimiento comprende la generación, mediante segundos medios de procesamiento de datos digitales que implementan el algoritmo de reconstrucción 362, de reconstrucciones 3D locales respectivas, relativamente más precisas que la reconstrucción 3D global 30, de dichos sectores asociados de la arcada dental.

En el ejemplo mostrado en la figura 6, se adquieren tres imágenes 2D locales 21, 22 y 23, en 161, 162 y 163, respectivamente, con ayuda del aparato local 62, por ejemplo un IOS. Estas adquisiciones son realizadas por el facultativo, en su consultorio. Ventajosamente, no es necesario respetar ninguna relación particular de proximidad (y en concreto de superposición parcial), ni orden de toma. En particular, no es necesario que el facultativo siga ninguna ruta de "*tracking*" para lograr esta serie de adquisiciones. Dicho de otra manera, las imágenes 2D adquiridas se pueden descorrelacionar completamente entre sí, la única restricción obviamente es que todas sean imágenes de la misma arcada dental (es decir, imágenes tomadas de la boca del mismo paciente) que la imagen 2D global 20.

En un tercer momento, por último, el facultativo puede hacer colocar, a continuación, todas o parte de las reconstrucciones 3D locales 41, 42 y 43 que ha obtenido, por encima del lienzo espacial constituido por la reconstrucción 3D global 30. Para esto, el procedimiento implementa la fusión, mediante un algoritmo de fusión 263 implementado por medios de procesamiento de datos digitales, de la reconstrucción tridimensional global y de las

reconstrucciones tridimensionales locales en una única y misma reconstrucción tridimensional total 50, que es una reconstrucción 3D especificada (es decir, hecha más precisa). En la reconstrucción especificada 50, las reconstrucciones 3D locales 41, 42 y 43 se sustituyen respectivamente por porciones de la reconstrucción tridimensional global 30 que corresponden a los sectores asociados de la arcada dental del paciente.

5 Los medios de procesamiento digitales que ejecutan el algoritmo de fusión 263 están configurados para ajustar cada reconstrucción 3D local 41-43 sobre la reconstrucción 3D global 30, es decir, llevar la reconstrucción local al marco de referencia de la reconstrucción global. Cabe señalar que estas dos reconstrucciones solo se diferencian entre sí por una transformación espacial relativa con 7 grados de libertad, es decir, 3 grados de libertad en rotación, 3 grados de libertad en traducción y 1 factor de escala.

15 Una vez que cada reconstrucción 3D local 41-43 está ajustada en la reconstrucción 3D global 30, basta con reemplazar, en la etapa 190, la porción de la reconstrucción global 30 cubierta por cada reconstrucción 3D local 41, 42 y 43 para tener una representación 3D más precisa en las porciones de la reconstrucción 3D global 30 en las cuales se han sustituido las reconstrucciones 3D locales. Se obtiene así, al final, la reconstrucción 3D global especificada 50. Esta reconstrucción 50 es una impresión óptica extendida de la arcada dental del paciente, que presenta la mejor precisión posible, determinada por el del aparato local 62, pero que podría generarse más rápida y fácilmente, sin la obligación para el facultativo de seguir el recorrido de "tracking".

20 A continuación, se describirá un modo de realización del ajuste en dos etapas, donde se estima por primera vez un primer ajuste relativamente burdo, a continuación, se afina mediante un segundo ajuste más fino.

25 En cuanto al ajuste burdo, se supone que las reconstrucciones globales y locales están segmentadas, es decir, cada punto 3D de cada una de estas reconstrucciones puede etiquetarse mediante una clase semántica, en este caso el tipo de dientes (molares, caninos, incisivos, etc., siguiendo la numeración convencional utilizada en odontología). Entonces es posible, para cada reconstrucción 3D, encontrar el centroide de cada clase semántica, es decir, el baricentro de cada tipo de diente. Haciendo coincidir cada vez (es decir, para cada reconstrucción 3D local) el centroide de cada tipo de diente entre las dos reconstrucciones 3D, global y local, entonces es posible encontrar la transformación relativa. El experto en la materia puede considerar, por ejemplo, el uso del algoritmo de Umeyama (ver el artículo de S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns", abril de 1991, IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-9 1991, Vol. 13, n.º 4, págs. 376-380).

