(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 981 320**

(51) Int. Cl.:

G06T 7/55 (2007.01)
A61C 9/00 (2006.01)
G03B 35/02 (2011.01)
G03B 37/02 (2011.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **08.02.2019 PCT/EP2019/053138**

(87) Fecha y número de publicación internacional: **22.08.2019 WO19158442**

(96) Fecha de presentación y número de la solicitud europea: **08.02.2019 E 19704779 (8)**

(97) Fecha y número de publicación de la concesión europea: **24.04.2024 EP 3752982**

(54) Título: **Escaneo intraoral con diferenciación de superficies**

(30) Prioridad:

16.02.2018 DK PA201870094

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
08.10.2024

(73) Titular/es:

3SHAPE A/S (100.0%)
Holmens Kanal 7
1060 Copenhagen K, DK

(72) Inventor/es:

ÖJELUND, HENRIK;
HOEDT, ASGER VEJEN y
HOLLENBECK, KARL-JOSEF

(74) Agente/Representante:

LINAGE GONZÁLEZ, Rafael

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

Escaneo intraoral con diferenciación de superficies

Campo

En el presente documento se divulga un sistema, método y producto de programa de ordenador de escáner para escanear la cavidad intraoral de un paciente. En particular, la divulgación se refiere a unificar una representación digital 3D de la cavidad intraoral, teniendo en cuenta la deformación del tejido durante el escaneo. Se pueden usar algoritmos de búsqueda de puntuaciones, tales como algoritmos de aprendizaje automático, para entrenar al sistema para diferenciar entre diversos tipos de superficies y ponderar las diferentes superficies para unificar vistas grabadas a una representación digital 3D combinada.

Antecedentes

En odontología, las mediciones de topografía 3D de los dientes del paciente y posiblemente de otras partes de la cavidad intraoral son necesarias como base para tratamientos de restauración u ortodoncia. Tradicionalmente, tal medición 3D se realizaba tomando inicialmente una impresión física. Debido a que este procedimiento es generalmente desagradable para los pacientes, se han usado más recientemente escaneos 3D intraorales para medir directamente la topografía de los dientes u otras partes de la cavidad intraoral.

Debido a las limitaciones de tamaño, los escáneres 3D intraorales graban típicamente pequeñas vistas a la vez, y cada vista contiene un mapa de distancias y posiblemente otra información tal como el color. Las vistas se unifican de manera incremental para una medición topográfica 3D combinada a medida que se mueve el escáner. Tal grabación, por ejemplo para los dientes de una única mandíbula y la encía circundante como región de interés, dura típicamente al menos un minuto y produce típicamente al menos 100 vistas. Los términos “registro” y “unificación” se usan en general indistintamente en la literatura. El modelo unificado, convertido también habitualmente en una superficie, se denomina a menudo “modelo 3D virtual” o “representación digital 3D” de la topografía 3D medida por el escáner.

Varios tipos de superficies dentro de la cavidad intraoral no son rígidas. Por ejemplo, las mejillas y la lengua pueden deformarse significativamente durante la grabación. La encía puede también deformarse, aunque típicamente en menor medida. Además, se encuentran objetos extraños durante la grabación con un escáner intraoral. Algunos objetos extraños son rígidos pero altamente móviles y solo están típicamente presentes en algunas vistas de la grabación, tales como instrumentos dentales. Otros objetos extraños, como almohadillas de algodón, pueden moverse menos y estar presentes en más vistas de la grabación, pero ocasionar más deformación. Típicamente, el dedo de un odontólogo se mueve y se deforma significativamente en sólo unas pocas vistas de una grabación.

Debido a que la unificación de vistas se basa generalmente en la suposición de que la superficie escaneada es estable y rígida, cualquier superficie móvil o deformable da típicamente como resultado una pérdida de precisión de la medición topográfica 3D combinada. Existen algoritmos de unificación no rígida, pero son caros informáticamente y requieren información adicional, por ejemplo puntos de referencia, que no está generalmente disponible durante el escaneo intraoral.

Se han introducido varios medios para reducir el impacto perjudicial de las superficies móviles o deformables en el escaneo 3D intraoral. Una estrategia es mantener tales superficies alejadas de las vistas, por ejemplo mediante el uso de un retractor de mejilla. Dado que a menudo el paciente percibe como desagradable el retractor de mejilla y que sólo soluciona parte del problema, se han introducido métodos de procesamiento de datos.

El documento US7698068 describe un método para distinguir dientes y otros tejidos intraorales en base al color, y usar únicamente la parte de las vistas que representan color de dientes para unificar durante la grabación. Como los dientes son rígidos y más blancos que otros tejidos, la calidad de la medición topográfica 3D combinada a menudo se puede mejorar. Sin embargo, los dientes pueden decolorarse, tanto de manera natural como mediante restauraciones, por lo que una clasificación únicamente por color puede ser inexacta. Además, el tejido no blanco sino más bien rígido, tal como el de las arrugas palatinas, puede ser útil para unificar, especialmente en casos desdentados, y, de este modo, no debe obviarse. Además, algunas superficies extrañas deformables, por ejemplo discos de algodón, pueden tener un color similar al de los dientes, pero deberían obviarse al unificar.

El documento US9629551 describe un método para detectar objetos móviles analizando la consistencia de múltiples vistas durante la grabación de la misma parte de la cavidad intraoral. Este método usa únicamente información geométrica, y, por consiguiente, es resistente a la variabilidad de color.

El documento US2015/164335 A1 describe la unificación/el registro de varios escáneres intraorales usando fluorescencia y profundidad, usándose el método de segmentación de Kronfeld et al., “Snake-Based Segmentation of Teeth from Virtual Dental Casts” (en Computer-Aided Design & Applications CAD Solutions, 2019) para excluir los tejidos blandos de la unificación. Además, Hazirbas et al., “FuseNet: Incorporating Depth into Semantic

Segmentation via Fusion-Based CNN Architecture" (en Asian Conference on Computer Vision, 2016), divulga un algoritmo para la segmentación que usa conjuntamente color y profundidad.

Sigue existiendo la necesidad de un escáner 3D intraoral y de un método de uso del escáner que sea generalmente resistente a superficies que se deforman o se mueven.

Sumario

En un aspecto, en el presente documento se divulga un método para generar una representación digital 3D de al menos una parte de una cavidad intraoral, comprendiendo el método:

- grabar una pluralidad de vistas que contienen datos de superficie que representan al menos la geometría de los puntos de superficie de la parte de la cavidad intraoral usando un escáner intraoral;

- determinar una ponderación para cada punto de superficie, al menos parcialmente en base a en puntuaciones, que son medidas de creencia de que ese punto de superficie representa un tipo particular de superficie;

- ejecutar un algoritmo de unificación que realiza una unificación ponderada de los puntos de superficie en dicha pluralidad de vistas para generar la representación digital 3D en base a las ponderaciones determinadas;

en el que las puntuaciones para los puntos de superficie se encuentran mediante al menos un algoritmo de búsqueda de puntuaciones que toma como entrada al menos la parte de geometría de los datos de superficie para ese punto de superficie y datos de superficie para puntos en un entorno de ese punto de superficie, y en el que el escáner también suministra alguna información de certeza de los datos de superficie medidos para las vistas grabadas, y en el que dicha información de certeza modifica al menos parcialmente las puntuaciones.

Pueden configurarse una o más unidades de procesamiento para aplicar un algoritmo, tal como un algoritmo de aprendizaje automático, entrenado en datos para diferenciar entre diversos tipos de superficie que indican tipos de tejido, otras superficies, y, posiblemente, datos erróneos en las vistas grabadas. Cada vista contiene datos de geometría de superficie, al menos algunos puntos $z(x, y)$ definidos en un sistema de coordenadas relativo al escáner. El método de la invención usa ponderación para unificar vistas grabadas para una representación combinada de la topografía 3D, denominada también representación digital 3D. La ponderación de un punto en la unificación está determinada al menos parcialmente por puntuaciones que son medidas de creencia de que ese punto representa al menos un tipo de superficie. Las medidas de creencia, y, por consiguiente, dichas puntuaciones, pueden ser medidas o probabilidades heurísticas.

Es posible que un escáner y un método de acuerdo con esta divulgación no necesariamente detecten directamente el movimiento o la deformación de tejido o de objeto extraño. Típicamente, diferencia por tipos de superficies en base a su supuesta propensión a moverse o deformarse, independientemente de si una superficie real de tal tipo se ha movido o deformado durante un escaneo. Los tipos de superficie pueden basarse en histología, por ejemplo dentina o mucosa. También pueden basarse en ubicación, por ejemplo la encía entre los dientes tiene una menor propensión a la deformación que la encía alrededor de un diente preparado. Los tipos de superficie también pueden ser heurísticos, por ejemplo si se desean o no en una representación digital 3D. Los tipos de superficie que tienen una propensión relativamente menor a la deformación o al movimiento son generalmente más deseables para la unificación.

En esta divulgación, la diferenciación por tipo de superficie puede basarse al menos parcialmente en datos de geometría de superficie, mientras que la técnica conocida requiere datos adicionales para la diferenciación, por ejemplo el color de superficie, como en el documento US7698068. Aún así, un escáner o método de acuerdo con una realización de esta divulgación también puede proporcionar y explotar datos de superficie adicionales en vistas, por ejemplo el color de superficie. Un escáner o método de acuerdo con esta divulgación también proporciona y explota la certeza de los datos de superficie que graba.

La unificación ponderada se puede realizar con una de las muchas variantes del algoritmo del punto más cercano iterativo (ICP) o con otros algoritmos apropiados en la técnica. Se describe la ponderación por pares; véase, por ejemplo, [1]. En otra formulación de unificación ponderada, los datos de superficie se pueden ordenar en base a sus ponderaciones, y, luego, sólo se usa en la unificación algún cuantil superior, o que supere algún umbral. La unificación ponderada en el sentido de esta divulgación es alguna formulación matemática que expresa diferenciación, de tal modo que algunos datos en las vistas tienen un impacto relativamente mayor en el resultado que otros.

