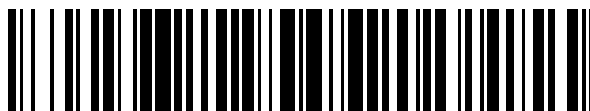


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 382 832**

51 Int. Cl.:
G06T 3/40 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **08100059 .8**
- 96 Fecha de presentación: **03.01.2008**
- 97 Número de publicación de la solicitud: **2045773**
- 97 Fecha de publicación de la solicitud: **08.04.2009**

54 Título: **Método y aparato para cambiar la resolución espacial de una imagen digital**

30 Prioridad:
02.10.2007 US 976999 P

45 Fecha de publicación de la mención BOPI:
13.06.2012

45 Fecha de la publicación del folleto de la patente:
13.06.2012

73 Titular/es:
**VESTEL ELEKTRONIK SANAYI VE TICARET A.S.
ORGANIZE SANAYI BÖLGESİ
45030 MANISA, TR**

72 Inventor/es:
**Akgun, Toygar;
Sezer, Osman Gokhan y
Altunbasak, Yucel**

74 Agente/Representante:
de Elzaburu Márquez, Alberto

ES 2 382 832 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

Método y aparato para cambiar la resolución espacial de una imagen digital.

5 La presente invención se refiere a un método y a un aparato para cambiar la resolución espacial de una imagen digital.

La presente invención se refiere, en general, al escalado y a la mejora de la resolución de una imagen digital. Las realizaciones preferentes permiten que las imágenes digitales sean escaladas con una cantidad mínima de borrosidad y bordes puntiagudos. Las técnicas pueden ser incluidas en cualquier sistema que necesite cambiar el tamaño de una imagen digital de entrada o una secuencia de imágenes digitales. Una aplicación práctica son los aparatos de televisión digital en los que una imagen de entrada debe ser visualizada a un tamaño diferente (normalmente, un tamaño mayor) de su tamaño original, sin perder detalles espaciales y sin introducir bordes puntiagudos. Otras aplicaciones incluyen impresoras en las que se requiere cambiar el tamaño de una imagen de entrada para imprimirla en un tamaño deseado (normalmente mayor), en las cámaras digitales, para permitir que una imagen capturada sea visualizada a una resolución diferente (normalmente mayor) y, en general, en el procesamiento digital de imágenes, tal como en imágenes médicas, para aumentar la resolución de la imagen, etc.

Se conocen diversas técnicas de escalado de imagen en la técnica anterior. Muchas de estas técnicas son del tipo denominado interpolador de B-spline. Los más simples de los interpoladores B-spline son los de orden cero y primero. Estos se conocen, respectivamente, como replicación de píxel e interpolación bilineal. En la replicación de píxeles, cada píxel en la imagen de salida de alta resolución se obtiene tomando el valor del píxel más cercano en la imagen de entrada de baja resolución. En la interpolación bilineal, cada píxel en la imagen de salida de alta resolución se obtiene calculando una combinación lineal de hasta cuatro píxeles en la imagen de entrada de baja resolución. Los interpoladores de orden más alto usan técnicas más sofisticadas, pero son computacionalmente intensivos.

Un enfoque diferente, basado en la clasificación, se divulga en la "Optimal Image Scaling Using Pixel Classification" por C.B. Atkins, C.A. Bouman y J.P. Allebach, en Proceedings of the 2001 International Conference on Image processing, 7-10 de octubre de 2001, volumen 3, página (s): 864 - 867 y, con más detalle, en la tesis doctoral de C.B. Atkins, titulada "Classification-base Methos in Optimal Image Interpolation". Las Figuras 1 y 2 muestran esquemáticamente la técnica divulgada por Atkins y col.

Haciendo referencia, en primer lugar, a la Figura 1, para cada píxel 1 de entrada en la imagen 2 de entrada de baja resolución, en este ejemplo una ventana 3 de 5x5 de píxeles en la imagen 2 de entrada, de baja resolución, está centrada en el píxel 1 de entrada. Esta ventana 5x5 es vectorizada en columnas en un vector Z de columna 25x1 (denominado vector Z de observación en los documentos Atkins). A continuación, en un bloque 11 extractor de características, se aplica un operador f de proyección de transformación no-lineal al vector Z de observación, para obtener un vector de agrupación o vector Y de características, es decir, $Y = f(Z)$ para cada píxel 1 de entrada. En términos generales, Y es de menor dimensión que Z. Esta asignación desde el vector Z de observación al vector Y de características afecta a la calidad de la imagen final, de alta resolución, interpolada. Son posibles un número de operadores de proyección f diferentes.

A continuación, el vector Y de características para cada píxel 1 de entrada es pasado a un bloque 12 clasificador en el que el vector Y de características para cada píxel 1 de entrada en la imagen 2 de entrada de baja resolución está asociado con (o "clasificado" o "agrupado a") un número limitado de clases de contexto. Estas clases de contexto modelan diferentes tipos de regiones de imagen en la imagen 2 de entrada, tales como bordes, regiones suaves, regiones texturadas, etc. Para este propósito, los parámetros θ de distribución son suministrados al bloque 12 clasificador, y el grado de asociación del vector Y de características con las clases individuales se determina, entre otras cosas, según los parámetros θ de distribución. Atkins et al proponen un marco estadístico para atribuir el píxel 1 de entrada a las clases de contexto. La clasificación de contexto es realizada comparando la distancia del vector de características del píxel de entrada de baja resolución a los prototipos de las clases de contexto. (El "prototipo" de una clase de contexto es el punto o vector en la clase de contexto que mejor representa a los miembros de esa clase de contexto). Para cada píxel 1 de entrada, el bloque 12 clasificador calcula y pone en su salida un número entre 0 y 1 para cada una de las clases de contexto, siendo estos números coeficientes de ponderación y siendo su suma igual a 1. Los coeficientes de ponderación asignados a cada clase por el bloque 12 clasificador son las probabilidades de que el píxel 1 de entrada pertenezca a las clases respectivas, es decir, w_i es la probabilidad de que el píxel 1 de entrada pertenezca a la clase i . Los parámetros θ de distribución se obtienen durante un procedimiento de entrenamiento, que será explicado más adelante.

Una vez que el vector Y de características para cada píxel 1 de entrada ha sido asociado con las clases de contexto, un bloque 13 de filtro lineal es usado para obtener los píxeles 5 de salida de la imagen 6 de salida de alta resolución. En particular, para cada píxel 1 de entrada de la imagen 2 de entrada de baja resolución, una ventana 3 de píxeles de 5x5, centrada en el píxel 1 de entrada, es pasada al bloque 13 de filtro lineal. El bloque 13 de filtro lineal toma como una entrada los coeficientes de ponderación de las clases de contexto, calculados por el bloque 12 clasificador para el píxel 1

de entrada particular y toma como otra entrada los coeficientes ψ de filtro de interpolación, que corresponden a las clases contexto respectivas, y calcula un bloque $L \times L$ de píxeles de salida para la imagen de salida de alta resolución, mediante una combinación lineal de las salidas de los filtros de interpolación en proporciones según los coeficientes de ponderación calculados por el bloque 12 clasificador para el píxel 1 de entrada particular. (L es el factor de escala. En el ejemplo mostrado en la Figura 1, $L = 2$) Si esta ventana $L \times L$ de píxeles de alta resolución es barrida por rastreo en un vector X, entonces la interpolación se puede ser expresada como

$$X = \sum_{i=1}^M w_i (A_i Z + \beta_i)$$

en la que A_i y β_i son la matriz de interpolación y el vector de desviación de la i-ésima clase de contexto, w_i son los coeficientes de ponderación para los filtros/las clases de contexto respectivos, y M es el número de filtros/clases de contexto. A_i y β_i , conjuntamente, son los coeficientes del filtro de interpolación. Esta fórmula muestra que el valor final de un píxel de alta resolución, interpolado, se obtiene como una combinación lineal ponderada de los valores de interpolación calculados usando los coeficientes del filtro de interpolación de todas las clases de contexto. Los coeficientes de ponderación de la combinación se calculan usando los parámetros obtenidos durante el entrenamiento fuera de línea, expuesto más adelante. El coeficiente de ponderación de una clase de contexto aumenta a medida que el vector característica del píxel de entrada de baja resolución se acerca al prototipo. El coeficiente de ponderación de una clase de contexto se incrementa conforme el vector de características del píxel de entrada de baja resolución se acerca al prototipo.

En resumen, en esta técnica anterior, los píxeles 1 de entrada son "clasificados" o asociados a cada una de las clases de contexto de la imagen, en diferentes grados o "coeficientes de ponderación". La ventana X de píxeles de salida interpolados se obtiene filtrando, en primer lugar, el vector Z de observación con los coeficientes ψ óptimos del filtro de interpolación para las clases individuales y, a continuación, combinando los resultados en una mezcla de proporciones, tal como se determinan en la fase de clasificación. Los coeficientes ψ del filtro de interpolación, introducidos al bloque 13 de filtro lineal, corresponden a la matriz A de interpolación y al vector β de desviación y son obtenidos, de nuevo, durante el procedimiento de entrenamiento, que se describirá más adelante.

Evidentemente, de importancia clave en esta técnica son los parámetros θ de distribución, relacionados con la distribución de los píxeles entre las clases de contexto, y los coeficientes ψ del filtro de interpolación, que, esencialmente, determinan las mezclas de los filtros de interpolación que son sumadas entre sí para obtener los píxeles de salida de la imagen de alta resolución. En los documentos de Atkins, los parámetros θ de distribución y los coeficientes ψ del filtro de interpolación se obtienen en un método de formación fuera de línea, que es realizado, típicamente, solo una vez. En particular, las imágenes de alta resolución son suministradas y, a continuación, se obtienen imágenes de baja resolución a partir de las imágenes de alta resolución (mediante una o varias técnicas diferentes). Con referencia a la Figura 2, un vector Y de características respectivo es obtenido para cada píxel de entrada en las imágenes de baja resolución. Estos se pasan a un bloque 14 de estimación de parámetros de distribución, en el que se obtiene una estimación inicial de los parámetros θ de distribución. Al hacer varias suposiciones, los valores de los parámetros de distribución son actualizados para encontrar un valor final, óptimo. En particular, se calculan estimaciones de probabilidad máxima de los parámetros θ de distribución, usando un algoritmo de maximización de esperanza. Esta etapa se corresponde con una "agrupación blanda" de los píxeles de entrada en las clases de contexto. (En este contexto, una "agrupación dura", o usando una "decisión dura", significa que un píxel es atribuido a una única clase de contexto. Por otro lado, una "agrupación-blanda", o usando una "decisión blanda", significa que, para cada píxel, un número no negativo o "coeficiente de ponderación" entre 0 y 1 es asignado a cada clase de contexto: el coeficiente de ponderación de cada clase de contexto es la probabilidad de que el píxel en cuestión pertenezca a esa clase de contexto). A continuación, los coeficientes ψ del filtro de interpolación son obtenidos en un bloque 15 de diseño del filtro de interpolación. Los parámetros θ de distribución estimados son introducidos a este bloque 15 para calcular los filtros ψ de interpolación óptimos que minimizan el error cuadrático medio.