35 Una vez obtenido el ajuste burdo, a continuación, se puede efectuar un ajuste más fino, por ejemplo, mediante el algoritmo de ICP (para "Iterative Closest Point" en inglés: véase por ejemplo el artículo de P.J. Besl y N.D. McKay, "A Method for Registration of 3-D Shapes", IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI, Los Alamitos, CA, EE. UU., IEEE Computer Society, vol. 14, n.º 2, febrero de 1992, págs. 239-256).

40 Para el ajuste burdo, se supuso que se sabe segmentar en tiempo real el modelo 3D. En un modo de realización, se puede reemplazar la red de aprendizaje profundo 361 con una red de aprendizaje más evolucionada, que es capaz de predecir para cada imagen 2D el mapa de profundidad así como un mapa de segmentación, es decir, un mapa 2D que da para cada píxel la clase semántica a la que pertenece.

45 Un ejemplo de este tipo de red de aprendizaje más evolucionada se puede encontrar en el artículo "Towards unified depth and semantic prediction from a single image", Wang *et al.*, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, págs. 2800-2809. Así, prediciendo para cada imagen 2D el mapa de segmentación correspondiente, entonces es posible hacer la proyección inversa, para etiquetar cada punto 3D.

50 Es necesario, para producir dicha red de aprendizaje más evolucionada, disponer de datos de entrenamiento. Además de los datos de entrenamiento para entrenar la red de aprendizaje profundo 361, también es necesario disponer, para cada imagen 2D, de mapas de segmentación asociados. Esto se puede hacer anotando manualmente las imágenes (es decir, asignando la clase semántica asociada a cada píxel, lo cual es fácil para el ojo humano). Este procedimiento es relativamente tedioso pero es convencional en la generación de datos de aprendizaje. Es posible sustituirlo por un procedimiento semiautomático, por ejemplo, etiquetando manualmente un modelo 3D y re proyectando la etiqueta en todas las imágenes que hayan servido para reconstruir este modelo 3D, pero la descripción de este mecanismo iría más allá del alcance de la presente descripción.

60 La descripción técnica de modos de implementación y modos de realización que se proporcionó anteriormente ha mostrado cómo cada tipo de aparato puede producir en tiempo real reconstrucciones 3D texturizadas, independientemente del número de sensores (que se pueden reducir a uno) y del número de imágenes (dependiendo de si el aparato se utiliza en modo foto o en modo vídeo). También se explicó cómo se pueden combinar estas reconstrucciones 3D para complementarse entre sí: la reconstrucción global, rápida de estimar pero menos precisa, puede mejorarse mediante reconstrucciones 3D locales efectuadas por un aparato local mientras sirve como estructura de anclaje para la colocación de dichas reconstrucciones 3D locales.

65 A continuación se describirán algunos casos de uso prácticos, o casos de uso ("use cases" en inglés), en los que dicha

combinación de un aparato global y un aparato local puede resultar ventajosa.

En un modo de realización, el aparato global se puede confundir con el aparato local, es decir, una cámara intraoral, proviniendo la diferencia de las condiciones de uso para la generación de la reconstrucción 3D global y para la de las reconstrucciones 3D locales, respectivamente. El uso como aparato global se realiza, por ejemplo, alejando la cámara de la escena en comparación con el uso como aparato local, para tener una vista de la misma global (toda la arcada dental) o semiglobal (semiarcada). La imagen procedente del sensor local como un IOS a esa distancia suele ser imprecisa, pero, sin embargo, puede resultar utilizable para obtener una superficie 3D burda, sirviendo de reconstrucción 3D global según los modos de implementación de la invención.

En un modo de realización, el aparato global 61 y el aparato local 62 pueden ser un único y mismo aparato en el sentido de que, por ejemplo, la cámara intraoral 62 puede diseñarse para integrar, por una parte, uno o más sensores locales que tienen alta precisión pero un campo estrecho (típicamente obteniendo imágenes de un área de 3 cm x 2 cm) y, por otra parte, uno o más sensores globales que tienen una precisión relativamente menor pero un campo relativamente mayor (típicamente obteniendo imágenes de una zona que comprende entre una semiarcada y una arcada completa). El facultativo efectúa una primera reconstrucción 3D global con ayuda de sensores globales, a continuación cambia a sensores locales para afinar localmente la reconstrucción 3D global con reconstrucciones 3D locales.