El sistema de coordenadas común para el modelo unificado puede ser el sistema de coordenadas local de la primera vista. Unificar una vista debe entenderse como unificar los datos $z(x, y)$ de geometría de superficie de la vista transformándolos a un sistema de coordenadas común, mientras se aplica la misma transformada geométrica a cualquier otro dato de superficie en la vista. Los puntos de datos de superficie dentro de una vista con ponderación cero o pequeña pueden incluirse en la transformada, o pueden eliminarse.

Una representación digital 3D se puede representar de varias maneras. La unificación por sí sola proporciona al menos una nube de puntos. A menudo es deseable aproximar una nube de puntos con una superficie, por ejemplo una malla triangular, nivelando el ruido en los datos de puntos y proporcionando una representación digital 3D que es una mejor base para CAD dental. Algunos algoritmos construyen tal superficie después de que se hayan unificado todas las vistas, por ejemplo [2]. Algunos algoritmos construyen algún modelo de superficie intermedio de manera incremental para cada vista grabada y unificada, posiblemente también usando ese modelo intermedio para mejorar la unificación, por ejemplo [3]. Después de grabar todas las vistas, a menudo se calcula una superficie final que reemplaza a la intermedia. Los puntos de datos de superficie con ponderaciones pequeñas, si no se retiran cuando se unificó la vista respectiva, a menudo se retiran efectivamente en este paso, porque se detectan como ruido.

En algunos ejemplos, los puntos *z(x, y)* están dispuestos como un mapa de distancias, es decir, como distancias *z(x, y)* desde alguna superficie de referencia definida con relación al escáner hasta la superficie escaneada. En algunos ejemplos, las coordenadas *(x, y)* existen en una cuadrícula sobre una superficie de referencia plana. Un punto de datos de superficie en el sentido de esta divulgación contiene al menos información de geometría, es decir, *z(x, y)*. También se puede aumentar con otros datos grabados para la superficie en *(x, y)*, por ejemplo con el color o con alguna medida de la certeza de *z(x, y)*, o con algún otro dato.

Para diferenciar por tipo de superficie para una ubicación *(x,y)*, un escáner de esta divulgación tiene en cuenta el valor *z(x, y)*, y también valores adicionales de *z* en un entorno de *(x,y)*. Considerar un entorno puede revelar alguna estructura geométrica típica de un tipo de superficie. El entorno puede ser un entorno inmediato o un conjunto de regiones cercanas que se extienden más allá del entorno inmediato. Puede resultar útil aplicar un núcleo para revelar una estructura geométrica, o un conjunto de núcleos. Considerar la información adicional contenida en entornos es otra mejora con respecto a la técnica conocida.

En algunas realizaciones, el al menos un algoritmo de búsqueda de puntuación es un algoritmo de aprendizaje automático. Se puede usar cualquier tipo de algoritmo de aprendizaje automático. Algunos ejemplos de algoritmos de aprendizaje automático incluyen redes neuronales artificiales, tales como redes neuronales artificiales profundas, redes neuronales artificiales convolucionales, o redes neuronales artificiales recurrentes. El método de aprendizaje automático de las realizaciones de esta divulgación puede aplicar métodos de reducción de dimensionalidad, tales como análisis de componentes principales o autocodificadores.

En algunas realizaciones, el al menos un algoritmo de aprendizaje automático comprende una red neuronal con al menos una capa convolucional. Las redes neuronales convolucionales proporcionan naturalmente una consideración de entornos. Para mapas de distancia o colores definidos o remuestreados en una cuadrícula de *(x, y)*, preferiblemente una cuadrícula equidistante, se pueden aplicar de manera análoga muchos algoritmos de aprendizaje automático publicados para el análisis de imágenes y la segmentación de imágenes. El algoritmo para diferenciar entre tipos de superficies puede también ser un algoritmo de aprendizaje automático más clásico, por ejemplo usando máquinas vectoriales de soporte. El algoritmo para diferenciar entre tipos de superficies puede también basarse en métodos estadísticos más clásicos, tales como la estadística bayesiana o un tipo de regresión. También se pueden usar en combinación varias de las clases de algoritmos anteriores.

En algunas realizaciones, el al menos un algoritmo de aprendizaje automático se entrena en una pluralidad de los tipos de superficies que comúnmente se graban con escáneres en cavidades intraorales. Al anotar las imágenes del conjunto de entrenamiento en los diversos tipos de superficies que normalmente se encuentran en las cavidades intraorales, tales como dientes, encías, lengua, paladar, etc., la determinación de la ponderación resultante será más sólida y consistente.

El entrenamiento de un algoritmo de aprendizaje automático para diferenciar entre tipos de superficies puede ser supervisado, semisupervisado o no supervisado. Para el aprendizaje semisupervisado o supervisado, el entrenamiento puede basarse, al menos parcialmente, en vistas anotadas, o en representaciones digitales 3D anotadas. Las anotaciones en una representación digital 3D se pueden retroproyectar para cada vista que contribuyó a esa representación digital 3D, porque la unificación dio también lugar a las transformaciones de cada vista para un sistema de coordenadas común. Por consiguiente, las anotaciones se pueden llevar a las vistas, y se pueden usar para entrenar a un algoritmo de aprendizaje automático en base a vistas. La anotación puede ser realizada por un ser humano y/o por algún algoritmo.

Durante el escaneo, el algoritmo de aprendizaje automático se ejecuta en modo de inferencia, detectando puntuaciones que son una medida de creencia de datos de superficie que pertenecen a uno o más tipos de superficie. Típicamente, las puntuaciones se pueden representar como un vector con un valor para cada tipo de superficie en la inferencia. En realizaciones que usan una red neuronal, las puntuaciones se obtienen típicamente de la capa de salida, posiblemente después de aplicar alguna transformación, tal como una transformación logarítmica.

La medida de creencia, tal como se usa en esta solicitud, indica alguna puntuación que indica un grado de certeza. Una medida de creencia puede ser una probabilidad, particularmente si se conoce o se supone que se conoce la

distribución de la variable aleatoria subyacente. Cuando tal conocimiento no existe ni ninguna suposición parece justificada, o si se prefiere por otras razones, una medida de creencia puede ser alguna evaluación y/o expresión más subjetiva de dicho grado de certeza.

Puede ser conveniente matemáticamente que una puntuación de uno represente la creencia cierta de que los datos de superficie pertenecen a un tipo de superficie particular, tal como diente o encía, mientras que una puntuación de cero representa la creencia cierta de que los datos de superficie no pertenecen a ese tipo de superficie particular. Las puntuaciones que aumentan de cero a uno representan una creencia cada vez mayor de que los datos de superficie pertenecen a ese tipo de superficie en particular.

Una ponderación para un punto de datos de superficie en la unificación se encuentra a partir de las puntuaciones para ese punto, por ejemplo en función de las puntuaciones. La realización en la que dicha función es 1 para el tipo de superficie con la puntuación más alta y 0 en caso contrario se conoce en la técnica como clasificación. Un ejemplo de un algoritmo de aprendizaje automático usado para la clasificación es [4]. Puede resultar ventajoso usar funciones más refinadas, por ejemplo devolver un valor de 1 sólo si la puntuación más alta es significativamente mayor que todas las demás, por ejemplo mayor que la suma de todas las demás. También puede ser ventajoso que la función devuelva valores distintos de cero para varios tipos de superficie, por ejemplo si hay motivos para creer que un punto de datos de superficie puede ser cualquiera de los varios tipos de superficie. La función puede también devolver 0 para todos los tipos de superficie en los casos en que ninguna puntuación sea grande, o en situaciones similares deficientemente determinadas.

En algunas realizaciones, el tipo de superficie representa más de un tipo de tejido intraoral. En algunos casos, puede resultar ventajoso agrupar diferentes tipos de tejidos intraorales, por ejemplo agrupar la superficie de diente junto con la parte superior de la encía, ya que esto es útil para unificar la representación digital 3D.

En algunas realizaciones, la ponderación de cada punto de superficie en la unificación se determina también mediante ponderaciones para los tipos de superficies. Esto significa que la ponderación para algún punto de datos de superficie particular en la unificación se encuentra a partir de las puntuaciones y de las ponderaciones de tipo de superficie, por ejemplo como una combinación lineal sobre todos los tipos de superficie de los productos de ponderaciones de tipo de superficie y dichas puntuaciones. Las ponderaciones de tipo de superficie se asignan preferiblemente *a priori*, recibiendo los tipos de superficie deseables para unificar ponderaciones más altas y otros con menos ponderación. En algunas realizaciones, algunas ponderaciones de tipo de superficie se establecen en cero, por lo que los datos de superficie de esos tipos de superficie se filtran de las vistas. Puede haber consideraciones adicionales que afecten la formulación de la ponderación, por ejemplo el tamaño de un parche de superficie que representa un punto de datos de superficie, por ejemplo porque es el punto de datos más cercano para todos los puntos dentro del parche.

En algunas realizaciones, la inferencia se puede ejecutar en tiempo real o casi en tiempo real, mientras se graban las vistas. Esto puede permitir que la unificación se realice en tiempo real o casi en tiempo real también. Es ventajoso realizar la unificación en tiempo real, porque se puede construir y visualizar una representación 3D de un sitio mientras el usuario escanea, lo que le ayuda a navegar por el sitio. De este modo, estas realizaciones son una clara mejora con respecto a otras aplicaciones de inferencia de aprendizaje automático, tales como, por ejemplo, los documentos US7720267, US8170306 y US9349178.

Otras realizaciones de la inferencia de aprendizaje automático de acuerdo con esta divulgación pueden ejecutarse más lentamente, tal como después de que se hayan grabado dos o más vistas, pero proporcionar una mayor precisión. También es posible combinar cierto grado limitado de ponderación de datos de superficie en base a alguna inferencia de datos únicos con ponderación de datos de superficie adicional en base a alguna inferencia de múltiples vistas, proporcionando potencialmente una buena combinación de velocidad y precisión.