En este procedimiento, tal como se señala en los documentos de Atkins, una de las principales suposiciones que se hacen es que el vector Y de características proporciona toda la información acerca de un píxel de entrada en la imagen de baja resolución (de entrenamiento) y la manera en que está relacionado con el bloque correspondiente de píxeles en la imagen fuente, de alta resolución. Debido a esta suposición, que simplifica los cálculos, la estimación de los coeficientes ψ del filtro de interpolación es desacoplada, de manera efectiva, de la estimación de los parámetros θ de distribución. En realidad, esto hace que el filtro de interpolación calculado sea subóptimo, lo que a su vez implica una pérdida de posible calidad de la imagen de alta resolución, interpolada.

Aquí, se hace referencia a los documentos EP-A-1814075, WO-A-2007/088113 y US-A-2007-0172152, de los presentes inventores, en los que se divulga, entre otras cosas, un método de obtención de filtros de interpolación y, opcionalmente,

vectores de características, en el que durante una fase de entrenamiento, los filtros de interpolación y vectores de características son actualizados, lo que permite que los filtros de interpolación y los vectores de características sean optimizados de mejor manera, para su uso en la fase de interpolación real.

5 Yoakeim, R y Taubman, D.: "Interpolation Specific Resolution Synthesis" Proc. IEEE Int. Conf. on Image Processing, 2007, ICIP 2007, 16-19 de Septiembre, 2007, Vol. 4, páginas 201-204, divulga un enfoque para una síntesis de resolución acoplada específicamente al problema de interpolación. Esto se consigue alineando explícitamente la clasificación y las subsiguientes actividades de interpolación con los modelos de imagen que conducen a una interpolación. Los modelos de imagen están predeterminados en lugar de ser descubiertos y representan un intervalo de bordes de orientación, perfil y posición arbitrarios.

10 Akgun, T. y Altunbasak, Y.: "A Coupled Feature-Filter Clustering Scheme for Resolution Synthesis" Proc. IEEE Int. Conf. on Image Processing", 2007, ICP 2007, 16-19 Sept., vol. 5, páginas 405-408, divulga una síntesis de resolución que implica un esquema de entrenamiento iterativo, en el que los filtros de interpolación están acoplados a las etapas de extracción de características y clasificación, para mejorar la agrupación.

20 Según un primer aspecto de la presente invención, se proporciona un método para cambiar la resolución espacial de una primera imagen digital, que tiene una primera resolución espacial, para proporcionar una segunda imagen digital, que tiene una segunda resolución espacial, que es diferente de la primera resolución espacial, comprendiendo el método, para cada uno de entre una pluralidad de píxeles en la primera imagen digital, obtener un vector de características para un píxel en la primera imagen digital; clasificar dicho píxel de manera que dicho píxel es asignado a una o más clases de contexto, comparando el vector de características de dicho píxel con un prototipo de cada una de dichas clases de contexto, teniendo cada una de dichas clases de contexto uno o más filtros de interpolación asociados a la misma, y, filtrando dicho píxel usando el uno o más filtros de interpolación asociados con la una o más clases de contexto a las cuales está asignado dicho píxel, lo que genera uno o más píxeles de la segunda imagen digital; en el que dicho vector de características para dicho píxel es una función de las diferencias entre pares contiguos de píxeles seleccionados en la primera imagen digital, caracterizado porque al menos uno de dichos pares contiguos de píxeles en la primera imagen digital no incluye dicho píxel.

30 Este método de obtención de los vectores de características permite el uso de un número relativamente pequeño de clases de contexto (por ejemplo, aproximadamente 11 o menos), obteniendo, al mismo tiempo, una alta calidad de imagen. Esto, a su vez, reduce la complejidad de la etapa de interpolación, facilitando la implementación a un coste razonable en equipos domésticos o de consumo.

35 Cuando se hace mención, en el texto anterior, por ejemplo a "la diferencia entre píxeles", se entenderá que significa la diferencia entre los valores de ciertos componentes de los píxeles. Por ejemplo, cuando la imagen de entrada está en formato RGB, las diferencias pueden ser en los valores respectivos de los componentes R, G, y B de los respectivos píxeles. Por ejemplo, cuando la imagen de entrada está en formato YUV o YCrCb, las diferencias pueden ser en los valores respectivos de la luminancia y dos componentes de crominancia, o de hecho sólo algunos de estos componentes, tal como solo el componente de luminancia.

40 En una realización, el vector de características es una función sólo de las diferencias entre los pares contiguos seleccionados de píxeles en la primera imagen digital, en los que al menos uno de los pares contiguos de píxeles en la primera imagen digital no incluye dicho píxel.

45 En una realización, dichos pares de píxeles son píxeles en una ventana de un tamaño predeterminado alrededor de dicho píxel. Es preferente que la ventana sea relativamente grande, tal como una ventana de 5x5 alrededor del píxel que está siendo procesado. Esto permite que las características locales sean capturadas con precisión, incluso en presencia de una pequeña cantidad de solapamiento. Esto resulta en clasificaciones que tienen más probabilidades de ser correctas o apropiadas y conduce a menos artefactos visuales.

50 En una realización, dicha función es tal que cada una las diferencias entre los pares contiguos seleccionados de píxeles en la primera imagen digital es elevada a una potencia n , donde n es un número par. Esto hace que las orientaciones de bordes similares, con polaridades opuestas, en la imagen de entrada sean forzadas a la misma clase de contexto.

55 En una realización, cada elemento de cada vector de características es una función de la diferencia entre un par de píxeles contiguo seleccionado de píxeles en la primera imagen digital, siendo obtenido cada vector de características normalizando los elementos de dicho vector de características, obtenido de esta manera. Esto ayuda a suprimir los efectos del ruido en la imagen de entrada.

60 En una realización, al menos uno de los elementos de cada vector de características es una función de la diferencia entre

pares de píxeles en filas contiguas de la primera imagen digital. Esto puede ser usado, de manera efectiva, para medir la similitud de las líneas pares (y/o impares), y/o la diferencia entre las líneas pares e impares, en el interior de la ventana desde la que se extraen los vectores de características. Esto, a su vez, puede ser usado para identificar píxeles ubicados en los artefactos de des-entrelazado y, de esta manera, permite que el procesamiento tenga en cuenta, de manera particular, los artefactos de des-entrelazado y, de esta manera, reducir su efecto sobre la calidad de la imagen.

En una realización, la primera imagen digital está en un formato que tiene al menos un componente de luminancia y dos componentes de crominancia, en la que:

dichas diferencias entre al menos algunos de los pares contiguos de píxeles son las diferencias en los valores respectivos de los componentes de luminancia de dichos pares contiguos de píxeles.

En una realización, la primera imagen digital está inicialmente en formato RGB y es convertida a dicho formato, que tiene al menos un componente de luminancia y dos componentes de crominancia, en el que se identifican aquellas combinaciones de componentes RGB para las que pueden ocurrir bordes de isoluminancia cuando la imagen de formato RGB es convertida a dicho formato que tiene al menos un componente de luminancia y dos componentes de crominancia, en la que:

para al menos algunos píxeles de la primera imagen digital, que inicialmente tiene valores RGB correspondientes a aquellas combinaciones de componentes RGB para las cuales pueden ocurrir bordes de isoluminancia cuando la imagen de formato RGB es convertida a dicho formato que tiene al menos un componente de luminancia y dos componentes de crominancia, dichas diferencias entre al menos algunos de los pares contiguos de dichos píxeles son las diferencias en los valores respectivos de al menos uno de las componentes de crominancia de dichos pares contiguos de píxeles.

En una realización, la clasificación de dicho píxel se realiza de manera que dicho píxel es asignado a una única clase de contexto comparando el vector de características de dicho píxel con un prototipo de cada una de dichas clases de contexto; y el filtrado de dicho píxel usa un único filtro de interpolación que está asociado con dicha única clase de contexto a la cual está asignado dicho píxel. Esta clasificación "dura" durante la fase de interpolación reduce la complejidad de los cálculos durante la fase de interpolación, pero puede usarse en la presente realización preferente, sin sacrificar la calidad de imagen.

En una realización, la segunda resolución es mayor que la primera resolución y la proporción de la primera resolución a la segunda resolución no es un número entero, y los píxeles en la segunda imagen digital se obtienen, para cada uno de dichos píxeles en la segunda imagen digital:

identificando la ubicación en la primera imagen digital que corresponde a la ubicación de dicho píxel en la segunda imagen digital; identificando el píxel de la primera imagen digital que corresponde a dicha ubicación en la primera imagen digital; y, filtrando dicho píxel en la primera imagen digital con el único filtro de interpolación respectivo que corresponde a dicha ubicación en la primera imagen digital, para obtener, de esta manera, los píxeles en la segunda imagen digital. Esto permite la obtención de un escalado no entero desde una resolución a otra, además de un escalado entero, permitiendo cambios muy flexibles desde una resolución a otra, y evitando la limitación de alguna técnica anterior de poder conseguir solo proporciones de escalado enteras, manteniendo, al mismo tiempo, una buena calidad de imagen.

En una realización, los valores de los elementos de un vector de características para dicho píxel se obtienen a partir de una tabla de consulta. Esto puede proporcionar una manera rápida y eficiente para obtener los elementos de los vectores de características y evita tener que calcular los elementos "desde cero".

Según un segundo aspecto de la presente invención, se proporciona un aparato para cambiar la resolución espacial de una primera imagen digital, que tiene una primera resolución espacial, para proporcionar una segunda imagen digital, que tiene una segunda resolución espacial, que es diferente de la primera resolución espacial, comprendiendo el aparato un extractor de vector de características para obtener un vector de características para un píxel en la primera imagen digital; un clasificador para clasificar dicho píxel de manera que dicho píxel es asignado a una o más clases de contexto comparando el vector de características de dicho píxel con un prototipo de cada una de dichas clases de contexto, teniendo cada una de dichas clases de contexto uno o más filtros de interpolación asociados a la misma; y un filtro para filtrar dicho píxel usando uno o más filtros de interpolación asociados con la una o más clases de contexto a la que ha sido asignado dicho píxel, para generar, de esta manera, uno o más píxeles de la segunda imagen digital; donde el extractor de vectores de características está dispuesto de manera que dicho vector de características para dicho píxel es una función de las diferencias entre pares contiguos seleccionados de píxeles en la primera imagen digital, al caracterizado porque al menos uno de dichos pares contiguos de píxeles en la primera imagen digital no incluye dicho píxel.