En modos de realización, el aparato global 61 puede ser un teléfono móvil equipado con un sensor 2D pasivo o activo, para integrarse en un enfoque de telemedicina. El sensor de imagen de un teléfono móvil es un sensor de campo amplio, pero que tiene una precisión relativamente baja en comparación con los estándares de imagenología en odontología. En este caso, la imagen 2D puede ser tomada en casa por el propio paciente, y esta imagen 2D o una reconstrucción 3D que se genera sobre la base de dicha imagen 2D, solamente podrá ser obtenida posteriormente por el facultativo en su consultorio dental. Con ayuda de este dispositivo, un paciente puede, de hecho, por sí mismo, solo o con la ayuda de un ser querido, efectuar tomas de imágenes 2D de su arcada dental. Estas imágenes 2D se pueden convertir en una reconstrucción 3D global mediante el algoritmo 261, como se ha descrito anteriormente.

Cuando el algoritmo 261 se ejecuta en el teléfono móvil, por ejemplo mediante un programa de aplicación instalado en dicho teléfono, a continuación, el paciente puede transmitir la reconstrucción 3D global a su facultativo para una opinión de diagnóstico, estética, etc. Tan pronto como obtenga la reconstrucción 3D global, a partir de esta reconstrucción 3D global, el facultativo puede formular un primer diagnóstico médico y/o estético. Para ello, el facultativo no necesita una reconstrucción 3D muy precisa. Puede proponer eventualmente al paciente programar una cita en el consultorio. Durante esta cita, el facultativo puede entonces perfeccionar localmente la reconstrucción 3D al menos para los sectores de la arcada dental (o de las arcadas dentales) afectados por el tratamiento previsto, con ayuda de un aparato local 62, por ejemplo un IOS.

La reconstrucción 3D global 30 se puede transmitir directamente al odontólogo o al cirujano-dentista, por correo electrónico, o por una red inalámbrica de transmisión de datos, como una red de comunicaciones celular 3G-LTE, 4G o 5G. Puede pasar a través de un servidor informático, o servidor de cálculo, en el que algoritmos automáticos pueden procesar esta reconstrucción 3D global para derivar de ella varios indicadores, por ejemplo la detección automática de patologías (caries, retracción, etc.), o hacer propuestas estéticas (alineación automática de los dientes), editar presupuestos orientativos sobre tratamientos protésicos o estéticos, etc. Esta información orientativa se puede ofrecer al paciente, que puede decidir enviar a su facultativo para tener su opinión.

Como variante, el algoritmo de reconstrucción 3D 261 se puede ejecutar en un servidor informático, sobre la base de la imagen 2D de la arcada tomada por el paciente con su teléfono móvil. En este caso, es la imagen 2D la que se transmite al profesional por correo electrónico o por la red 3G-LTE, 4G o 5G, a través de dicho servidor. Un servidor de este tipo puede albergar una plataforma de servicios, que conecta pacientes y facultativos suscritos. Esta plataforma podrá disponer de medios de cálculo configurados para implementar el algoritmo de reconstrucción 3D, en concreto. A continuación, el facultativo obtiene la reconstrucción 3D indirectamente del paciente, a través de dicho servidor de servicios.

En otra variante, el aparato global puede ser un par de gafas de realidad aumentada, por ejemplo disponible en el consultorio del odontólogo o cirujano dentista. El sensor de este tipo de dispositivos tiene un campo relativamente amplio, es decir, capaz de abarcar una porción extendida de una arcada dental, por ejemplo al menos una semiarcada dental. Así, durante una primera inspección de la cavidad bucal mientras se lleva este dispositivo, el facultativo puede obtener, él mismo, directa y rápidamente una reconstrucción global a partir de la primera imagen adquirida. Esta solución tiene la ventaja de permitir integrar automáticamente la imagen 2D de la arcada dental en la cadena digital desmaterializada ("*digital workflow*") del facultativo.

Por ejemplo, la reconstrucción 3D global se puede efectuar en tiempo real cuando el facultativo mira la cavidad bucal a través de gafas de realidad aumentada.