En algunas realizaciones, los datos de superficie también comprenden información de color. Añadir información de color a los datos de superficie puede hacer que la determinación del tipo de tejido sea más segura.

En algunas realizaciones, al menos un algoritmo de aprendizaje automático fue entrenado al menos parcialmente usando datos grabados por un escáner intraoral. Dado que puede haber variaciones en la sensibilidad y la calidad de la imagen entre escáneres de diferentes fabricantes, el resultado será más preciso cuanto más coincidan los datos usados para entrenar el algoritmo de aprendizaje automático con los escaneos que posteriormente adquirirá un usuario.

En algunas realizaciones, al menos un algoritmo de aprendizaje automático fue entrenado al menos parcialmente por un operario del escáner intraoral. El sistema de escáner de esta divulgación se puede suministrar al usuario con al menos un algoritmo de aprendizaje automático previamente entrenado. En otras realizaciones, el usuario del escáner realiza al menos algún entrenamiento después de haber recibido el escáner. Los datos de entrenamiento adicionales podrían contener, por ejemplo, imágenes en color o datos de geometría de superficie en los que aparezca la clase especial de guantes o de rollos de algodón que usa un odontólogo. Los datos de entrenamiento adicionales también podrían proceder de un grupo étnico de pacientes del que el odontólogo tiene una proporción

superior a la media. Con entrenamiento adicional, el escáner también puede adaptarse al estilo de escaneo del usuario. Se puede realizar entrenamiento adicional en una o más unidades de procesamiento del sistema de escáner o en la nube. Puede resultar ventajoso personalizar el algoritmo de aprendizaje automático con entrenamiento adicional porque probablemente funcionará mejor.

En algunas realizaciones, se selecciona un algoritmo de búsqueda de puntuación para un tipo de aplicación y se selecciona al menos otro algoritmo para otro tipo de aplicación. Puede resultar ventajoso entrenar varios algoritmos de aprendizaje automático para diferentes tipos de aplicaciones, para su posterior selección durante la inferencia. La selección del algoritmo apropiado puede ser hecha, por ejemplo, por el usuario del escáner en una interfaz de usuario.

Los tipos de aplicaciones pueden diferir en el conjunto de tipos de superficie entrenados o inferidos. Por ejemplo, un algoritmo con un conjunto que contenga un tipo de superficie que represente las papilas interdentales y las bolsas gingivales puede ser relevante para el seguimiento de pacientes con gingivitis. En otro ejemplo, un algoritmo con un conjunto que contenga un tipo de superficie que represente parte de las encías podría ser relevante para pacientes edéntulos, donde los datos sobre la superficie dental son escasos y normalmente no suficientes para unificar.

En otros ejemplos, un tipo de aplicación se caracteriza al menos parcialmente por al menos uno de un grupo particular de edad de paciente, una etnia particular de paciente, un estilo particular de tratamiento, una indicación médica particular, un tipo particular de equipo usado junto con el escáner, o una región particular de la cavidad intraoral. Por ejemplo, un algoritmo puede ser más adecuado para niños y otro para adultos, o para una etnia frente a otras. Los tipos de aplicación también pueden representar diferentes estilos de tratamiento dental, por ejemplo según lo determinado por equipos o procedimientos operativos estándar específicos de la organización o región, o similares.

En algunos ejemplos, las puntuaciones se suman en la pluralidad de vistas.

Al unificar subescaneos, se puede crear una representación digital 3D provisional. Luego, se puede obtener una imagen de cada vóxel en la representación provisional desde múltiples vistas, y, luego, las puntuaciones se pueden sumar en las múltiples vistas, para realizar una determinación de puntuación más sólida.

En algunas realizaciones, se aplican otros algoritmos o criterios para filtrar datos de las vistas grabadas.

En algunas realizaciones, uno de dichos otros algoritmos evalúa la coherencia geométrica en una pluralidad de vistas. Un ejemplo de ello es la detección de objetos móviles en base a la consistencia geométrica como se divulga en el documento US9629551 B1. Durante la inferencia, filtrar los datos de superficie en base a otros criterios antes de encontrar puntuaciones simplifica el problema de unificación, mientras que filtrar los datos de superficie en base a otros criterios después de encontrar dichas puntuaciones puede mejorar los resultados globales. Sin embargo, durante el entrenamiento, puede resultar ventajoso no filtrar en base a otros criterios, reteniendo relativamente más datos de entrenamiento de esta manera.

Una realización ventajosa de esta divulgación usa una combinación de filtrado en base a la consistencia geométrica y a la segmentación semántica. En esta realización, se construye un volumen excluido a partir de todos los datos en el mismo espacio que la representación digital 3D que se construye sólo a partir de aquellos datos de superficie que pertenecen a segmentos de tipos de superficie deseables. Entonces se pueden retirar partes de la representación digital 3D que están en el volumen excluido, tal como después de que se hayan recopilado todas las vistas y, por consiguiente, se haya recopilado la mayor parte de la información sobre el volumen excluido. También es factible unificar en base a datos que pasan únicamente el filtrado, pero también conservar los datos filtrados para algún análisis posterior.

El escáner también suministra alguna información de certeza de los datos de superficie medidos para las vistas grabadas, y dicha información de certeza determina al menos parcialmente las puntuaciones. En algunos ejemplos, el escáner es un escáner de enfoque, tal como el escáner de enfoque divulgado en el documento US8878905. Un escáner de enfoque puede suministrar una certeza de los datos medidos $z(x, y)$ a partir de la distintividad de una medida de enfoque. Otras clases de escáneres 3D pueden también proporcionar información sobre la certeza de datos de superficie medidos. Por ejemplo, los escáneres que usan triangulación o patrones de luz proyectados grabados con al menos dos cámaras pueden proporcionar dos vistas simultáneas y derivar la certeza del grado de coherencia entre ellas. Otros escáneres 3D pueden deducir la certeza del contraste de imagen o de otra información. Sin embargo, otros escáneres pueden proporcionar certeza sobre otros datos de superficie, tales como el color.

Las certezas de los datos de superficie se usan para modificar adicionalmente sus ponderaciones en la unificación, y se pueden usar durante el entrenamiento y la inferencia.

La certeza, u otros datos de superficie en una vista, se pueden expresar matemáticamente, por ejemplo como canales adicionales en un mapa de distancia aumentada. Muchos algoritmos de aprendizaje automático publicados para el análisis de imágenes multicanal y la segmentación de imágenes se pueden aplicar de manera análoga en

esta divulgación.

En otro aspecto de esta divulgación, en el presente documento se divulga un sistema de escáner para generar una representación digital 3D de al menos una parte de una cavidad bucal de acuerdo con la reivindicación 10.

La unidad de procesamiento de datos que ejecuta el algoritmo de aprendizaje automático puede ser parte del escáner intraoral o puede estar contenida en otro recinto al que esté conectado el escáner portátil. La demanda de energía y el requisito reglamentario de que el escáner portátil se mantenga relativamente frío hacen que sea ventajoso colocar los medios de procesamiento en un recinto separado. La una o más unidades de procesamiento pueden ser un PC, una FPGA, o similar, pueden también contener una GPU, y pueden también realizar otro procesamiento de datos. Las unidades de procesamiento pueden estar conectadas a un dispositivo de exhibición visual en el que se muestre el modelo virtual según está siendo unificado durante el escaneo.

En algunas realizaciones, el al menos un algoritmo de búsqueda de puntuación es un algoritmo de aprendizaje automático.

En algunos ejemplos, el escáner tiene un sistema óptico al menos casi telecéntrico. Es típicamente más fácil entrenar y usar para inferir un algoritmo de aprendizaje automático cuando las vistas no se ven afectadas por la escala, es decir, cuando se obtienen imágenes de un tipo de superficie dado con la misma resolución y tamaño en toda la profundidad de campo del escáner. Un escáner con un sistema óptico telecéntrico proporciona esta ventaja por construcción, mientras que un escáner con un sistema óptico casi telecéntrico, tal como uno con un ángulo de visión mayor que cero pero inferior a 10 grados, proporciona una aproximación a la misma. Para escáneres con un ángulo de visión mayor, puede resultar ventajoso volver a muestrear las vistas antes de usarlas en el aprendizaje automático. Por ejemplo, se puede calcular una vista ortonormal aparente dado el conocimiento del sistema óptico desde su construcción o calibración. Como el remuestreo puede compensar los efectos de escala en el tamaño, pero no en la resolución, un escáner con un sistema óptico casi telecéntrico puede ser preferible a un escáner que use remuestreo.

En algunos ejemplos de esta divulgación, el escáner es un escáner intraoral confocal.

En algunos ejemplos de esta divulgación, el escáner también puede suministrar y explotar cierta información de certeza de los datos de superficie medidos. En algunos de tales ejemplos, el escáner es un escáner de enfoque, tal como el escáner de enfoque divulgado en el documento US8878905. Un escáner de enfoque puede suministrar una certeza de los datos $z(x, y)$ medidos a partir de la distintividad de una medida de enfoque. Otros tipos de escáneres 3D pueden también proporcionar información sobre la certeza de los datos de superficie medidos. Por ejemplo, los escáneres que usan la triangulación de patrones de luz proyectados grabados con dos cámaras pueden proporcionar dos vistas simultáneas y derivar la certeza del grado de coherencia entre ellas. Otros escáneres 3D pueden deducir la certeza del contraste de imagen o de otra información. Aún otros escáneres pueden proporcionar certeza sobre otros datos de superficie, tal como el color.

Las certezas de los datos de superficie se pueden usar para modificar adicionalmente sus ponderaciones en la unificación, o se pueden usar durante el entrenamiento y la inferencia. La certeza u otros datos de superficie en una vista se pueden expresar matemáticamente, por ejemplo como canales adicionales en un mapa de distancia aumentada. Muchos algoritmos de aprendizaje automático publicados para el análisis de imágenes multicanal y la segmentación de imágenes se pueden aplicar de manera análoga de acuerdo con los ejemplos de esta divulgación.