El aparato y/o el método preferentes pueden ser incorporados en cualquier aparato y/o método que es usado para mejorar o sino cambiar la resolución de una imagen digital, incluyendo, por ejemplo, un procesador de imagen usado en un aparato de televisión o similar, impresoras, cámaras digitales, tarjetas de captura de señal de televisión, software de procesamiento digital de imagen que puede ser usado en muchas aplicaciones, etc. Los métodos descritos en la presente memoria pueden ser realizados por un software apropiado que es ejecutado en un equipo informático apropiado. El software puede estar incluido en un circuito integrado, estando adaptado el circuito integrado para realizar, o para su uso en la realización de los procedimientos pertinentes. Muchas de las etapas de procesamiento pueden ser realizadas usando un software, un hardware dedicado (tal como ASIC), o una combinación.

Ahora, se describirán, a modo de ejemplo, las realizaciones de la presente invención, con referencia a los dibujos adjuntos, en los que:

- La Fig. 1 muestra, esquemáticamente, una técnica anterior para escalar una imagen óptica usando clasificación de píxeles;
- La Fig. 2 muestra, esquemáticamente, cómo se obtienen, en la técnica anterior, los parámetros de distribución y los coeficientes del filtro de interpolación para el método de la Figura 1;
- La Fig. 3 muestra, esquemáticamente, un ejemplo de escalado de imagen según una realización de la presente invención;
- La Fig. 4 muestra las máscaras espaciales que corresponden a cada elemento del vector de características para un ejemplo específico de un método para la obtención de un vector de características para un píxel;
- La Fig. 5 muestra un conjunto alternativo de máscaras para un ejemplo específico de un método para la obtención de un vector de características para un píxel;
- La Fig. 6 muestra, esquemáticamente, un ejemplo de un escalado no entero según una realización de la presente invención;
- La Fig. 7 muestra, esquemáticamente, la fase de entrenamiento, en la que se obtienen los filtros de interpolación y las clases de contexto; y,
- La Fig. 8 muestra, esquemáticamente, un ejemplo de un método de entrenamiento preferente.

En resumen, el presente método preferente y el presente aparato preferente usan una fase de entrenamiento inteligente, fuera de línea, antes de la fase de interpolación real. La fase de entrenamiento, fuera de línea, preferente, es computacionalmente exigente, pero solo es necesario realizarla una sola vez. La fase de interpolación preferente es bastante sencilla computacionalmente, ya que puede usar un número relativamente pequeño de clases de contexto y debido a que puede requerir un filtrado con un único filtro de interpolación. Una vez aprendidos los coeficientes del filtro y los prototipos de agrupación de todas las clases de contexto a partir de los datos de entrenamiento, estos parámetros son pasados a un módulo de interpolación. En la fase de interpolación, el contexto de un píxel de entrada determinado es extraído por un extractor de características. El vector de características extraído es pasado a un clasificador de contextos. El clasificador de contextos decide a qué clase de contexto pertenece el píxel real de baja resolución y pasa el índice de clase correspondiente a un filtro lineal. A continuación, un bloque de filtrado de interpolación lineal aplica el conjunto correspondiente de coeficientes de filtro para obtener los píxeles de alta resolución, interpolados.

En mayor detalle, y haciendo referencia, inicialmente, a la Figura 3, un ejemplo de un procedimiento de interpolación según una realización preferente de la presente invención es básicamente similar al procedimiento de interpolación de la técnica anterior, divulgada por Atkins et al., y mostrada, esquemáticamente, en la Figura 1. En particular, para cada píxel 101 de entrada en una imagen 102 de entrada de baja resolución, una ventana 103 de 5x5 es vectorizada en columnas en un vector Z de columna 25x1, que puede denominarse un vector Z de observación. En un bloque 200 extractor de características, el vector Z de observación es asignado a un vector Y de características usando un operador f de proyección de transformación no lineal, es decir, $Y = f(Z)$ para cada píxel 101 de entrada. En términos generales, Y es de menor dimensión que Z. Esta asignación desde el vector Z de observación al vector Y de características afecta a la calidad de la imagen final de alta resolución, interpolada. Son posibles un número de operadores f de proyección diferentes.

En el presente contexto, el operador de proyección preferente es tal que puede usar un número relativamente pequeño de clases de contexto sin sacrificar (al menos significativamente) la calidad de la imagen de salida. En una realización preferente, con el fin de obtener los elementos del vector Y de características para el píxel 101 que está siendo procesado, esto se consigue mediante el operador de proyección usando, principal o exclusivamente, las diferencias de primer orden entre los píxeles, es decir, las diferencias en los valores de los componentes de la señal de pares contiguos de píxeles en la ventana 103 de píxeles alrededor del píxel 101 que está siendo procesado (además de o en lugar de las diferencias que usan el píxel 101 que está siendo procesado). (Los componentes de señal posibles que pueden ser usados para este propósito se exponen más adelante). Usando las diferencias entre los píxeles, en lugar de o además de las diferencias entre el píxel que está siendo procesado y uno o más de sus píxeles circundantes, ayuda a capturar el borde local dominante o estructura similar bajo un ligero solapamiento y otros ruidos de señal. A esto contribuye el uso de una

5 ventana 103 relativamente grande de píxeles alrededor del píxel 101 que está siendo procesado, tal como una ventana 103 de 5x5. Mediante el uso de diferencias de primer orden, se evitan las dificultades con el procesamiento de las señales de entrada con ruido, ya que las diferencias de mayor orden (es decir, diferencias entre píxeles que no son contiguos unos a los otros, sino que, por el contrario, están separados unos de otros por otros píxeles) son más sensibles al ruido de la señal.

10 Además, es preferente que las orientaciones de borde similares, con polaridades opuestas, sean forzadas a la misma agrupación o clase de contexto, ya que el procedimiento de interpolación preferente no se ve afectado por estas polaridades diferentes. Por ejemplo, un borde horizontal perfecto, con una transición de alto a bajo, y un borde horizontal perfecto, con una transición de bajo a alto, son forzados, idealmente, a la misma agrupación. Esto puede conseguirse en un ejemplo, elevando las diferencias de primer orden a una potencia par. Esto suprime también las diferencias de ruido, ya que las diferencias dominantes elevadas a la segunda potencia o a una potencia superior reduce, de manera efectiva, la influencia de los términos de diferencias más pequeños, que podrían haber sido causados por valores de ruido de píxel o textura insignificantes. La normalización de los elementos del vector Y de características, tal como se expone más adelante, ayuda también a suprimir los efectos de las diferencias de ruido.

15 En un ejemplo específico, los elementos Y_i del vector Y de características pueden ser obtenidos tal como se indica a continuación:

20
$$Y_1 = (Z_7 - Z_{13})^2$$

$$Y_2 = (Z_8 - Z_{13})^2$$

$$Y_3 = (Z_9 - Z_{13})^2$$

25
$$Y_4 = (Z_{12} - Z_{13})^2$$

$$Y_5 = (Z_{14} - Z_{13})^2$$

$$Y_6 = (Z_{17} - Z_{13})^2$$

30
$$Y_7 = (Z_{18} - Z_{13})^2$$

$$Y_8 = (Z_{19} - Z_{13})^2$$

35
$$Y_9 = (Z_1 - Z_7)^2$$

$$Y_{10} = (Z_2 - Z_6)^2$$

$$Y_{11} = (Z_{17} - Z_{21})^2$$

40
$$Y_{12} = (Z_{16} - Z_{22})^2$$

$$Y_{13} = (Z_{25} - Z_{19})^2$$

45
$$Y_{14} = (Z_{24} - Z_{20})^2$$

$$Y_{15} = (Z_5 - Z_9)^2$$

$$Y_{16} = (Z_{10} - Z_4)^2$$

50 Preferentemente, estos elementos Y_1, \dots, Y_{16} están normalizados como:

55
$$Y_i = \frac{Y_i}{\left[\sum_{j=1}^{16} (Y_j)^2 \right]^k}$$

Aquí, k es, más preferentemente, $0,75^{3/2}$ ó 0,375. Sin embargo, pueden usarse valores superiores o inferiores para k para controlar la nitidez del procedimiento de interpolación y dependiendo de la fase de entrenamiento y el tipo de imágenes de entrada, donde $0 < k < 1$.

Tal como se verá, en este ejemplo, todos los elementos Y_i del vector Y de características son diferencias de primer orden, es decir, diferencias entre píxeles contiguos, en la ventana 103 5x5 de píxeles alrededor del píxel 101 que se está procesando. En particular, además de tomar las diferencias entre el píxel 101 que está siendo procesado y los píxeles contiguos a ese píxel 101, las diferencias entre otros pares de píxeles en la ventana 103 son usadas también (en este caso, para obtener los elementos Y_9 a Y_{16}). Se ha encontrado que tomando sólo la diferencia de primer orden entre el píxel 101 y sus vecinos inmediatos como vectores de características puede no representar con precisión la estructura 2D de la imagen que rodea el píxel 101. Por lo tanto, la ocurrencia de picos, depresiones o ensilladuras de intensidad alrededor del píxel 101 sólo puede ser detectada apropiadamente si las diferencias entre otros pares de píxeles en la ventana 103, tales como diferencias que son ortogonales a las diferencias centro-circundantes estándares, son añadidas al conjunto de características. Además, las diferencias entre píxeles contiguos en las esquinas de la ventana 103 son usadas también en este ejemplo con el fin de obtener una mejor representación de la estructura de la imagen. Todas las diferencias son elevadas al cuadrado, de manera que las orientaciones de borde similares, con polaridades opuestas, son forzadas a la misma clase de contexto. A continuación, los elementos obtenidos inicialmente son normalizados para encontrar los valores finales para los elementos del vector Y de características para el píxel 101 que está siendo procesado.