La reconstrucción 3D global 30 se puede representar en transparencia alfa en las gafas de realidad aumentada a medida que avanza la estimación. Gracias al procedimiento de reconstrucción 3D descrito anteriormente, esta

visualización es fácil de configurar, dado que la imagen 2D que sirve para efectuar la reconstrucción 3D global 30 es la misma que la vista por las gafas. Dicha colocación de la reconstrucción 3D en la imagen 2D, conocida en sí misma, permite al facultativo saber en cualquier momento qué zona(s) de la arcada ha reconstruido por medio de las gafas, para poder completarla si es necesario cambiando el ángulo de visión.

5 En un modo de realización, la reconstrucción global 30 se sigue visualizando en las gafas de realidad aumentada, a medida que el médico la completa gracias un aparato de adquisición de imágenes local. Dado que cada reconstrucción 3D local 41, 42, 43, etc., está ajustada en la reconstrucción 3D global 30, entonces es posible visualizarla en transparencia alfa en gafas de realidad aumentada, por encima de la reconstrucción global 3D 30, por ejemplo usando un color diferente. Esto permite al médico saber en cualquier momento qué zona(s) ya ha especificado gracias al aparato local.

10 La presente invención ha sido descrita e ilustrada en la presente descripción detallada y en las figuras de los dibujos adjuntos, en posibles formas de realización.

15 En la presente descripción, el término "comprender" o "constar de" no excluye otros elementos u otras etapas. Los signos de referencia en los dibujos no han de comprenderse como que limitan el alcance de la invención.

REIVINDICACIONES

1. Procedimiento de generación de una reconstrucción tridimensional de una arcada dental de un paciente, llamada impresión óptica extendida, que comprende las siguientes etapas implementadas bajo el control de un facultativo:

- la obtención de al menos una imagen bidimensional global de la arcada dental adquirida con un primer sensor de imagen que tiene un campo de visión relativamente amplio y una precisión relativamente baja, y la generación, mediante unos primeros medios de procesamiento de datos digitales, de una reconstrucción tridimensional global, relativamente burda, de la arcada dental a partir de dicha imagen bidimensional;
- la adquisición de una pluralidad de imágenes bidimensionales locales respectivas de sectores asociados respectivos de la arcada dental con un segundo sensor de imagen que tiene un campo de visión relativamente más restringido y una precisión relativamente más alta que el primer sensor de imagen, y la generación, mediante segundos medios de procesamiento de datos digitales, de reconstrucciones tridimensionales locales respectivas, relativamente más precisas que la reconstrucción tridimensional global, de dichos sectores asociados de la arcada dental;
- la fusión, mediante un algoritmo de fusión implementado por terceros medios de procesamiento de datos digitales, de la reconstrucción tridimensional global y de las reconstrucciones tridimensionales locales en una única y misma reconstrucción tridimensional total en la que las reconstrucciones tridimensionales locales se sustituyen respectivamente por porciones de la reconstrucción tridimensional global que corresponden a los sectores asociados de la arcada dental, desempeñando la reconstrucción global un papel de guía espacial absoluta, comprendiendo dicha fusión un ajuste de las reconstrucciones tridimensionales locales en la reconstrucción tridimensional global, estando adaptado dicho ajuste para minimizar una distancia entre las nubes de puntos correspondientes a cada uno de los sectores asociados de la arcada dental y la nube de puntos correspondientes de la reconstrucción tridimensional global.

2. Procedimiento según la reivindicación 1, en el que el primer sensor de imagen es un sensor de imagen de un aparato digital portátil del paciente, como un teléfono móvil, y en el que la imagen bidimensional global es adquirida por el propio paciente con ayuda de su aparato digital portátil y se transmite desde dicho aparato digital portátil al facultativo para la generación de la reconstrucción tridimensional global.

3. Procedimiento según la reivindicación 1, que comprende la adquisición de la imagen bidimensional global por el facultativo con ayuda del primer sensor de imagen.

4. Procedimiento según la reivindicación 3, en el que el primer sensor de imagen está comprendido en unas gafas de realidad aumentada que comprenden además los primeros medios de procesamiento digitales para la generación de la reconstrucción tridimensional global.