Típicamente, es más fácil entrenar y usar para inferencia un algoritmo de aprendizaje automático cuando las vistas no se ven afectadas por la escala, es decir, cuando se visualiza un tipo de superficie dado con la misma resolución y tamaño en toda la profundidad de campo del escáner. Un escáner con un sistema óptico telecéntrico proporciona esta ventaja por construcción, mientras que un escáner con un sistema óptico casi telecéntrico, tal como uno con un ángulo de visión mayor que cero pero inferior a 10 grados, proporciona una aproximación a la misma. Para escáneres con un ángulo de visión mayor, puede resultar ventajoso volver a muestrear las vistas antes de usarlas en aprendizaje automático. Por ejemplo, se puede calcular una vista ortonormal aparente dado el conocimiento del sistema óptico desde su construcción o calibración. Como el remuestreo puede compensar los efectos de escala en el tamaño, pero no en la resolución, un escáner con un sistema óptico casi telecéntrico puede ser preferible a un escáner que use remuestreo.

En otro aspecto, en el presente documento se divulga un producto de programa de ordenador de acuerdo con la reivindicación 9.

Breve descripción de los dibujos

Los objetos, características y ventajas anteriores y/o adicionales de las realizaciones de esta divulgación se describirán adicionalmente mediante la siguiente descripción detallada ilustrativa y no limitante de las realizaciones de la presente divulgación, con referencia al dibujo o dibujos adjuntos, en los que:

la figura 1 muestra un sistema de escáner de acuerdo con una realización de esta divulgación;

la figura 2 muestra una vista esquemática de una parte portátil de ejemplo de un sistema de escáner de acuerdo con una realización de esta divulgación;

la figura 3 muestra una vista de un patrón de tablero de ajedrez de período espacial usado para determinar mapas de distancias de acuerdo con una realización de esta divulgación;

la figura 4 muestra curvas ejemplares para dos grupos de píxeles que tienen su máximo en la misma ubicación de acuerdo con una realización de esta divulgación;

la figura 5a muestra una vista en corte transversal de una parte de la cavidad bucal inferior de acuerdo con una realización de esta divulgación;

la figura 5b muestra un procedimiento para entrenar un algoritmo de aprendizaje automático de acuerdo con una realización de esta divulgación;

la figura 6 muestra la arquitectura de una red neuronal convolucional adecuada de acuerdo con una realización de esta divulgación;

la figura 7a muestra un ejemplo de cómo se puede aplicar un algoritmo de aprendizaje automático entrenado en modo de inferencia de acuerdo con una realización de esta divulgación;

la figura 7b muestra otro ejemplo de cómo se puede aplicar un algoritmo de aprendizaje automático entrenado en modo de inferencia de acuerdo con una realización de esta divulgación; y

la figura 8 muestra un ejemplo de cómo el algoritmo de la figura 7 (b) se expande para usar también filtrado en base al volumen excluido de acuerdo con una realización de esta divulgación.

Descripción detallada

La figura 1 muestra una aplicación de ejemplo de un sistema de escáner de acuerdo con una realización de esta divulgación. El odontólogo sostiene el escáner 100 de mano. Este escáner de mano suele estar conectado mediante un cable 60 a un ordenador portátil 50 con una pantalla 55. En algunos casos, el escáner de mano puede ser inalámbrico. La representación digital 3D que es resultado de la unificación se exhibe visualmente durante el escaneo en la pantalla 55. La exhibición visual se actualiza con el tiempo, a medida que se graban y unifican nuevas vistas a la representación digital 3D. El portátil 50 es el recinto para una o más unidades de procesamiento que incluyen una CPU y una GPU que ejecutan los algoritmos de las realizaciones de esta divulgación. Algunos escáneres se comunican con los medios de procesamiento mediante transferencia inalámbrica en lugar de mediante un cable.

La figura 2 muestra esquemáticamente un ejemplo de una parte portátil 100 del escáner intraoral de acuerdo con una realización de esta divulgación. Comprende una fuente 110 de luz LED, una lente 111, un patrón 130 (una línea en una vista en corte transversal real, pero que se muestra aquí en ángulo para mayor claridad), un divisor 140 de haz y un sensor 180 de imagen. El escáner es un escáner de enfoque debido a una lente móvil 150. Un espejo 170 que dobla la trayectoria de haz hacia una parte de la cavidad intraoral que se está escaneando, que comprende un diente 301 y una encía 302. También se muestran como líneas discontinuas algunos rayos de luz emitidos desde la fuente 110 de luz, transmitidos a través del sistema óptico a una ubicación en la superficie de diente a la que se enfocan, regresan a través del sistema óptico, y se obtienen imágenes en el sensor 180, al que se enfocan también. Para la posición de la lente 150 de enfoque mostrada en la figura 2, es visible parte de la encía 302, pero sin enfocarse. Podría enfocarse para otras posiciones de la lente 150.

Las figuras 3 y 4 muestran cómo se procesan los datos del escáner portátil 100 ejemplar para dar lugar a mapas de distancia. La figura 3(a) es una vista de un patrón 130 de tablero de ajedrez de período espacial según es visto por el sensor 180 de imagen cuando la imagen de ese período espacial en la superficie escaneada está enfocada. En el caso extremo de una imagen completamente desenfocada del patrón, el sensor de imagen vería un gris uniforme. Es ventajoso si los píxeles del sensor de imagen y el patrón están alineados de tal manera que los bordes del patrón se obtengan en imágenes en los bordes de los píxeles. Como muestra la figura 3(b), a los píxeles que se espera que sean oscuros cuando la superficie está enfocada se les asigna una ponderación $f(i) = -1$, mientras que a los que se espera que sean brillantes se les asigna una ponderación $f(i) = 1$, donde i es un índice de píxeles.

A medida que se mueve la lente de enfoque, se toman imágenes. Las posiciones s de la lente de enfoque para cada imagen se encuentran mediante un codificador y una interpolación adecuada si es necesario. Para todos los grupos de píxeles en la imagen para la posición p, se calcula una medida de correlación como

$$A(s) = \sum_{i=1}^{n} f(i)I(i)$$

donde *I(i)* son las intensidades medidas en los píxeles. Para el ejemplo de la figura 3(b), *i* va de 1 a n = (6 * 6) = 36. Obsérvese que hay típicamente muchos períodos espaciales en un patrón típicamente rectangular 130 y su imagen en el sensor 180, por ejemplo N = 100 * 120 = 12.000 grupos de píxeles, de modo que la imagen constaría de al menos 36 * 12.000 píxeles en ese ejemplo.

Un grupo de píxeles está enfocado cuando A está en su máximo en todos los s, es decir, en todas las imágenes de una serie obtenidas dentro de una pasada de la lente de enfoque. La ubicación de ese máximo determina entonces la distancia *z(s)* a la superficie medida, porque la óptica del escáner se conoce desde la construcción, y, por consiguiente, se conoce la ubicación del plano de enfoque para todos los grupos de píxeles y todas las posiciones de la lente 150. La distancia z *(s)* se puede también encontrar o refinar mediante calibración. Con el plano y algún punto del sensor de imagen definiendo un sistema de coordenadas relativo al escáner, y cada grupo de píxeles teniendo *(x, y)* en ese sistema de coordenadas, tal como el centro del grupo de píxeles, la distancia *z(s)* para un grupo de píxeles dado da lugar a un punto *z(x, y)*. Como todos los grupos de píxeles tienen el mismo tamaño, todas las ubicaciones *(x, y)* de una cuadrícula.

Obsérvese que, si un grupo de píxeles está completamente desenfocado en alguna posición s, es decir, que todos los píxeles tienen el mismo valor, *A* = 0 en esa s. Obsérvese también que es posible que algunos grupos de píxeles nunca se enfoquen, por ejemplo cuando no hay ninguna superficie para obtener imágenes o cuando existe una superficie pero está fuera del intervalo de enfoque. No es posible determinar una distancia para tales grupos de píxeles.

La lente de enfoque tiene que moverse rápidamente de modo que un grupo de píxeles dado represente al menos aproximadamente la misma región de la superficie escaneada incluso en presencia de algún movimiento de la mano. Por ejemplo, la lente de enfoque puede pasar a través de todas las s con una frecuencia de ciclo de 10 Hz, viajando hacia adelante y hacia atrás, por tanto, con 20 pasos por segundo. Al mismo tiempo, el número de imágenes durante un pase debe ser bastante alto para dar como resultado una buena resolución de medición de distancia, tal como 200. Esto significa que el sensor de imagen debe ser bastante rápido; en este ejemplo, necesitaría tomar imágenes a una frecuencia de 10 Hz * 2 * 200 = 4000 Hz.

La figura 4 muestra dos curvas ejemplares A(s) para un grupo de píxeles. Ambos tienen su máximo en la misma ubicación s = 10. La curva A1 tiene un máximo más distinto que A2 en el sentido de que ambos valores máximos son mayores y el ancho de la zona distinta de cero es menor para A1 que para A2. De este modo, una medición de distancia en base a la curva A1 puede considerarse más cierta que una medición en base a la curva A2. Puede haber varias razones para que las mediciones de distancia sean menos ciertas. Algunas se deben a propiedades de la superficie escaneada, por ejemplo a que esa superficie esté inclinada con respecto al ángulo de visión, o a que tenga una reflectividad variable. El movimiento de la mano a menudo también conduce a mediciones de distancia menos ciertas. La certeza podría cuantificarse, por ejemplo, como el valor máximo, el logaritmo del valor máximo, el recíproco del ancho de la zona en la que A es superior a la mitad de su valor máximo, o similar.

En resumen, el escáner ejemplar proporciona mapas de distancia con un valor *z* por grupo de píxeles con coordenadas *(x, y)* con certeza asociada *q*, o no definidas. La combinación de *(z, q) (x, y)* puede denominarse mapa de distancia aumentada, análogo a una imagen con dos canales. Luego, se obtiene una representación 3D completa de la parte escaneada de la cavidad intraoral unificando los mapas de distancia aumentada obtenidos con diversas posturas del escáner.