En un segundo ejemplo, extendido, los elementos Y_i del vector Y de características son obtenidos tal como se indica a continuación:

20

$$Y_1 = (Z_7 - Z_{13})^2$$

$$Y_2 = (Z_8 - Z_{13})^2$$

25

$$Y_3 = (Z_9 - Z_{13})^2$$

$$Y_4 = (Z_{12} - Z_{13})^2$$

$$Y_5 = (Z_{14} - Z_{13})^2$$

30

$$Y_6 = (Z_{17} - Z_{13})^2$$

$$Y_7 = (Z_{18} - Z_{13})^2$$

$$Y_8 = (Z_{19} - Z_{13})^2$$

35

$$Y_9 = (Z_1 - Z_7)^2$$

$$Y_{10} = (Z_2 - Z_6)^2$$

40

$$Y_{11} = (Z_{17} - Z_{21})^2$$

$$Y_{12} = (Z_{16} - Z_{22})^2$$

$$Y_{13} = (Z_{25} - Z_{19})^2$$

45

$$Y_{14} = (Z_{24} - Z_{20})^2$$

$$Y_{15} = (Z_5 - Z_9)^2$$

50

$$Y_{16} = (Z_{10} - Z_4)^2$$

$$Y_{17} = (Z_8 - Z_{12})^2$$

$$Y_{18} = (Z_{12} - Z_{18})^2$$

55

$$Y_{19} = (Z_{18} - Z_{14})^2$$

$$Y_{20} = (Z_{14} - Z_8)^2$$

60

$$\begin{aligned}
 Y_{21} &= (Z_1 + 2 \cdot Z_6 + 4 \cdot Z_{11} + 2 \cdot Z_{16} + Z_{21} - Z_2 - 2 \cdot Z_7 - 4 \cdot Z_{12} - 2 \cdot Z_{17} - Z_{22})^2 \\
 Y_{22} &= (Z_4 + 2 \cdot Z_9 + 4 \cdot Z_{14} + 2 \cdot Z_{19} + Z_{24} - Z_5 - 2 \cdot Z_{10} - 4 \cdot Z_{15} - 2 \cdot Z_{20} - Z_{25})^2 \\
 Y_{23} &= (Z_2 + 2 \cdot Z_7 + 4 \cdot Z_{12} + 2 \cdot Z_{17} + Z_{22} - Z_3 - 2 \cdot Z_8 - 4 \cdot Z_{13} - 2 \cdot Z_{18} - Z_{23})^2 \\
 Y_{24} &= (Z_3 + 2 \cdot Z_8 + 4 \cdot Z_{13} + 2 \cdot Z_{18} + Z_{23} - Z_4 - 2 \cdot Z_9 - 4 \cdot Z_{14} - 2 \cdot Z_{19} - Z_{24})^2
 \end{aligned}$$

La Figura 4 muestra las máscaras espaciales que corresponden a cada elemento del vector de características para este ejemplo. Los elementos del vector de características, Y_i , para $i = 1, \dots, 24$, son normalizados como:

$$Y_i = \frac{Y_i}{\left[\sum_{j=1}^{24} (Y_j)^2 \right]^k}$$

en la que k puede ser, por ejemplo, 0,75 ó 0,375.

En este ejemplo, las entradas adicionales del vector de características se incluyen también en el vector de características final para conseguir robustez contra ciertos artefactos de compresión o de des-entrelazado no deseados. Por ejemplo, las entradas del vector de características que miden la semejanza de las líneas pares (y/o impares), y/o la diferencia entre las líneas pares e impares, dentro de la ventana 103, de la que se extraen los vectores de características, pueden ser usadas para identificar píxeles ubicados en artefactos de des-entrelazado. Es ventajoso detectar dichos píxeles y evitar el uso de filtros de mejora de resolución que simplemente mejorarían el contenido de frecuencia de una región artefacto y la harían aún más visible. Teniendo en cuenta que la tendencia actual es la transmisión de imágenes de televisión en formato progresivo y que dichas imágenes se obtienen frecuentemente a partir de imágenes de televisión entrelazadas originales, es un hecho que las imágenes transmitidas en formato progresivo contienen, frecuentemente, artefactos de des-entrelazado no deseados.

De esta manera, en el segundo ejemplo específico proporcionado anteriormente, los últimos cuatro elementos Y_{21} a Y_{24} están diseñados para ser sensibles a los artefactos de des-entrelazado, que se caracterizan por líneas alternas. En particular, las diferencias entre pares de píxeles adyacentes en filas adyacentes en la imagen 102 de entrada son usadas de manera que el vector Y de características puede acomodar mejor los artefactos de des-entrelazado. Discriminando, de manera efectiva, artefactos de des-entrelazado, puede evitarse la amplificación de los artefactos de des-entrelazado y, de hecho, puede incluso suprimirse los artefactos de des-entrelazado. A esto contribuye también el uso de muestras de entrenamiento apropiadas en los datos de entrenamiento, tal como se expone adicionalmente, más adelante. Por ejemplo, si la fase de entrenamiento es suministrada con marcos de baja resolución con artefactos de des-entrelazado (obtenidos, por ejemplo, introduciendo borrosidad y convirtiendo a una menor resolución un vídeo entrelazado de alta resolución el cual es obtenido, a su vez, a partir de un contenido progresivo original, de alta resolución) y los marcos progresivos de alta resolución correspondientes, el método puede ser entrenado, en una manera mejor, para suprimir artefactos de des-entrelazado. (Se entenderá que los coeficientes de 1, 2 y 4, usados en los últimos cuatro elementos de este ejemplo, se pueden variar).

Cabe señalar que se pueden usar diferentes máscaras de diferencias para obtener las diferencias de primer orden. Por ejemplo, un conjunto alternativo de máscaras para obtener diferencias promedio horizontal, vertical y diagonal, robustas frente al ruido, se muestra en la Figura 5. Las entradas correspondientes del vector de características se proporcionan como:

$$\begin{aligned}
 Y &= (0.25 \cdot Z_7 + 0.5 \cdot Z_{12} + 0.25 \cdot Z_{17} - Z_{13})^2 \\
 Y &= (0.25 \cdot Z_{17} + 0.5 \cdot Z_{18} + 0.25 \cdot Z_{19} - Z_{13})^2 \\
 Y &= (0.25 \cdot Z_{12} + 0.5 \cdot Z_{17} + 0.25 \cdot Z_{18} - Z_{13})^2
 \end{aligned}$$

Otras diferencias direccionales pueden ser modificadas en una manera similar. De nuevo, se entenderá que el esquema de ponderación (aquí, 0,25, 0,5, 0,25) puede ser diseñado también para que se ajuste a ciertas características de la señal de entrada.

5 Se entenderá que la elevación a la segunda potencia no es la única manera de conseguir el objetivo de forzar a las orientaciones de borde similares, con polaridades opuestas, a la misma agrupación o clase de contexto. Por ejemplo, es posible usar valores absolutos combinados con una asignación de tabla de consulta, diseñada para imitar una regla de elevación a una potencia para conseguir el mismo objetivo (o cualquier otra asignación no lineal que conseguiría efectivamente el mismo objetivo). La tabla de consulta puede ser ampliada para que incluya también la etapa de normalización. De hecho, una tabla de consulta puede ser usada para evitar tener que calcular los diversos valores de los elementos del vector Y de características sobre la marcha. Una operación de umbral puede ser realizada inmediatamente después de obtener las diferencias, como una contra-medida adicional contra el ruido.

15 Además, los vectores de características, obtenidas como diferencias de primer orden, pueden ser procesados en maneras preferentes antes de la etapa de normalización indicada anteriormente. Una posible etapa de procesamiento que puede ser incluida en el camino de extracción de características es la eliminación de media. Para realizar una eliminación de media, se calcula el valor medio de los elementos del vector de características, sumando todos los elementos del vector de características (por ejemplo, Y_1 a Y_{24} en el segundo ejemplo indicado anteriormente) y dividiendo por la longitud del vector de características. El valor medio resultante es restado de todos los elementos del vector de características, de uno en uno, para obtener el vector de características con media eliminada.

20 En el caso de señales de entrada de color, que están en formato RGB, la etapa de extracción de características puede ser realizada sobre cada uno de los componentes R, G y B, por separado. En otras palabras, cuando se ha indicado anteriormente, por ejemplo, "la diferencia entre los píxeles", en este caso, esto se entenderá como "la diferencia en los valores respectivos de los componentes R, G, y B de los píxeles respectivos".

25 En el caso de señales de entrada de color que son transformadas desde RGB a YUV (o YCbCr), la etapa de extracción de características puede ser realizada solo en las componentes de luminancia. En otras palabras, cuando se ha indicado anteriormente, por ejemplo "la diferencia entre píxeles", en este caso, esto se entenderá que significa "la diferencia en los valores de luminancia de los píxeles respectivos". Usando sólo los componentes de luminancia, se reduce la complejidad computacional, pero sin prácticamente ninguna pérdida de calidad visual en la imagen de salida final. Hay una peculiaridad que es abordada, preferentemente, de manera explícita, en tal caso, concretamente, los bordes con isoluminancia, es decir, los bordes de color para los cuales la componente de luminancia no tiene un borde. Dichos bordes no ocurren con frecuencia, y la presencia de dicho borde no implica necesariamente un artefacto que suponga una perturbación visual. La única desventaja de ignorar simplemente dichos bordes es una menor definición de los bordes, lo que resulta en una transición borrosa. Una forma de resolver este problema consiste en deducir, de antemano, los componentes de color RGB para los cuales los bordes con isoluminancia pueden ocurrir para la transformación del espacio de color específica, usada en la canalización de vídeo. Una vez identificados los valores RGB específicos que conducen a los bordes con isoluminancia, para estos valores RGB, la etapa de extracción de características puede ser realizada en uno o en ambos componentes de crominancia (es decir, aquí, los componentes UV o CbCr).

30 Haciendo referencia de nuevo a la Figura 3, habiendo obtenido el vector Y de características para el píxel 101 que está siendo procesado, o habiendo obtenido los vectores Y de características respectivos para todos los píxeles en la imagen 102 de entrada original, el vector Y de características o cada vector Y de características es pasado a un clasificador 201 de contextos. Aquí, el vector Y de características de cada píxel en la imagen 102 de entrada es usado para clasificar cada píxel de entrada de baja resolución como un miembro de un número limitado de clases de contexto. Estas clases modelan diferentes regiones de imagen, tales como bordes, regiones lisas y texturadas. Hay un conjunto diferente de coeficientes de filtro de interpolación para cada contexto, siendo ajustados los coeficientes a las características espaciales de ese contexto.

35 En este ejemplo, la clasificación de contextos es realizada comparando el vector de características del píxel 101 actual con los prototipos de las clases de contexto. En un ejemplo, esta comparación se basa en una distancia euclídea ponderada. La clase cuyo prototipo proporciona la distancia más corta al vector de características actual es declarada como la clase de contexto del píxel 101 actual. Se observará que esta clasificación contexto es dura, es decir, cada vector de características es clasificado a una única clase de contexto.

40 La salida del clasificador 201 de contextos es pasada a un filtro 202 de interpolación lineal. Debido a que la presente clasificación de contextos es dura, sólo se usa un único filtro de interpolación para cada píxel 101 de la imagen 102 de entrada. El bloque de filtro de interpolación lineal usa una ventana 5x5, centrada en el píxel 101 de entrada, de baja resolución, para interpolar un único píxel x 104 de la imagen 105 de salida, de alta resolución, como

$$\mathbf{x} = \mathbf{AZ}$$

5 en la que Z es tal como se ha definido anteriormente y A es el filtro de interpolación optimizado para la clase de contexto del píxel 101 actual.

La etapa de interpolación puede implicar también un término β de desviación adicional, de manera que:

10

$$\mathbf{x} = \mathbf{AZ} + \beta$$

En experimentos, se ha observado que la eliminación del término de desviación adicional resulta en una pérdida de calidad visual insignificante, mientras que proporciona sustanciales ahorros computacionales. Sin embargo, en una configuración en la que la calidad visual es de suma importancia, el término de desviación puede ser incluido para maximizar la calidad del resultado de interpolación.