5. Procedimiento según una cualquiera de las reivindicaciones anteriores, en el que el segundo sensor de imagen es un escáner intraoral, que comprende los segundos medios de procesamiento de datos digitales para la generación de las reconstrucciones tridimensionales locales.

6. Procedimiento según una cualquiera de las reivindicaciones 1 a 5, en el que la reconstrucción tridimensional global y/o las reconstrucciones tridimensionales locales se efectúan generando, para cada una, un mapa de profundidad a partir de una única imagen 2D adquirida por el primer sensor o por el segundo sensor de imagen, respectivamente, sobre la base de una arquitectura de aprendizaje profundo previamente entrenada para tal fin.

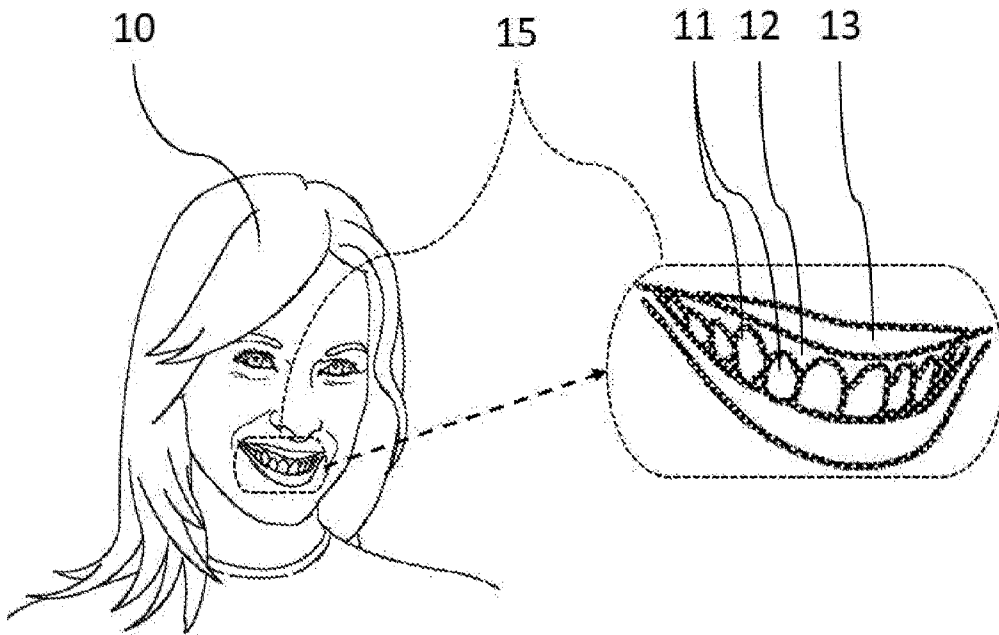
7. Programa informático que comprende instrucciones que, cuando el programa informático se carga en la memoria de un ordenador y es ejecutado por un procesador de dicho ordenador, causan la implementación por el ordenador de todas las etapas del procedimiento según una cualquiera de las reivindicaciones 1 a 6.

8. Dispositivo de generación de una reconstrucción tridimensional de una arcada dental de un paciente, llamada impresión óptica extendida, que comprende:

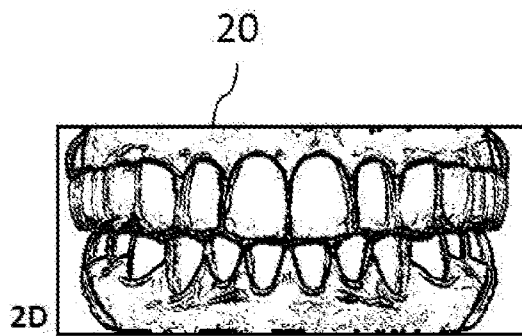
- un primer sensor de imagen que tiene un campo de visión relativamente amplio y una precisión relativamente baja para adquirir al menos una imagen bidimensional global de la arcada dental, y primeros medios de procesamiento de datos digitales para generar una reconstrucción tridimensional global, relativamente burda, de la arcada dental a partir de dicha imagen bidimensional;
- un segundo sensor de imagen que tiene un campo de visión relativamente más restringido y una precisión relativamente más alta que el primer sensor de imágenes, para adquirir una pluralidad de imágenes bidimensionales locales de sectores asociados respectivos de la arcada dental, y segundos medios de procesamiento de datos digitales para generar reconstrucciones tridimensionales locales respectivas, relativamente más precisas que la reconstrucción tridimensional global, de dichos sectores asociados de la arcada dental; y
- terceros medios de procesamiento de datos digitales configurados para implementar un algoritmo de fusión adaptado para fusionar la reconstrucción tridimensional global y las reconstrucciones tridimensionales locales en un