Otros tipos de escáneres 3D incluyen escáneres láser 3D de triangulación y escáneres 3D de luz estructurada. Un escáner láser 3D de triangulación usa luz láser para sondear el ambiente o el objeto. Un láser de triangulación apunta con un láser al objeto y usa una cámara para buscar la ubicación del punto láser. Dependiendo de la distancia a la que incide el láser sobre una superficie, el punto láser aparece en diferentes lugares del campo de visión de la cámara. Esta técnica se llama triangulación porque el punto láser, la cámara y el emisor láser forman un triángulo. Se puede usar una franja láser, en lugar de un único punto láser, que, luego, se barre por el objeto para acelerar el proceso de toma.

Los escáneres 3D de luz estructurada proyectan un patrón de luz sobre el objeto y observan la deformación del patrón en el objeto. El patrón puede ser unidimensional o bidimensional. Un ejemplo de un patrón unidimensional es una línea. La línea se proyecta sobre el objeto usando, por ejemplo, un proyector LCD o un láser de barrido. Una cámara, ligeramente desplazada del proyector de patrones, observa la forma de la línea y usa una técnica similar a la triangulación para calcular la distancia de cada punto de la línea. En el caso de un patrón de una sola línea, la línea se barre a lo largo del campo de visión para reunir información de distancia una franja por vez. Otros principios del escáner 3D son bien conocidos en la técnica.

Para ver un ejemplo de cómo llegar a un conjunto de tipos de superficie, se hace referencia a la figura 5a. La figura muestra en una vista en corte transversal la boca inferior, con un diente incisivo 600, su margen gingival 610, encía

630 y labio inferior 650. Un posible conjunto de tipos de superficie estaría simplemente de acuerdo con estos tres tipos de tejido. Sin embargo, para un mejor entrenamiento, puede resultar ventajoso reducir el número de tipos de superficies en el conjunto. Tomando como ejemplo un tipo de aplicación de uso general que tiene como objetivo grabar dientes y encía cercana, se puede agrupar "diente" con algo de tejido cerca de la encía hasta, por ejemplo, el punto 620, siendo el resto un segundo tipo de superficie. El primero es "deseable" para unificar, el segundo es "indeseable". El punto 620 (una línea en 3D) no necesita tener un significado anatómico, pero podría, por ejemplo, establecerse desde la extensión del campo de visión del escáner cuando el escáner está obteniendo imágenes de al menos una parte de un diente. Cualquier superficie de un instrumento dental, de un dedo, de un algodón o de otros artefactos también se agrupa en el tipo "indeseable" en este ejemplo.

En la figura 5b se muestra un ejemplo de un procedimiento para entrenar un algoritmo de aprendizaje automático de acuerdo con una realización de esta divulgación.

En el paso 501, el escáner se usa para crear una representación digital 3D de una parte de una cavidad intraoral grabando múltiples mapas de distancia aumentada que se unifican, como se explica para las figuras 1 a 4. Algunas vistas y por consiguiente algunos mapas de distancia aumentada podría contener un instrumento dental, un dedo, un algodón u otros artefactos. Estos artefactos pueden representarse al menos parcialmente en la representación digital 3D, lo que hace que el modelo sea una representación relativamente peor de la cavidad intraoral.

En el paso 502, porciones de la representación digital 3D están anotadas por el tipo de superficie como se definió anteriormente para la figura 5a. La anotación puede ser realizada, por ejemplo, por un ser humano en una interfaz gráfica de usuario. La anotación también podría estar semiautomatizada, y el ser humano recibiría sugerencias obtenidas con otros métodos, por ejemplo algoritmos tradicionales de análisis de imágenes.

En el paso 503, las porciones anotadas se proyectan de nuevo a los mapas de distancia aumentada individuales, usando las inversas de las transformaciones encontradas en la unificación. Esto proporciona un mapa $c(x,y)$ de clase para cada mapa de distancia aumentada, donde c es un indicador de clase. Debido a que la anotación supuestamente proporciona información perfecta, se usa codificación en caliente para llegar a un vector de puntuaciones $\boldsymbol{p}$ de una parte de un mapa de distancia aumentada que pertenece a cualquiera de los tipos de superficie del paso 502. Por consiguiente, en $\boldsymbol{p}$, el elemento para la clase c se establece en 1 y todos los demás se establecen en cero. En el ejemplo usado aquí, $\boldsymbol{p}$ tiene dos elementos, uno para "deseable" y otro para "indeseable".

Los pasos 501-503 se realizan para muchas grabaciones similares R_1, R_2, ..., R_m, tales como al menos $m = 2$ grabaciones. El número de grabaciones m también podría ser, por ejemplo, de 10, 20, 50, 100, 200, 500 o 1000, o cualquier número de grabaciones entre éstos o superior. Las grabaciones son similares con respecto a la parte de la cavidad intraoral que se escanea, y pueden también ser similares con respecto a los artefactos en algunas vistas, afectando de manera análoga a las m representaciones digitales 3D creadas a partir de las m grabaciones.

En el paso 504, se entrena una red neuronal para predecir $\boldsymbol{p}(x, y)$ para $(z, q)\ (x, y)$. La incertidumbre q se toma como el logaritmo del valor máximo de A, como se explica en la figura 4. A medida que el escáner genera un mapa de distancia definido en una cuadrícula rectangular como se explicó anteriormente, el mapa de distancia aumentada $(z, q)\ (x, y)$ puede interpretarse como una imagen rectangular de dos canales que se convierte en la entrada a la red neuronal.

Obsérvese que los pasos 501 - 504 podrían también realizarse para mapas de distancia única, para mapas de distancia única aumentada, ambos casos dando como resultado la no unificación y retroproyecciones triviales, o para mapas de distancia múltiples, si bien es preferible efectuarlos para múltiples mapas de distancia aumentada, ya que esa constelación proporciona la mayor parte de la información. Para un escáner que también proporciona datos de color, el entrenamiento y la predicción podrían formularse como $\boldsymbol{p}(x, y)$ para $(z, q, r, g, b)\ (x, y)$, donde r, g, b son los componentes rojo, verde y azules de una medición de color; de manera análoga para otros datos o datos adicionales proporcionados por el escáner.

La figura 6 muestra un ejemplo de arquitectura de una red neuronal convolucional adecuada, con:

I: la imagen de entrada con ancho 120 y alto 100 (los valores de ejemplo usados en la explicación de las figuras 3 y 4) y dos canales para z y q, respectivamente.

P: relleno de ceros con un borde de 15 píxeles de ancho alrededor de la imagen.

C: una capa convolucional seguida de una unidad lineal rectificada.

CR: una capa de recorte para reducir el tamaño y adaptarse a la operación de la capa posterior

B: una capa de mejora de escala, usando un factor 2 e interpolación bilineal

M: una capa de agrupación máxima

+: una capa para la adición por elementos de canales correspondientes

O: la imagen de salida con un número de canales igual a la dimensión de ***p*** (dos en el ejemplo usado aquí)

Las dimensiones de los datos en las diversas capas y operaciones también se muestran en la figura 6. Son posibles muchas arquitecturas de red. Preferiblemente, la red neuronal tiene capas convolucionales con núcleos pequeños, tales como 3 x 3 x M, donde M es el número de canales. El impacto de entorno de puntos de superficie se deriva del uso de capas convolucionales. El impacto de entornos más amplios se deriva del uso de capas de agrupación máxima.

Debido a que está definida para el mismo *(x, y)* que la imagen de entrada, la imagen de salida tiene el mismo ancho y alto que la imagen de entrada, por lo que para cada punto de datos de superficie de entrada, hay una salida ***p***.

La figura 7 (a) muestra un primer ejemplo de cómo un algoritmo de aprendizaje automático entrenado se aplica en modo de inferencia.

En el paso 700, se elige un conjunto de tipos de superficie para predecir. En el paso 700 sólo se pueden elegir los tipos de superficie definidos durante el entrenamiento, en el paso 502, pero posiblemente se puedan agrupar algunos tipos de entrenamiento. Para el ejemplo que nos ocupa, se supone que los conjuntos de entrenamiento e inferencia son los mismos, con los dos tipos de superficie “deseable” e “indeseable”. Sin embargo, se puede usar cualquier número de tipos de superficie definidos durante el entrenamiento.

En el paso 701, se graba un mapa *(z, q) (x, y)* de distancia aumentada.

En el paso 702, el vector ***p*** *(x, y)* de puntuación para *(z, q) (x, y)* se infiere para todos los puntos *(x, y)* en el mapa de distancia aumentada grabado. El algoritmo de aprendizaje automático proporciona la inferencia. Para el ejemplo de red neuronal convolucional de la figura 6, ***p*** *(x, y)* viene dado por los valores con coordenadas *(x, y)* en la imagen **O** de salida multicanal; es decir, que cada canal proporciona un elemento del vector ***p***.

En el paso 703, se encuentra una ponderación *w* para ***p*** *(x, y)* en la unificación a partir de ponderaciones de tipo de superficie y de una función de puntuaciones, por ejemplo un producto escalar.

$w = \boldsymbol{w}\, g(\boldsymbol{p}(x,y))$

donde ***w*** es un vector de ponderaciones de tipo superficie y ***g*** es un vector de evaluaciones de una función *g* de las puntuaciones. Por ejemplo, ***w*** podría elegirse como [1, 0] (refiriéndose, el primer elemento, al tipo de superficie “deseable”). La función *g* podría elegirse, por ejemplo, como

$$g(p(x,y)) = \begin{cases} 1 \text{ si } p_{\text{deseable}}(x,y) > 0{,}6 \\ 0 \quad \text{en caso contrario} \end{cases}$$

Se podrían usar valores distintos de 0,6 dependiendo de la preferencia, preferiblemente valores superiores a 0,5, pero en principio se podrían también usar valores inferiores a 0,5. En el paso 704, los puntos *z (x, y)* se unifican con la representación digital 3D construida a partir de vistas anteriores, ponderados de acuerdo con sus ponderaciones encontradas en el paso 703. Por ejemplo, se usa un algoritmo ICP estándar para la unificación, y todos los puntos con $w > 0$ se incluyen en el problema de minimización subyacente. Desde la primera vista, la representación digital 3D se establece en los puntos *z (x, y)*.