Cabe señalar que, a diferencia de los esquemas de la técnica anterior, tales como los divulgados en los documentos de Atkins, indicados anteriormente, el presente método preferente interpola los píxeles de salida, uno cada vez. Esto posibilita proporciones de un escalado no entero con el presente método. Con referencia a la Figura 6, se muestra un ejemplo de un escalado no entero de una imagen de baja resolución (LR) a una imagen de alta resolución en la que, solamente a modo de ejemplo y sin limitación, se usa un factor de escala de 3,5. Considerando, en primer lugar, el píxel de la imagen de alta resolución de coordenada de imagen (150, 100) en la imagen de alta resolución. Este píxel de alta resolución es asignado a la coordenada cartesiana (42, 85, 28, 57) en la imagen de baja resolución. El píxel actual, de baja resolución, correspondiente, está ubicado en la coordenada de imagen (42, 28) en la imagen de baja resolución. De esta manera, con el fin de conseguir una asignación de tipo uno-a-varios entre imágenes de baja y alta resolución, el píxel de baja resolución es dividido (teóricamente) en una rejilla. Esta división puede ser realizada de diversas maneras. A modo de ilustración, una rejilla de filtro 3x3, de nueve particiones de rejilla, se muestra en la Figura 6. A cada partición, se le asigna un filtro de F1 a F9 respectivo, siendo obtenidos los filtros F1 a F9 en una fase de entrenamiento (explicada más adelante). Para cada píxel en la imagen de alta resolución, se identifica su proyección o ubicación correspondiente en la imagen de baja resolución, y el filtro F1-F9 correspondiente de la partición correspondiente en la imagen de baja resolución es usado y aplicado al píxel de baja resolución correspondiente, para construir el píxel de alta resolución en la imagen de alta resolución. Por supuesto, puede emplearse la misma estrategia para proporciones de escalado enteras, para aumentar la resolución. Se apreciará que una rejilla de filtro 3x3 es usada en este ejemplo, ya que la relación de escala deseada de 3,5, es cercana a 3. Pueden usarse diferentes rejillas de filtro, con un número correspondiente de filtros F_n diferentes, para diferentes factores de escala, en una manera correspondiente.

En resumen, en la fase de interpolación de la realización preferente, se hace uso de las diferencias de primer orden entre los píxeles en la imagen 102 de entrada, es decir, se hace uso de las diferencias en los valores de los componentes de la señal de pares contiguos de píxeles en la ventana 103 de píxeles alrededor del píxel 101 que está siendo procesado (además de o en lugar de las diferencias que usan el píxel 101 que está siendo procesado). Usando las diferencias entre los píxeles distintos de o además de las diferencias entre el píxel que está siendo procesado y uno o más de sus píxeles circundantes, ayuda a capturar el borde local dominante o estructura similar bajo un ligero solapamiento y ruido de señal. Esto permite obtener una alta calidad de imagen usando sólo un número relativamente pequeño de clases de contexto. Esto es importante, debido a que esta fase de interpolación se realiza en el lado del receptor, por ejemplo, en un esquema de transmisión de señal de televisión, o se realiza, típicamente, en otros equipos de consumo, tales como una impresora o una cámara digital, etc., ya que permite obtener una alta calidad de imagen a un costo relativamente bajo. Además, la realización preferente permite conseguir factores de escala no enteros, desde una resolución a otra, en una manera computacionalmente eficiente y sin ninguna pérdida significativa de calidad de imagen.

La fase de entrenamiento, en la que se obtienen los filtros de interpolación y las clases de contexto, se muestra, esquemáticamente, en la Figura 7. Al igual que en la técnica anterior, un número de imágenes de origen, de alta resolución, son adquiridas y las correspondientes imágenes de baja resolución son obtenidas a partir de las imágenes de alta resolución, disminuyendo la resolución y/o reduciendo la escala, etc., de manera apropiada, para obtener imágenes de entrenamiento. A continuación, los vectores Y de característica son obtenidos a partir de las imágenes de entrenamiento, preferentemente, usando el mismo operador de proyección para obtener los elementos Y_i del vector Y de características, tal como se usará en la fase de interpolación (véase más arriba). A continuación, en el bloque 300, se realiza la agrupación con respecto a los vectores Y de características, es decir, se obtiene la clase o las clases de contexto más cercanas al vector de características actual.

En el método de entrenamiento preferente, para poder iniciar las iteraciones de entrenamiento, se requiere un conjunto de

valores iniciales para los coeficientes del filtro de interpolación. Hay muchas estrategias diferentes para la inicialización de los coeficientes del filtro de interpolación. Es un hecho bien conocido que para los algoritmos iterativos, tales como el método de entrenamiento preferente, las condiciones iniciales afectan a la velocidad de convergencia y a los resultados finales. De esta manera, es importante elegir sabiamente los coeficientes iniciales del filtro. En una realización preferente, se usan un único filtro, no direccional, de tipo Gaussiano, y varios filtros direccionales, de tipo Gaussiano, para inicializar los coeficientes del filtro. Los filtros de tipo Gaussiano se obtienen del núcleo Gaussiano bidimensional estándar:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2\rho(x-m_x)(y-m_y)}{\sigma_x\sigma_y} - \frac{(y-m_y)^2}{\sigma_y^2}\right)\right]$$

y pueden ser ajustados para tener la orientación de borde deseada, ajustando el coeficiente de correlación ρ y las desviaciones estándar σ_x , σ_y . Otros métodos para elegir los coeficientes iniciales del filtro son posibles.

En la realización preferente, se usa una agrupación borrosa robusta con una función de coste modificada. Esto permite imponer fácilmente restricciones sobre los tamaños y las formas de las agrupaciones. También, pueden introducirse restricciones adicionales para manipular vectores de características con ruido y valores atípicos.

Una función de coste preferente es:

$$J = \sum_{i=1}^M \sum_{k=1}^N (u_{ik})^m d^2(y_k, p_i) + \sum_{i=1}^M \eta_i \sum_{k=1}^N (1-u_{ik})^m$$

en la que

M es el número de clases de contexto,

N es el número de píxeles de entrenamiento (y, por tanto, el número de vectores de características),

u_{ik} es el grado de pertenencia de k-ésimo vector de características a la i-ésima clase de contexto,

$$d^2(y_k, p_i) = \rho_i^{1/r} |C_i|^{1/r} (y_k - m_i)^T C_i^{-1} (y_k - m_i)$$

es la distancia métrica,

ρ_i es el parámetro que controla el tamaño de la i-ésima clase,

r es la dimensión de los vectores de características,

m_i es el vector de características promedio de la i-ésima clase,

C_i es la matriz de covarianza de la i-ésima clase,

η_i es un parámetro que controla la asociación de los valores atípicos y las muestras de ruido a la i-ésima clase, y

m es el parámetro de borrosidad que controla la borrosidad de la agrupación.

La agrupación borrosa se consigue minimizando la función de coste J sobre todas las particiones borrosas. Las relaciones de pertenencia resultantes para todos los vectores de características son usadas para la derivación del filtro. Se comprenderá que la clasificación durante el filtrado puede ser borrosa o dura. Se entenderá también que los métodos de agrupación borrosa incluyen métodos de agrupación dura, tales como medias k como un caso especial (en el caso límite en el que $m \rightarrow 1$). En ciertas configuraciones, puede ser ventajoso entrenar el método propuesto con un método de agrupación dura de medias k.

Otras funciones de coste, diferentes a la mostrada anteriormente, pueden ser adaptadas para diferentes características de las imágenes de entrada. Por ejemplo, pueden imponerse diferentes formas y tamaños de grupo, diseñando

apropiadamente la distancia métrica $d^2 (y_k, p_i)$.

$$\sum_{i=1}^M \eta_i \sum_{k=1}^N (1-u_{ik})^m$$

Pueden incluirse términos adicionales, tales como η_i para aumentar la robustez frente a ruido y valores atípicos. En este término específico, η_i representa una distancia fija de clase específica. Mientras que el primer componente de la función de coste J, tal como se ha indicado anteriormente, está diseñado para encontrar relaciones de pertenencia que minimicen la suma ponderada de las distancias de todos los vectores de características a los prototipos de clases (donde los coeficientes de ponderación son la relación de pertenencia a las clases), el término adicional está diseñado para que los vectores de características tengan relaciones de pertenencia pequeñas a todas las clases, es decir, los vectores de características que están muy lejos de todos los prototipos de clases. Debido a que es probable que estos vectores de características sean valores atípicos o muestras muy ruidosas, son "asignados" efectivamente a una clase de ruido, y se suprime su influencia negativa sobre los prototipos de clases actualizados.

El bloque 301 representa la fase en la que se obtienen los filtros de interpolación. Para cada agrupación obtenida en el bloque 300 de agrupación, se calcula el filtro de interpolación con el mínimo error cuadrático medio (LSE). Mientras se resuelve para el filtro LSE, las relaciones de pertenencia de los píxeles son usadas como coeficientes de ponderación.

De esta manera, se obtienen las clases de contexto, y sus prototipos, y los coeficientes del filtro de interpolación, para su uso en la fase de interpolación, expuesta anteriormente.

En resumen, por lo tanto, y con referencia al diagrama de flujo esquemático de la Figura 8, en el método de entrenamiento preferente, se realiza, en primer lugar, una etapa 400 de entrenamiento fuera de línea. Todos los píxeles vecinos, de baja-alta resolución, extraídos de las imágenes de entrenamiento disponibles son introducidos al entrenamiento. En la etapa 401, se realiza una agrupación con respecto a los vectores de características. La agrupación se basa en robustas técnicas de agrupación borrosa. A continuación, en la etapa 402, se calculan los filtros de interpolación. Para cada agrupación obtenida en la etapa 401, se calcula el filtro de interpolación con el mínimo error cuadrático medio (LSE). Mientras se resuelve para el filtro LSE, las relaciones de pertenencia de los píxeles son usadas como coeficientes de ponderación.

La etapa 500 es la etapa de interpolación. En la etapa 501, se realiza una comprobación de las condiciones para terminar la etapa de interpolación, para comprobar que todavía hay píxeles de baja resolución para ser procesados. Los filtros de interpolación son aplicados a todos los píxeles de baja resolución, en el orden de exploración de trama.

La etapa 502 es la primera etapa en la interpolación. Se extrae una ventana $K \times K$, de baja resolución, centrada en el píxel de baja resolución actual, a ser filtrado. (En los ejemplos específicos anteriores, $K = 5$). En la etapa 503, los píxeles $K \times K$ son reordenados en un vector Z de columna de $K^2 \times 1$. En la etapa 504, para cada píxel de entrada de baja resolución, se obtiene un vector Y de características a partir del vector Z de columna, en el que:

$$Y = f(Z)$$

La extracción del vector de características es una asignación desde Z a Y, que es, normalmente, de dimensionalidad más baja. El vector de características preferente afecta a la calidad de la imagen de alta resolución, interpolada. Por lo tanto, la selección del vector de características es una parte importante del método, tal como se ha descrito anteriormente, en más detalle.