- 5 a única y misma reconstrucción tridimensional total en la que las reconstrucciones tridimensionales locales se sustituyen respectivamente por porciones de la reconstrucción tridimensional global que corresponden a los sectores asociados de la arcada dental, desempeñando la reconstrucción global un papel de guía espacial absoluta, comprendiendo el algoritmo de fusión un ajuste de las reconstrucciones tridimensionales locales en la reconstrucción tridimensional global, estando adaptado dicho ajuste para minimizar una distancia entre las nubes de puntos correspondientes a cada uno de los sectores asociados de la arcada dental y la nube de puntos correspondientes de la reconstrucción tridimensional global.
- 10 9. Dispositivo según la reivindicación 8, que comprende gafas de realidad aumentada que comprenden el primer sensor de imagen y, además, los primeros medios de procesamiento digitales para la generación de la reconstrucción tridimensional global.
- 15 10. Dispositivo según la reivindicación 8 o la reivindicación 9, que comprende un escáner intraoral que comprende el segundo sensor de imagen y, además, los segundos medios de procesamiento de datos digitales para la generación de las reconstrucciones tridimensionales locales.
- 20 11. Dispositivo según una cualquiera de las reivindicaciones 8 a 10, en el que los primeros medios de procesamiento digitales y/o los segundos medios de procesamiento digitales están configurados para generar la reconstrucción tridimensional global y las reconstrucciones tridimensionales locales, respectivamente, generando para cada una un mapa de profundidad a partir de una única imagen 2D adquirida por el primer sensor o por el segundo sensor de imagen, respectivamente, sobre la base de una arquitectura de aprendizaje profundo previamente entrenada para tal fin.
- 25 12. Aparato de toma de impresión óptica que comprende un ordenador que comprende los terceros medios de procesamiento de datos digitales del dispositivo según una cualquiera de las reivindicaciones 8 a 11, así como medios adaptados para la implementación del procedimiento según una cualquiera de las reivindicaciones 1 a 6.

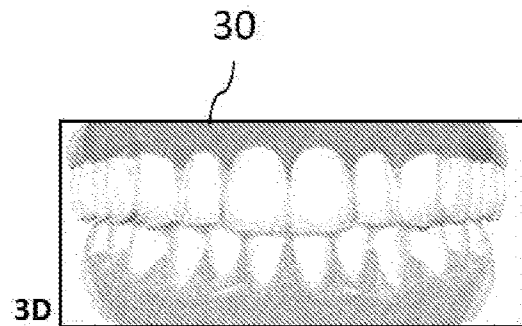
[Fig. 1]



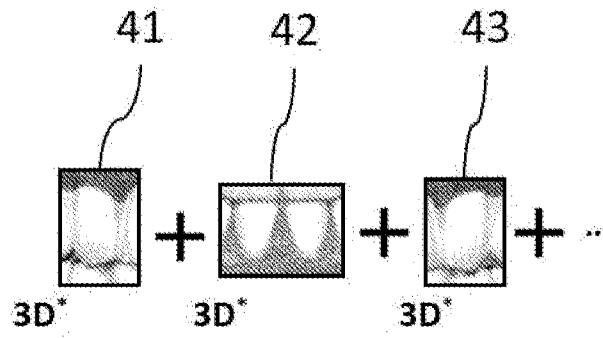
[Fig. 2]