Luego, el procedimiento puede repetirse desde el paso 701 para vistas adicionales, típicamente tomadas desde diferentes posturas a medida que el operario mueve el escáner, o puede terminar, típicamente si el operario decide que la representación digital 3D está completa. Esa primera representación digital 3D se convierte a menudo en una segunda, en una malla.

La figura 7 (b) muestra un segundo ejemplo de cómo se aplica un algoritmo de clasificación de aprendizaje automático entrenado en modo de inferencia.

Los pasos 700 a 702 son como en el primer ejemplo de la figura 7 (a).

En el paso 714, todos los puntos *z (x, y)* se unifican con una representación digital 3D provisional construida a partir de vistas anteriores, y también lo hacen los valores de ***p*** *(x, y)*, recibiendo, de este modo, coordenadas espaciales en el espacio de la representación digital 3D como *(X, Y, Z)* y ***p*** *(X, Y, Z)*, respectivamente. Con fines prácticos, el espacio 3D del modelo virtual provisional se representa con vóxeles y los valores de ***p*** *(X, Y, Z)* se asignan al vóxel más cercano. En cada vóxel, los valores de ***p*** se suman en las vistas y se mantiene un recuento de cuántos valores se añaden, de tal modo que, por ejemplo, se pueda calcular un promedio más adelante.

El procedimiento puede entonces repetirse desde el paso 701 para vistas adicionales, típicamente tomadas desde diferentes posturas a medida que el operario mueve el escáner, o puede terminar, típicamente si el operario decide que la representación digital 3D provisional está completa.

En el paso 713, las ponderaciones se encuentran de manera análoga al paso 703, pero, para el promedio ***p***, ese promedio se encuentra dividiendo la suma por el número de valores. Los vóxeles con ponderación 0 se filtran, al igual que los vóxeles sin ningún valor.

En el paso 715, se construye un modelo virtual final a partir de los puntos *(X, Y, Z)* en los vóxeles restantes, usando, por ejemplo, el algoritmo de cubos en marcha.

La figura 8 muestra un ejemplo de cómo el algoritmo de la figura 7 (b) se expande para usar también filtrado en base al volumen excluido. Para empezar, el volumen excluido está vacío.

Los pasos 700 - 702 y 714 son como se explicaron anteriormente para la figura 7.

En el paso 814, el volumen excluido se actualiza con los puntos unificados *(X, Y, Z)* del paso 714. Un volumen excluido es, por ejemplo, el espacio desde una superficie vista hasta el escáner y el cuerpo de escáner. Con fines prácticos, la exclusión podría expresarse en el mismo espacio de vóxel que se usa en el paso 714. Una marca binaria puede indicar si un vóxel es parte del espacio excluido.

El procedimiento puede repetirse entonces desde el paso 701 para vistas adicionales como se describió anteriormente. A medida que se graban más vistas, la cantidad de vóxeles marcados como estando en el espacio excluido puede crecer, pero nunca disminuir. El paso 713 es como se explicó anteriormente para la figura 7.

En el paso 813, se borran todos los vóxeles que todavía contienen valores de ***p*** sumado, pero que están en el volumen excluido. Este paso es más fácil de implantar si los espacios de vóxeles son idénticos; de lo contrario, se puede usar una búsqueda de vecino más próximo para encontrar el vóxel más cercano en el espacio de vóxel de volumen excluido.

El paso 715 es como se explicó anteriormente para la figura 7.

Aunque se han descrito y mostrado en detalle algunas realizaciones y ejemplos, se debe entender que se pueden realizar modificaciones estructurales y funcionales sin apartarse del alcance de la presente invención, definida por el conjunto de reivindicaciones adjunto.

Se debe enfatizar que el término “comprende/que comprende” cuando se usa en esta especificación se toma para especificar la presencia de características, números enteros, pasos o componentes indicados/as, pero no excluye la presencia o adición de uno/a o más diferentes características, números enteros, pasos, componentes o grupos de estos elementos.

Bibliografía

[1] Khoshelham K, Dos Santos DR, Vosselman G., “Generation and weighting of 3D point correspondences for improved registration of RGB-D data”, Procedimientos de ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 11 de noviembre de 2013, 5, T2.

[2] Kazhdan M, Hoppe H., “Screened Poisson surface reconstruction”, ACM Transactions on Graphics (TOG), 1 de junio de 2013, 32(3), 29.

[3] Newcombe RA, Izadi S, Hilliges O, Molyneaux D, Kim D, Davison AJ, Kohi P, Shotton J, Hodges S, Fitzgibbon A., *KinectFusion:* “Real-time dense surface mapping and tracking”, International Symposium on Mixed and Augmented Reality (ISMAR), 2011, Décimo simposio internacional IEEE del 26 de octubre de 2011, págs. 127-136.

[4] Long J, Shelhamer E, Darrell T., “Fully convolutional networks for semantic segmentation”, Procedimientos de la Conferencia IEEE sobre visión por ordenador y reconocimiento de patrones de 2015, págs. 3431-3440.

REIVINDICACIONES

1. Un método para generar una representación digital 3D de al menos una parte de una cavidad intraoral, comprendiendo el método:

- grabar una pluralidad de vistas que contienen datos de superficie que representan al menos la geometría de los puntos de superficie de la parte de la cavidad intraoral usando un escáner intraoral;

- determinar una ponderación para cada punto de superficie al menos parcialmente en base a puntuaciones que son medidas de creencia de que ese punto de superficie representa un tipo particular de superficie;

- ejecutar un algoritmo de unificación que realiza una unificación ponderada de los puntos de superficie en dicha pluralidad de vistas para generar la representación digital 3D en base a las ponderaciones determinadas;

- en el que las puntuaciones para los puntos de superficie se encuentran mediante al menos un algoritmo de búsqueda de puntuaciones que toma como entrada al menos la parte de geometría de los datos de superficie para ese punto de superficie y datos de superficie para puntos en un entorno de ese punto de superficie;

caracterizado porque el escáner también suministra alguna información de certeza de datos de superficie medidos para las vistas grabadas, y dicha información de certeza modifica al menos parcialmente las puntuaciones.

2. El método de acuerdo con la reivindicación 1, en el que el al menos un algoritmo de búsqueda de puntuación es un algoritmo de aprendizaje automático.

3. El método de acuerdo con la reivindicación 2, en el que el al menos un algoritmo de aprendizaje automático comprende una red neuronal con al menos una capa convolucional.

4. El método de acuerdo con la reivindicación 2 ó 3, en el que el al menos un algoritmo de aprendizaje automático fue entrenado en una pluralidad de los tipos de superficies que se graban comúnmente con escáneres en cavidades intraorales.

5. El método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que los datos de superficie comprenden también información de color.

6. El método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que al menos un algoritmo de aprendizaje automático fue entrenado al menos parcialmente usando datos grabados por un escáner intraoral antes de la generación de la representación digital 3D.

7. El método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que al menos un algoritmo de aprendizaje automático fue entrenado al menos parcialmente por un operario de un/el escáner intraoral.

8. El método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que se aplican otros algoritmos para filtrar datos desde las vistas grabadas, en el que uno de dichos otros algoritmos evalúa la coherencia geométrica en una pluralidad de vistas.

9. Un producto de programa de ordenador que comprende instrucciones que, cuando el programa es ejecutado por un ordenador, hacen que el ordenador lleve a cabo los pasos del método de acuerdo con cualquiera de las reivindicaciones anteriores.

10. Un sistema de escáner para generar una representación digital 3D de al menos una parte de una cavidad bucal, comprendiendo el sistema de escáner:

- un escáner intraoral portátil; y

- una o más unidades de procesamiento configuradas para realizar los pasos del método de acuerdo con cualquiera de las reivindicaciones 1-8.