En la etapa 505, el vector de características de cada píxel de entrada, de baja resolución, es usado para clasificar el píxel de entrada, de baja resolución, como un miembro de un número limitado de clases de contexto, ponderando las clases tal como se ha expuesto anteriormente. Una vez más, estas clases modelan diferentes regiones de imagen, tales como bordes suaves, regiones suaves, regiones texturadas, etc. Hay un conjunto diferente de coeficientes del filtro de interpolación para cada contexto, los cuales están ajustados a las características especiales de ese contexto, y que se obtienen en el método de entrenamiento, fuera de línea, expuesto anteriormente.

A continuación, en la etapa 506, se realiza la interpolación lineal, tal como se ha indicado anteriormente. En este ejemplo, una ventana 5×5 está centrada en el píxel de entrada, de baja resolución, para interpolar un píxel 104, de alta resolución, denominado x, como:

$$x = AZ$$

en la que A es el filtro de interpolación.

A continuación, el flujo vuelve a la etapa 501 de comprobación de la condición. Finalmente, cuando ya no hay píxeles de baja resolución, se obtiene la imagen 105 de salida, de alta resolución.

5 Las realizaciones de la presente invención han sido descritas con referencia particular a los ejemplos ilustrados. Sin embargo, se apreciará que pueden realizarse variaciones y modificaciones a los ejemplos descritos, incluidas en el alcance de la presente invención, definido por las reivindicaciones adjuntas.

REIVINDICACIONES

- 5 1. Un método para cambiar la resolución espacial de una primera imagen (102) digital que tiene una primera resolución espacial para proporcionar una segunda imagen (105) digital que tiene una segunda resolución espacial, que es diferente de la primera resolución espacial, comprendiendo el método, para cada uno de entre una pluralidad de píxeles en la primera imagen digital:
- 10 obtener un vector (Y) de características para un píxel (101) en la primera imagen (102) digital;
 clasificar dicho píxel (101) de manera que dicho píxel es asignado a una o más clases de contexto, comparando el vector (Y) de características de dicho píxel con un prototipo de cada una de dichas clases de contexto, teniendo cada una de dichas clases de contexto uno o más filtros (202) de interpolación asociados a la misma; y
 15 filtrar dicho píxel (101) usando el uno o más filtros (202) de interpolación asociados con la una o más clases de contexto a las cuales está asignado dicho píxel, generando, de esta manera, uno o más píxeles (104) de la segunda imagen (105) digital;
 en el que dicho vector (Y) de características para dicho píxel (101) es una función de las diferencias entre pares contiguos seleccionados de píxeles en la primera imagen (102) digital, **caracterizado porque**
 20 al menos uno de dichos pares contiguos de píxeles en la primera imagen (102) digital no incluye dicho píxel (101).
- 25 2. Un método según la reivindicación 1, en el que el vector (Y) de características es una función sólo de las diferencias entre pares contiguos seleccionados de píxeles en la primera imagen (102) digital, en el que al menos uno de entre los pares contiguos de píxeles en la primera imagen (102) digital no incluye dicho píxel (101).
- 30 3. Un método según la reivindicación 1 o la reivindicación 2, en el que dichos pares de píxeles son píxeles en una ventana (103) de tamaño predeterminado alrededor de dicho píxel (101).
4. Un método según cualquiera de las reivindicaciones 1 a 3, en el que dicha función es tal que cada una de las diferencias entre los pares contiguos seleccionados de píxeles en la primera imagen (102) digital es elevada a una potencia n, donde n es un número par.
- 35 5. Un método según cualquiera de las reivindicaciones 1 a 4, en el que cada elemento de cada vector (Y) de características es una función de la diferencia entre un par contiguo seleccionado de píxeles en la primera imagen (102) digital, siendo obtenido cada vector de características normalizando los elementos de dicho vector de características, obtenidos de esta manera.
- 40 6. Un método según cualquiera de las reivindicaciones 1 a 5, en el que al menos uno de los elementos de cada vector (Y) de característica es una función de la diferencia entre pares de píxeles en filas contiguas de la primera imagen (102) digital.
- 45 7. Un método según cualquiera de las reivindicaciones 1 a 6, en el que la primera imagen (102) digital está en un formato que tiene al menos un componente de luminancia y dos componentes de crominancia, en el que:
 dichas diferencias entre al menos algunos de los pares contiguos de píxeles son las diferencias en los valores respectivos de los componentes de luminancia de dichos pares contiguos de píxeles.
- 50 8. Un método según la reivindicación 7, en el que la primera imagen (102) digital está inicialmente en formato RGB y es convertido en dicho formato, que tiene al menos un componente de luminancia y dos componentes de crominancia, en el que se identifican aquellas combinaciones de componentes RGB para las que pueden ocurrir bordes con isoluminancia cuando la imagen en formato RGB es convertida a dicho formato que tiene al menos un componente de luminancia y dos componentes de crominancia, en el que:
 55 para al menos algunos píxeles de la primera imagen (102) digital, que inicialmente tiene valores RGB correspondientes a aquellas combinaciones de los componentes RGB para las que pueden ocurrir bordes con isoluminancia cuando la imagen formato RGB es convertida a dicho formato, que tiene al menos un componente de luminancia y dos componentes de crominancia, dichas diferencias entre al menos algunos de los pares contiguos de dichos píxeles son las diferencias en los valores respectivos de al menos uno de las componentes de crominancia de dichos pares contiguos de píxeles.
- 60 9. Un método según cualquiera de las reivindicaciones 1 a 8, en el que:

la clasificación de dicho píxel (101) es tal que dicho píxel es asignado a una única clase de contexto, comparando el vector (Y) de características de dicho píxel con un prototipo de cada una de dichas clases de contexto; y

5 el filtrado de dicho píxel usa un único filtro (202) de interpolación que está asociado con dicha única clase de contexto a la que está asignado dicho píxel.

10. Un método según la reivindicación 9, en el que la segunda resolución es mayor que la primera resolución y la relación de la primera resolución a la segunda resolución no es un número entero, siendo obtenidos los píxeles (104) en la segunda imagen digital, para cada uno de dichos píxeles en la segunda imagen (105) digital:

10 identificando la ubicación en la primera imagen digital que corresponde a la ubicación de dicho píxel en la segunda imagen (104) digital;

identificando el píxel en la primera imagen (102) digital que corresponde a dicha ubicación en la primera imagen (102) digital; y

15 filtrando dicho píxel en la primera imagen digital con el único filtro (202) de interpolación respectivo que corresponde a dicha ubicación en la primera imagen (102) digital, obteniendo, de esta manera, los píxeles en la segunda imagen (104) digital.

11. Un método según cualquiera de las reivindicaciones 1 a 10, en el que los valores de los elementos de un vector (Y) de características para dicho píxel se obtienen desde una tabla de consulta.

12. Aparato para cambiar la resolución espacial de una primera imagen (102) digital, que tiene una primera resolución espacial, para proporcionar una segunda imagen (105) digital, que tiene una segunda resolución espacial, que es diferente de la primera resolución espacial, comprendiendo el aparato:

25 un extractor (200) de vectores de características para obtener un vector (Y) de características para un píxel (101) en la primera imagen (102) digital;

un clasificador (201) para clasificar dicho píxel (101) de manera que dicho píxel es asignado a una o más clases de contexto, comparando el vector (Y) de características de dicho píxel con un prototipo de cada una de dichas clases de contexto, teniendo cada una de dichas clases de contexto uno o más filtros de interpolación asociados a la misma; y,

30 un filtro (202) para filtrar dicho píxel usando el uno o más filtros (202) de interpolación asociados con la una o más clases de contexto a las que está asignado dicho píxel, para generar, de esta manera, uno o más píxeles (101) de la segunda imagen (105) digital;

35 en el que el extractor (200) de vectores de características está dispuesto de manera que dicho vector (Y) de características para dicho píxel (101) es una función de las diferencias entre pares contiguos seleccionados de píxeles en la primera imagen (102) digital, caracterizado porque al menos uno de dichos pares contiguos de píxeles en la primera imagen (102) digital no incluye dicho píxel (101).

40 13. Aparato según la reivindicación 12, en el que el extractor (200) de vectores de características está dispuesto de manera que el vector (Y) de características es una función sólo de las diferencias entre pares contiguos seleccionados de píxeles en la primera imagen (102) digital, en el que al menos uno de dichos pares de píxeles en la primera imagen digital no incluye dicho píxel (101).

45 14. Aparato según la reivindicación 12 o la reivindicación 13, en el que el extractor (200) de vectores de características está dispuesto de manera que dichos pares de píxeles son píxeles en una ventana (103) de tamaño predeterminado alrededor de dicho píxel.

50 15. Aparato según cualquiera de las reivindicaciones 12 a 14, en el que el extractor (200) de vectores de características está dispuesto de manera que dicha función es tal que cada una de las diferencias entre pares contiguos seleccionados de píxeles en la primera imagen (102) digital es elevada a una potencia n, donde n es un número par.

55 16. Aparato según cualquiera de las reivindicaciones 12 a 15, en el que el extractor (200) de vectores de características está dispuesto de manera que cada elemento de cada vector (Y) de características es una función de la diferencia entre un par contiguo seleccionado de píxeles en la primera imagen (102) digital y cada vector (Y) de características es obtenido normalizando los elementos de dicho vector de características, obtenido de esta manera.

60 17. Aparato según cualquiera de las reivindicaciones 12 a 16, en el que el extractor (200) de vectores de características está dispuesto de manera que al menos uno de los elementos de cada vector (Y) de características

es una función de la diferencia entre pares de píxeles en filas contiguas de la primera imagen (102) digital.

5 18. Aparato según cualquiera de las reivindicaciones 12 a 17, en el que el aparato está dispuesto para operar sobre una primera imagen (102) digital que está en un formato que tiene al menos un componente de luminancia y dos componentes de crominancia, estando dispuesto el extractor (200) de vectores de características de manera que dichas diferencias entre al menos algunos de los pares contiguos de píxeles son las diferencias en los valores respectivos de las componentes de luminancia de dichos pares contiguos de píxeles.

10 19. Aparato según la reivindicación 18, en el que el aparato está dispuesto para operar sobre una primera imagen (102) digital que está inicialmente en formato RGB y es convertida a dicho formato que tiene al menos un componente de luminancia y dos componentes de crominancia, en el que se han identificado aquellas combinaciones de componentes RGB para las que pueden ocurrir bordes con isoluminancia cuando la imagen de formato RGB es convertida a dicho formato que tiene al menos un componente de luminancia y dos componentes de crominancia, en el que:

15 el extractor (200) de vectores de características está dispuesto de manera que, para al menos algunos píxeles en la primera imagen (102) digital, que inicialmente tiene valores RGB correspondientes a aquellas combinaciones de componentes RGB para las que pueden ocurrir bordes de isoluminancia cuando la imagen de formato RGB es convertida a dicho formato que tiene al menos un componente de luminancia y dos componentes de crominancia, dichas diferencias entre al menos algunos de los pares contiguos de dichos píxeles son las diferencias en los valores respectivos de al menos uno de los componentes de crominancia de dichos pares contiguos de píxeles.