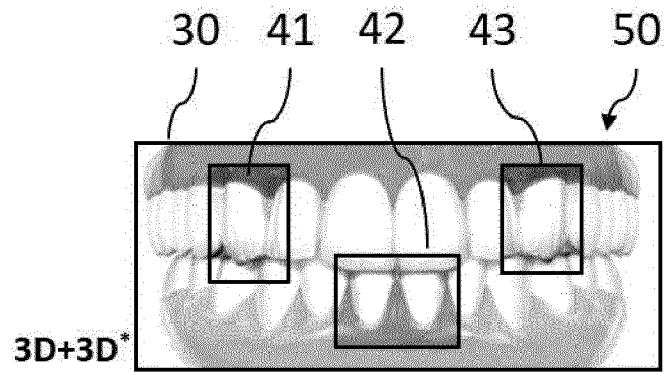
[Fig. 3]



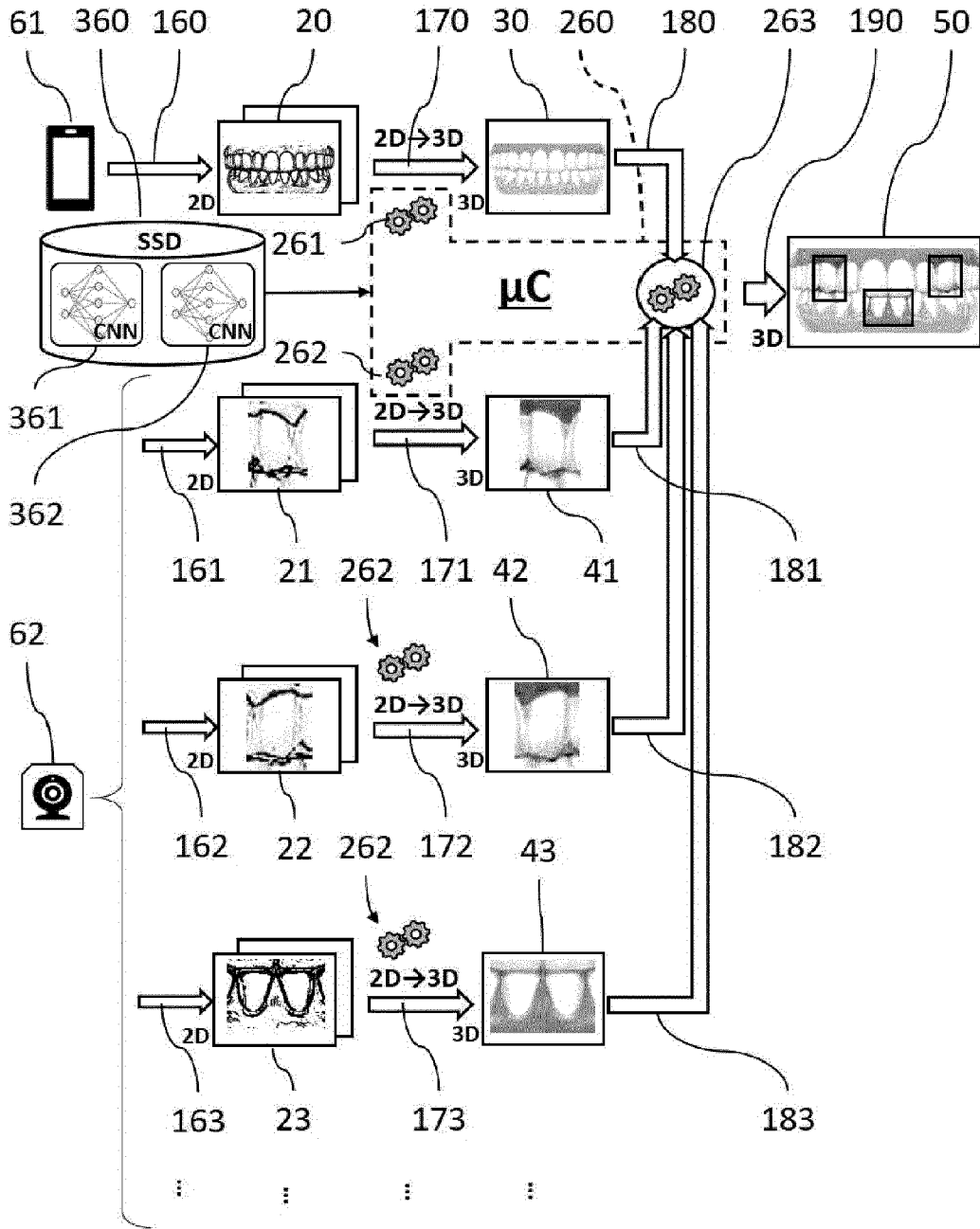
[Fig. 4]