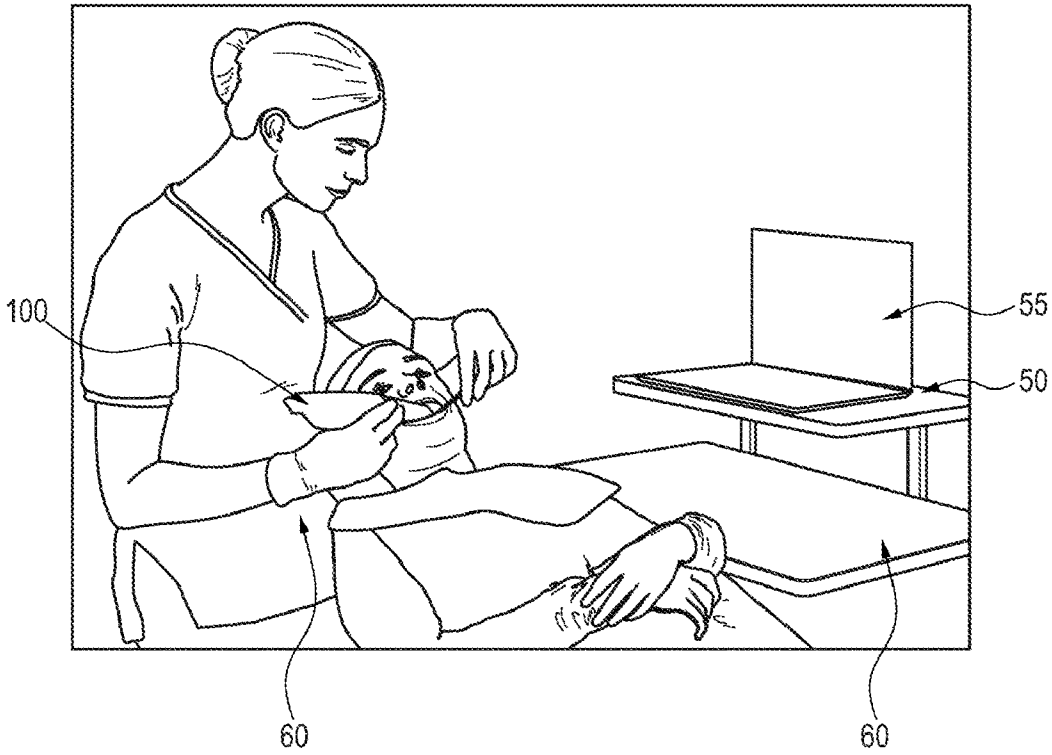


Fig. 1

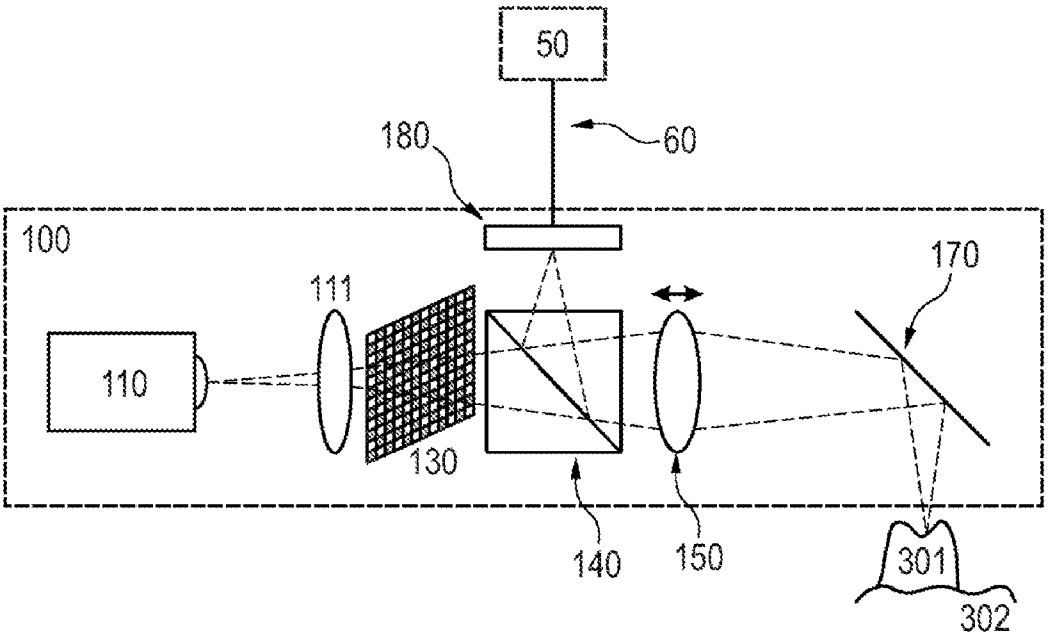
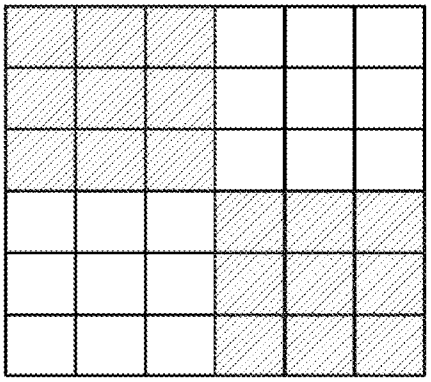


Fig. 2



(a)

-1	-1	-1	1	1	1
-1	-1	-1	1	1	1
-1	-1	-1	1	1	1
1	1	1	-1	-1	-1
1	1	1	-1	-1	-1
1	1	1	-1	-1	-1

(b)