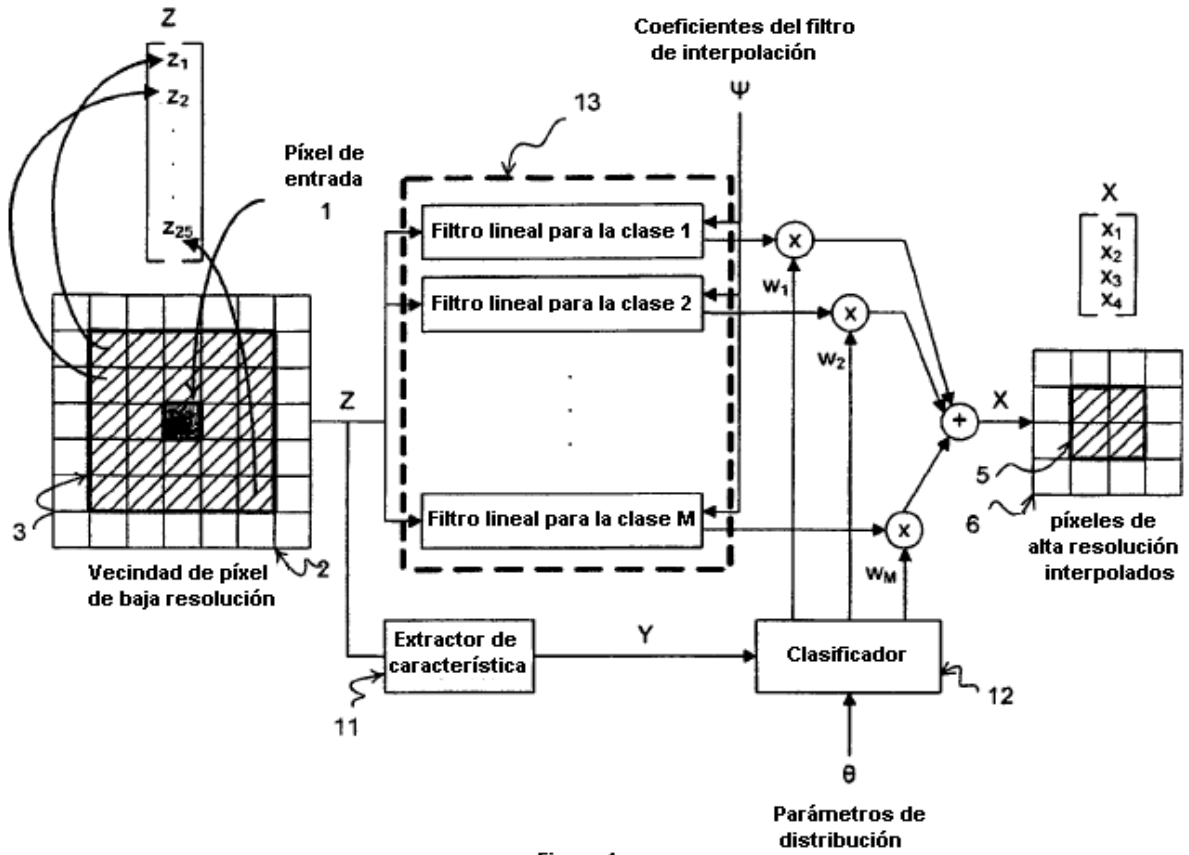
20 20. Aparato según cualquiera de las reivindicaciones 12 a 19, en el que:

25 el clasificador (201) está dispuesto de manera que dicho píxel (101) es asignado a una única clase de contexto comparando el vector (Y) de características de dicho píxel (101) con un prototipo de cada una de dichas clases de contexto; y
30 el filtro (202) está dispuesto para filtrar dicho píxel (101) usando un único filtro de interpolación que está asociado con dicha única clase de contexto a la que está asignado dicho píxel.

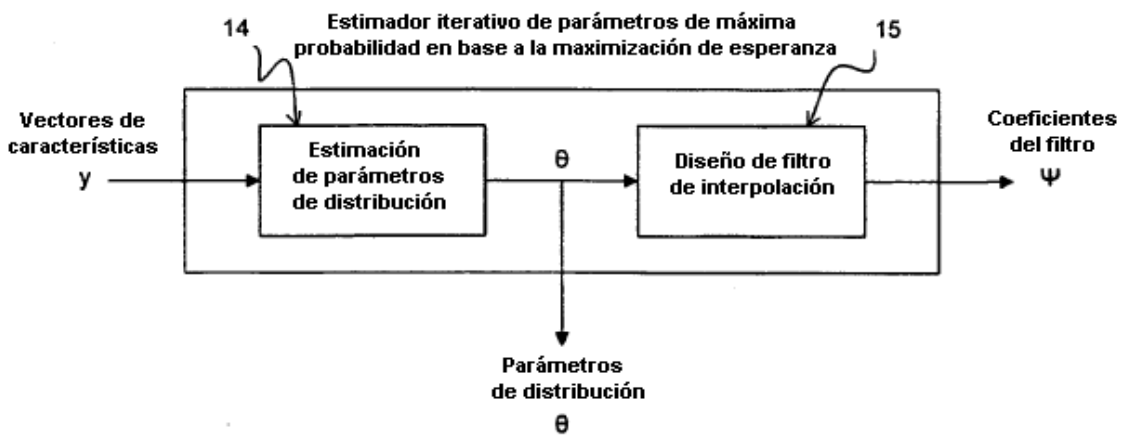
21. Aparato según la reivindicación 20, en el que la segunda resolución es mayor que la primera resolución y la relación de la primera resolución a la segunda resolución no es un entero, estando dispuesto el aparato para obtener los píxeles en la segunda imagen (105) digital, para cada uno de dichos píxeles en la segunda imagen digital:

35 identificando la ubicación en la primera imagen (102) digital que corresponde a la ubicación de dicho píxel en la segunda imagen digital;
identificando el píxel en la primera imagen (102) digital que corresponde a dicha ubicación en la primera imagen digital; y,
40 filtrando dicho píxel en la primera imagen (102) digital con el único filtro (202) de interpolación respectivo que corresponde a dicha ubicación en la primera imagen (102) digital, para obtener, de esta manera, los píxeles en la segunda imagen (104) digital.

45 22. Aparato según cualquiera de las reivindicaciones 12 a 21, que comprende una tabla de consulta que contiene los valores de los elementos de los vectores (Y) de características.