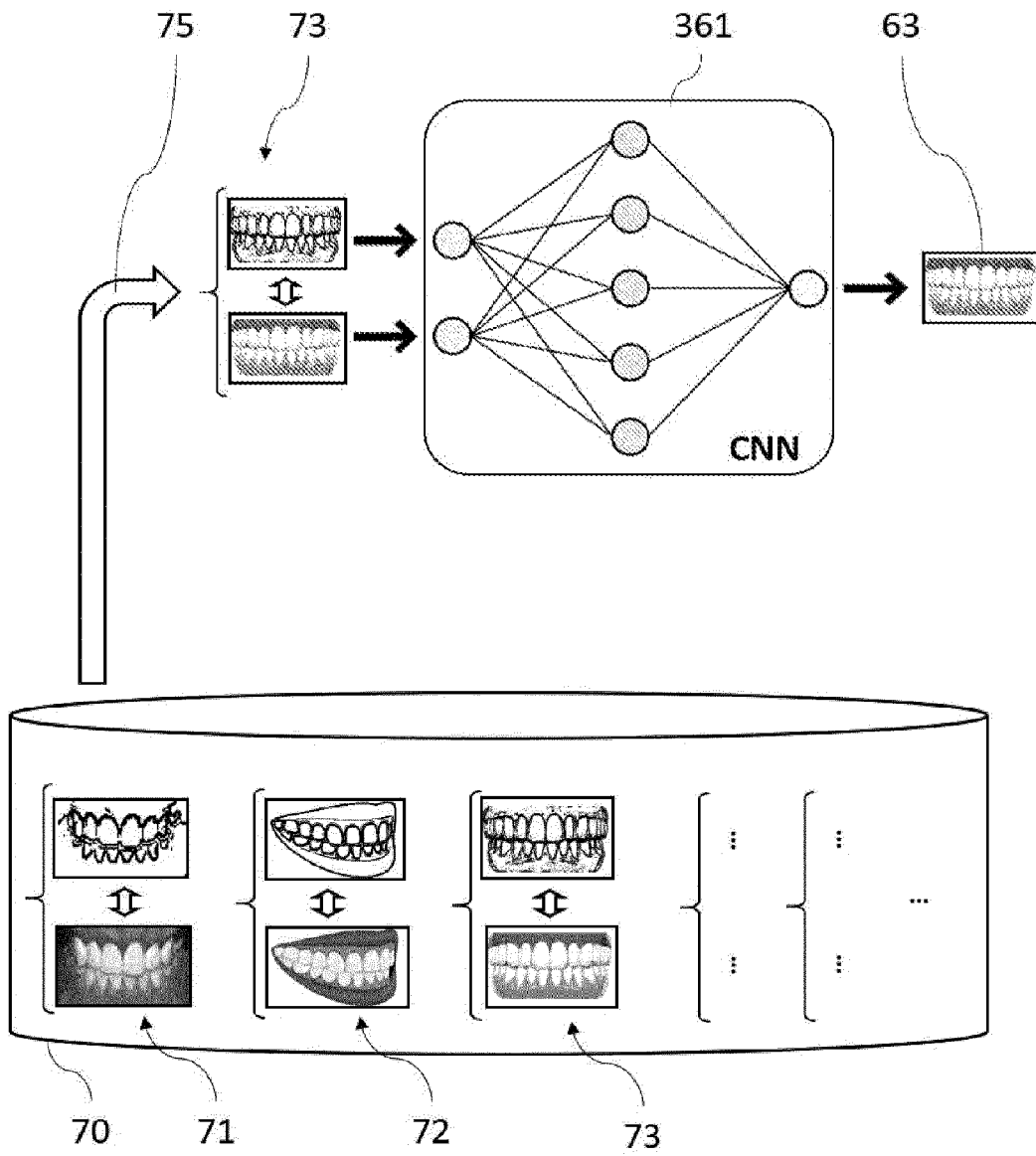
[Fig. 5]



[Fig. 6]



[Fig. 7]



[Fig. 8]