Fig. 3

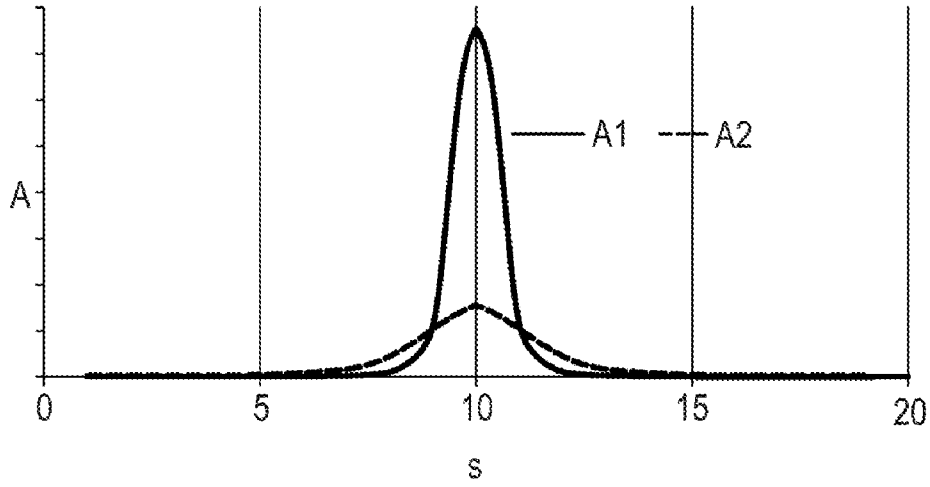


Fig. 4

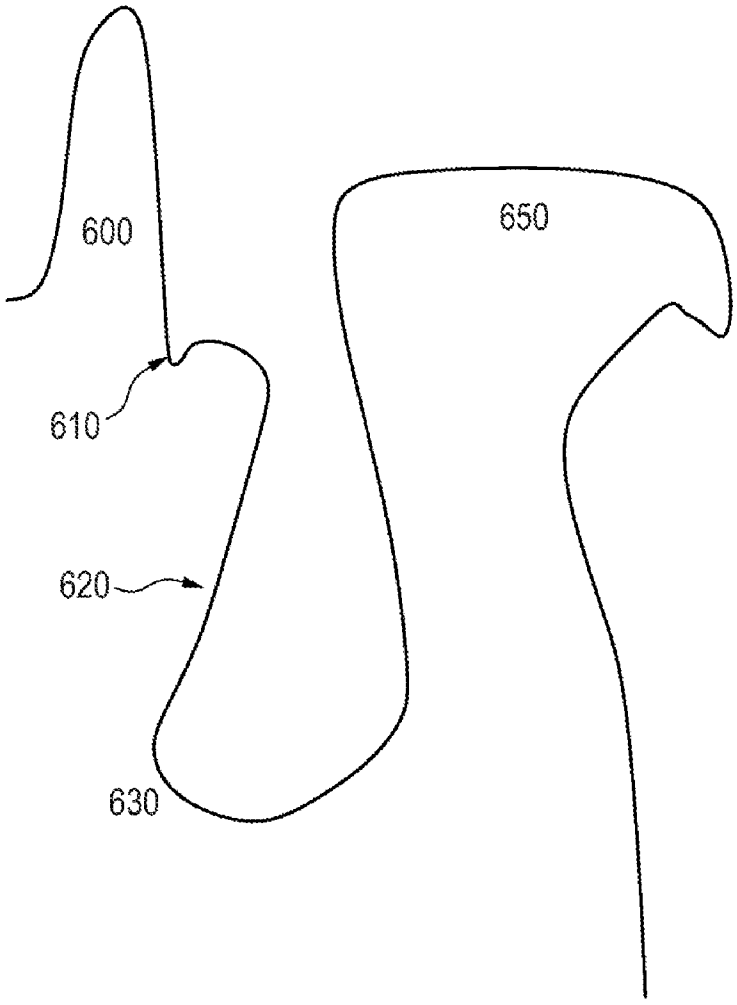


Fig. 5a

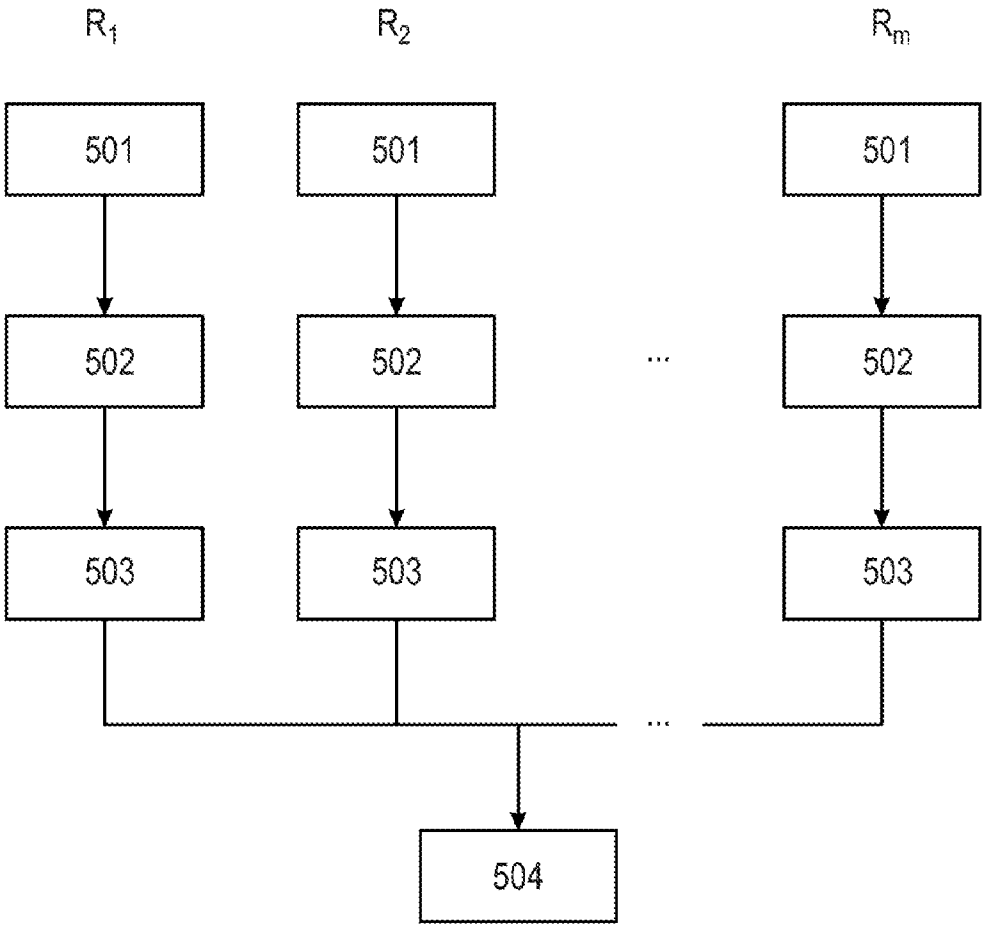


Fig. 5b

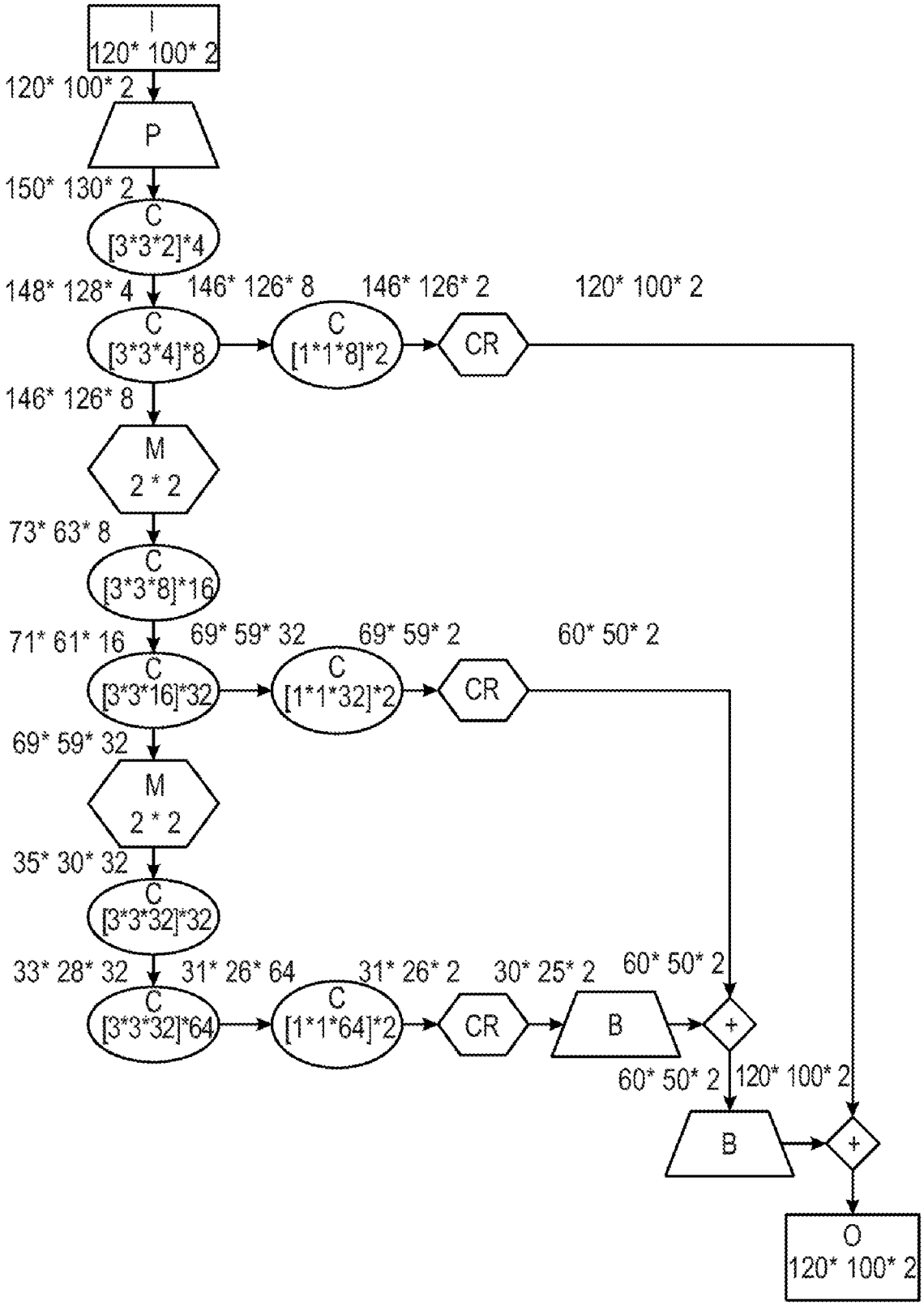

I
120* 100* 2
120* 100* 2
P
150* 130* 2
C
[3*3*2]*4
148* 128* 4
146* 126* 8
146* 126* 2
120* 100* 2
C
[3*3*4]*8
C
[1*1*8]*2
CR
146* 126* 8
M
2 * 2
73* 63* 8
C
[3*3*8]*16
71* 61* 16
69* 59* 32
69* 59* 2
60* 50* 2
C
[3*3*16]*32
C
[1*1*32]*2
CR
69* 59* 32
M
2 * 2
35* 30* 32
C
[3*3*32]*32
33* 28* 32
31* 26* 64
31* 26* 2
30* 25* 2
60* 50* 2
C
[3*3*32]*64
C
[1*1*64]*2
CR
B
+
60* 50* 2
120* 100* 2
B
+
O
120* 100* 2

Fig. 6

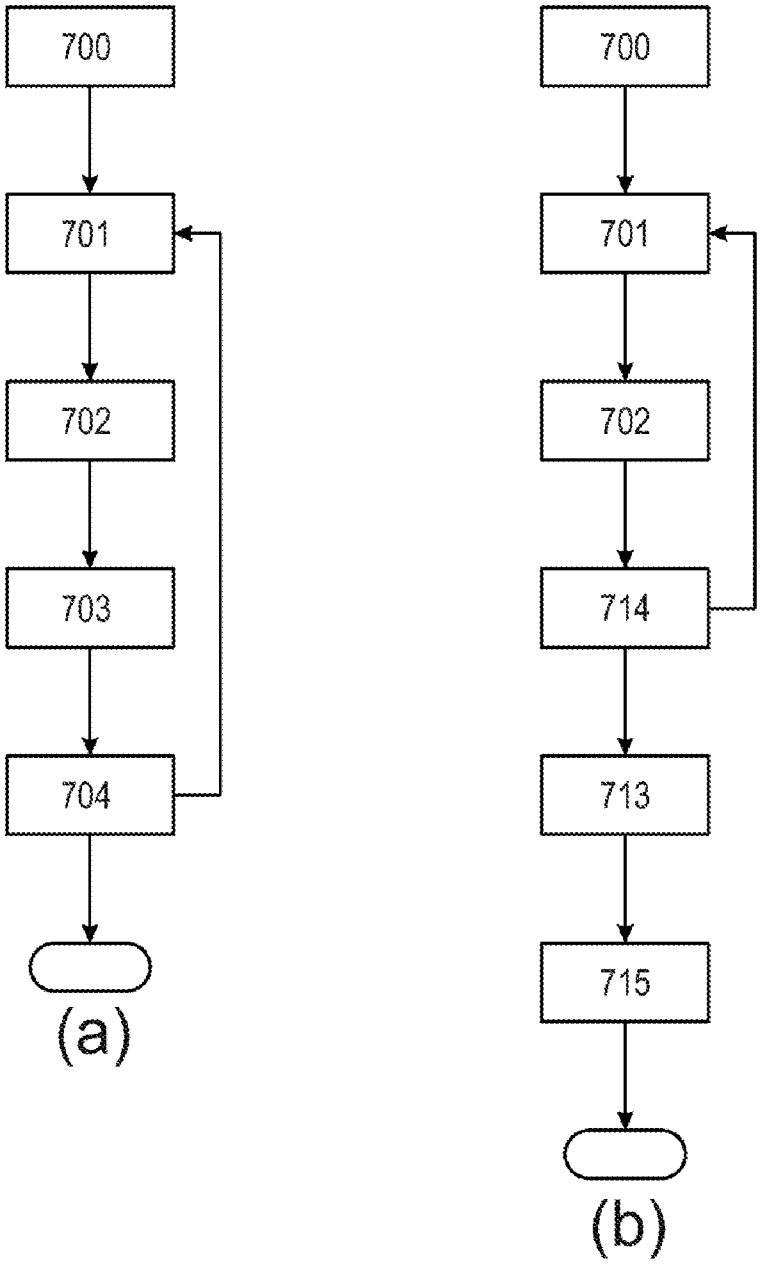


Fig. 7

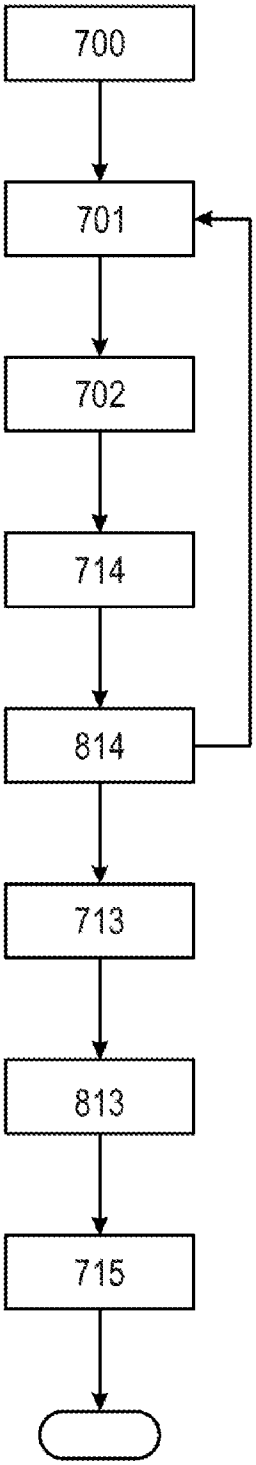


Fig. 8