TÉCNICA ANTERIOR



TÉCNICA ANTERIOR

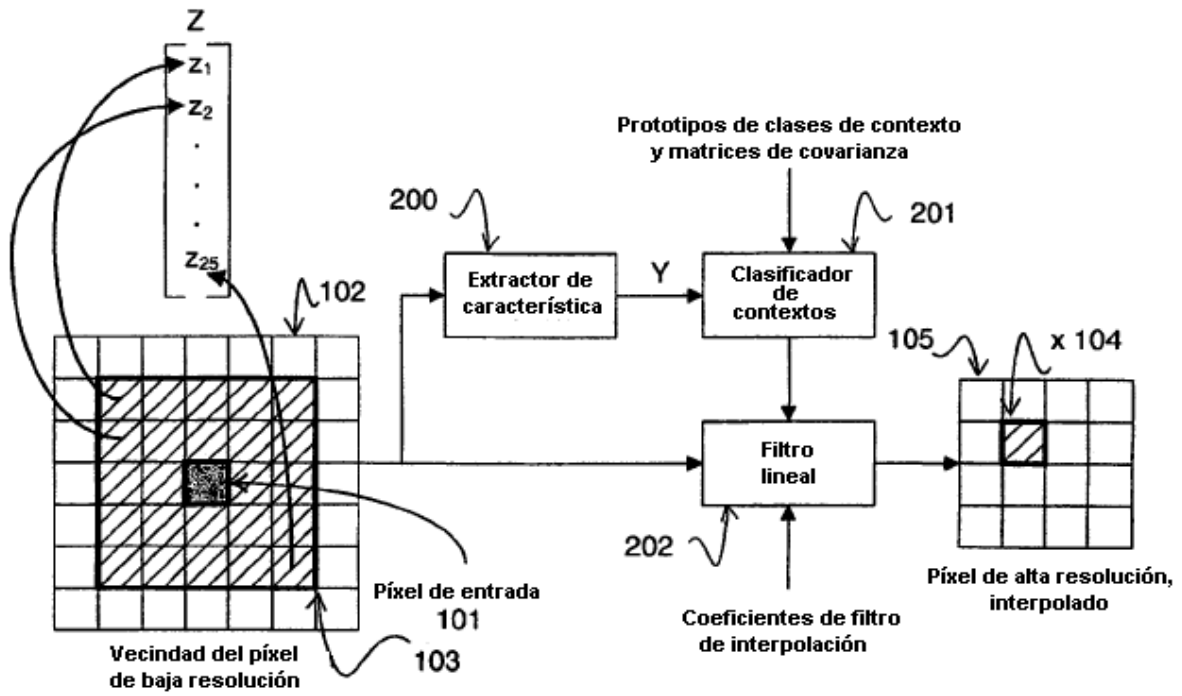


Figura 3

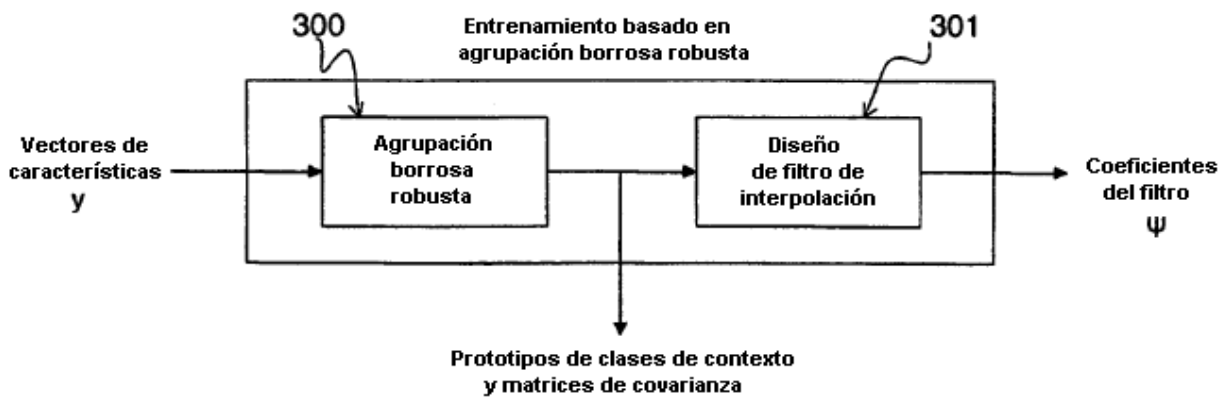


Figura 7

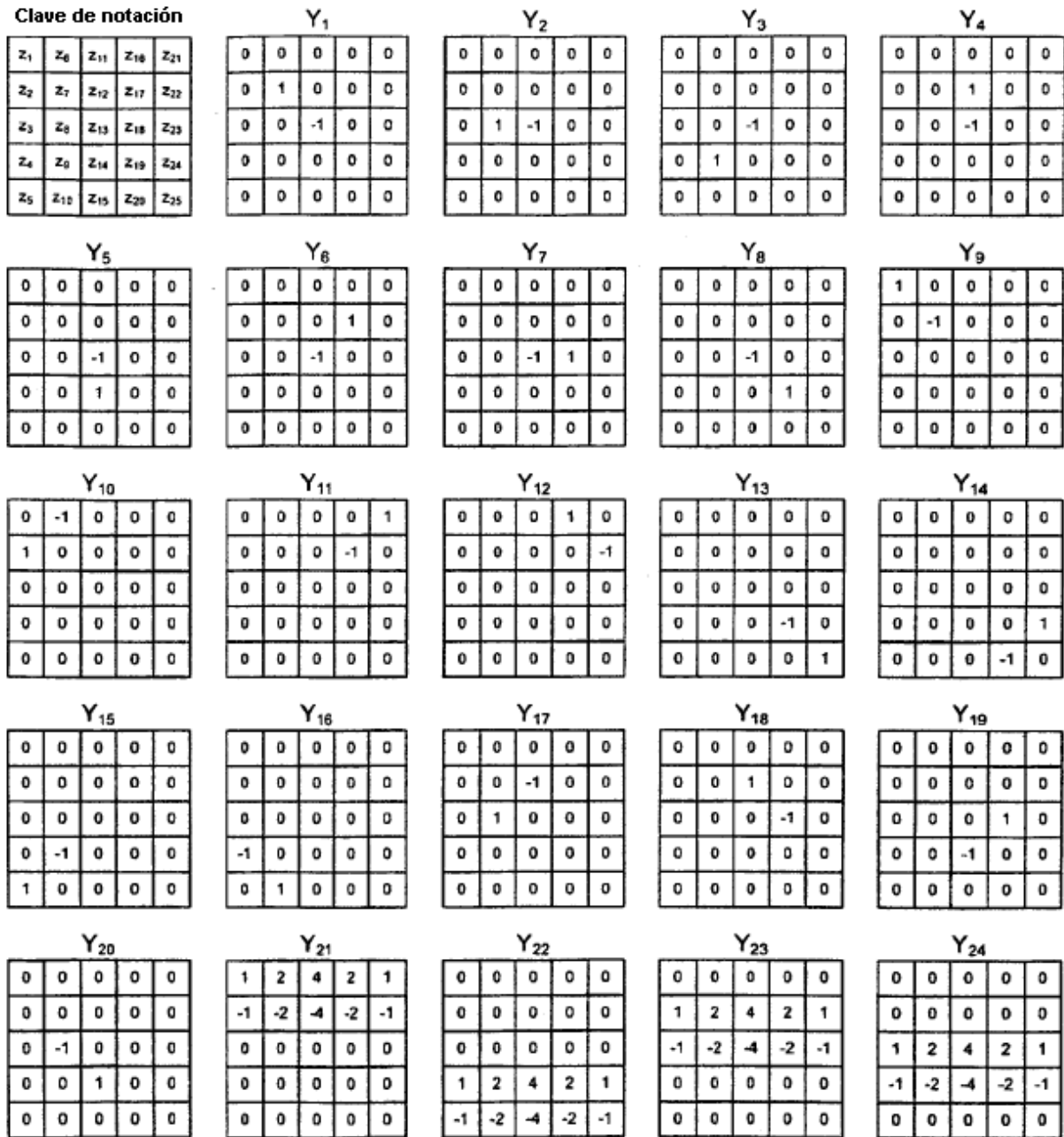


Figura 4

Arriba vertical	Derecha horizontal	Arriba-derecha diagonal
0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
0 1/4 1/2 1/4 0	0 0 0 1/4 0	0 0 1/4 1/2 0
0 0 -1 0 0	0 0 -1 1/2 0	0 0 -1 1/4 0
0 0 0 0 0	0 0 0 1/4 0	0 0 0 0 0
0 0 0 0 0	0 0 0 0 0	0 0 0 0 0

Figura 5

Ejemplo de escalado no entero
El factor de escala operativo es de 3,5

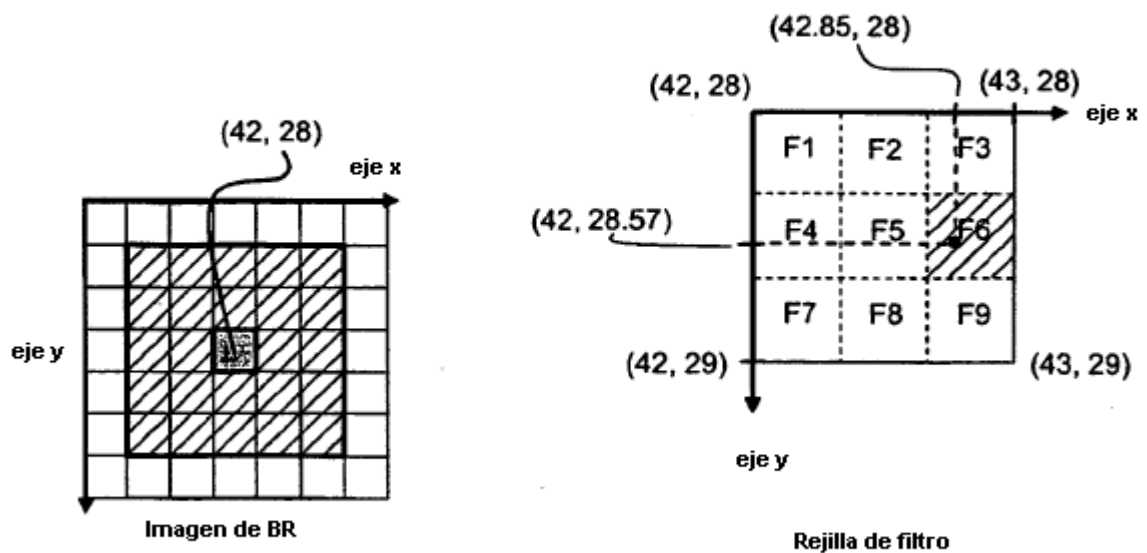
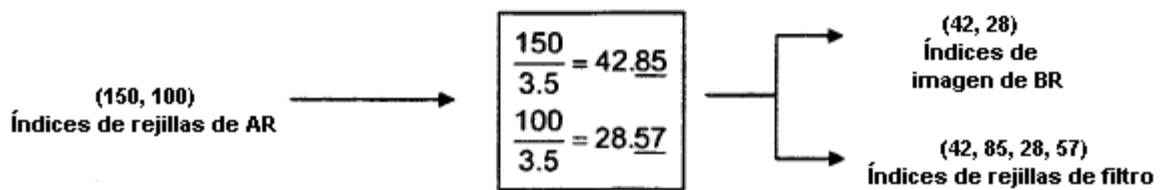


Figura 6

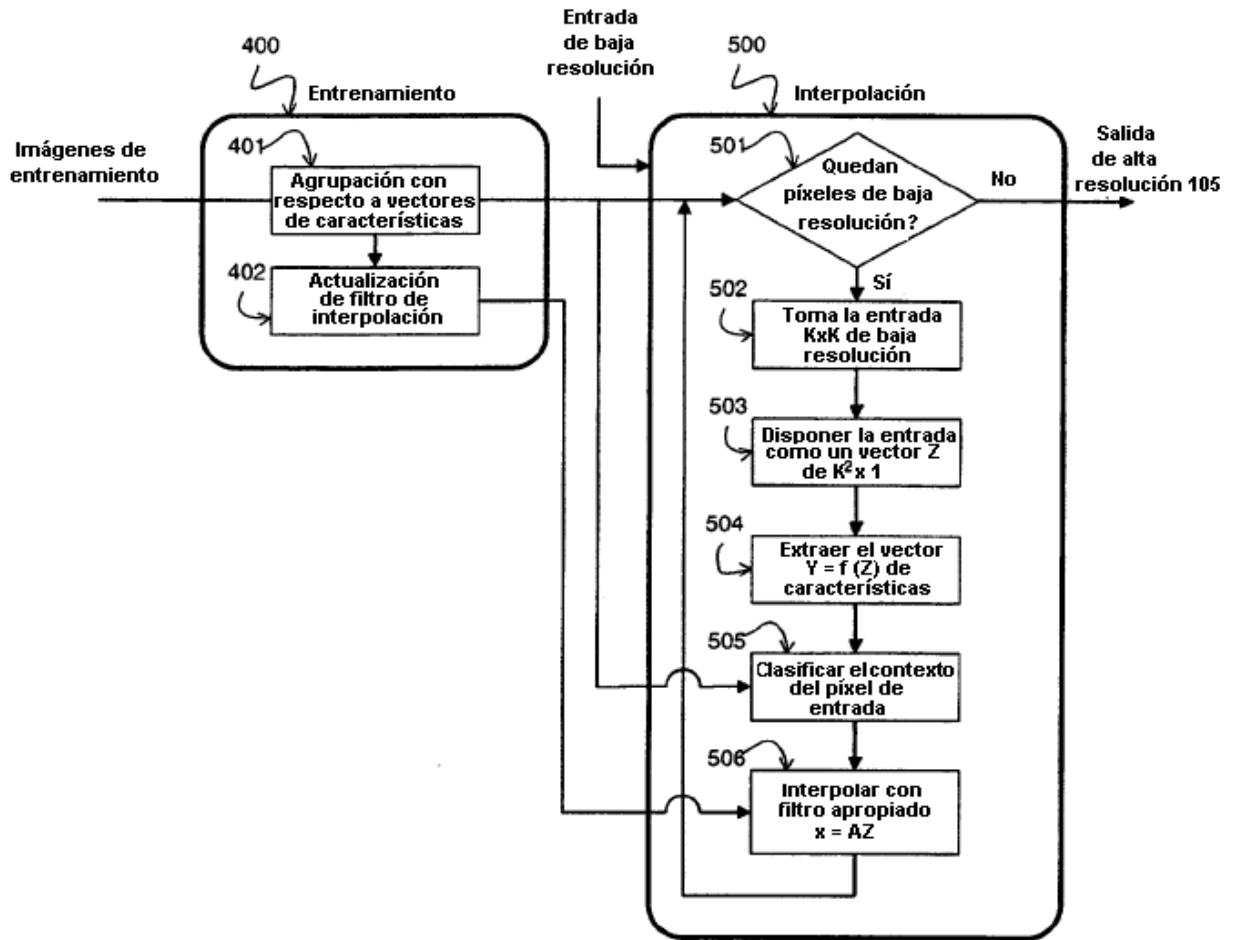


Figura 8