

IDENTIFICACIÓN DE LOS DATOS DEL PROYECTO.

Tema y Título: IMAGEN Y TELEVISIÓN. "Análisis de imágenes basado en correlación de imágenes"

Autor: Eduardo Martín Pastor

Titulación: Sonido e Imagen

Tutor: José Manuel Díaz López

Departamento: DIAC

Tribunal:

Presidente: Salvador Sánchez Fernández

Vocal: José Manuel Díaz López

Vocal Secretario: José Luis Rodríguez Vázquez

RESUMEN DEL PROYECTO

El objetivo principal del proyecto es la realización de una aplicación en el programa MATLAB.

En primer lugar, realizaremos un estudio teórico relativo al tema de nuestro proyecto. En nuestro caso como el tema es Imagen y Televisión, explicaremos de forma teórica la información principal acerca del Tratamiento Digital de la Imagen.

Una vez conocida las técnicas principales utilizadas en el tratamiento digital, realizaremos un estudio exhaustivo en las técnicas actuales que existen acerca del análisis de imágenes. Daremos una breve explicación mostrando en qué consiste esta técnica, los diferentes pasos que se llevan a cabo en una imagen para su análisis, explicando brevemente cada uno de ellos y enumerando algunas técnicas para la realización de cada una de ellas.

Tras esta primera parte, nos centraremos en las técnicas de correlación de imágenes (DIC). Explicaremos como han surgido estas técnicas, cual son sus principales conceptos, sus inicios y las ventajas e inconvenientes que tienen.

Dentro de las diferentes técnicas de correlación de imágenes, explicaremos de forma detallada la correspondencia por áreas, ya que es la técnica que vamos a utilizar para la realización del proyecto. Explicaremos en qué consiste, y desarrollaremos teóricamente cual son los pasos que se deben realizar en las imágenes para realizar esta técnica. Explicaremos cual es su terminología, y cuáles son los posibles defectos que puede tener esta técnica.

Finalmente, una vez estudiada la teoría, realizaremos una sencilla aplicación que nos permita evaluar y encontrar las diferencias en una secuencia de imágenes. El programa utilizado para este proyecto es MATLAB, que es un programa matemático, utilizado enormemente en el ámbito de la ingeniería. Mediante esta aplicación obtendremos dos figuras, una de ellas donde veremos los vectores de movimiento que existen entre las dos imágenes y la segunda, donde obtendremos el factor de correlación que hay entre las dos imágenes.

ABSTRACT OF MY PROJECT

The main objective of the project is the development of an application in MATLAB program.

Firstly carry out a theoretical study on the topic of our project. In our case as the theme is Picture and Television, we explain the main information about Digital Image Processing.

Once known the main techniques used in digital images, we will make a study on current techniques that exist about image analysis. We will give a brief explanation showing what this technique is, the different steps that are performed on an image for analysis, briefly explaining each of them and listing some techniques for performing each.

After this first part, we will focus on the techniques of image correlation (DIC). We explain how these techniques have emerged, which are the main concepts, the beginning and the advantages and disadvantages they have.

There are different image correlation techniques. We will explain in detail the correspondence areas, as it is the technique that we will use for the project. Explain what it is, which is theoretically and we develop steps that must be performed on the images for this technique. We explain what their terminology is, and what are the possible defects that may have this technique.

Finally, having explored the theory images, we will make a simple application that allows us to evaluate and find differences in a sequence of images. The program used for this project is MATLAB, a mathematical program, widely used in the field of engineering. Using this application will get two figures, one where we will see the motion vectors between the two images and the second where we get the correlation factor between the two images.

Análisis de imágenes basado en correlación de imágenes

Proyecto fin de
carrera

Eduardo Martín Pastor

Quiero dar las gracias a todas las personas que me han ayudado durante estos 5 años de carrera, especialmente a mi familia y amigos, que siempre han estado a mi disposición tanto en los buenos como en los malos momentos.



Análisis de imágenes basado en correlación de imágenes

Contenido

1. Planteamiento del proyecto.....	1
2. Estudios previos	2
2.1 Teoría tratamiento digital de la imagen (TDI)	2
2.1.1 Definición de la imagen digital	2
2.1.2 Sistema de procesamiento de imágenes	5
2.1.3 Transformaciones.....	7
3. Análisis de imágenes	46
3.1 Extracción de características de una imagen	46
3.1.1 Características especiales de una imagen.....	46
3.1.2 Características de la transformada	47
3.1.3 Detección de bordes, texturas y movimientos	48
3.1.4 Texturas.....	58
3.1.5 Detección de movimiento	60
3.2 Segmentación de imágenes	60
3.2.1 Segmentación basada en píxeles	61
3.2.2 Segmentación basada en bordes	66
3.2.3 Segmentación orientada a regiones.....	68
3.2.4 Segmentación basada en texturas	69
3.3 Transformaciones morfológicas.....	69
3.3.1 Transformaciones morfológicas en imágenes binarias.....	70
3.3.2 Transformaciones morfológicas en imágenes con varios niveles de gris	77
3.4 Representación y descripción de contornos y regiones	78
3.4.1 Representación del contorno.....	78
3.4.2 Representación de regiones.....	79
3.4.3 Descriptores de contornos.....	80
3.4.4 Descriptores de regiones.....	81

Análisis de imágenes basado en correlación de imágenes

4. Correlación Digital de Imágenes (DIC).....	82
4.1 Introducción	82
4.2 Antecedentes.	82
4.3 Conceptos básicos.	82
4.4 Dificultades en la correspondencia de imágenes.....	84
4.5 Distorsiones geométricas de las entidades de correspondencia	86
4.6 Estrategias de correspondencia por áreas	87
4.6.1 Correlación	89
4.6.2 Correspondencia por mínimos cuadrados	91
4.6.3 Correspondencia basada en características.....	94
4.6.4 Correspondencia relacional.....	99
5. Procedimiento Experimental.....	105
6. Conclusiones.....	109
7. Objetivos conseguidos	113
8. Bibliografía	114

1. Planteamiento del proyecto

Antes de introducirnos en el proyecto por completo, vamos a dar una pequeña explicación acerca de los conceptos que vamos a aclarar y explicar.

En primer lugar, daremos una pequeña definición acerca del tratamiento digital de la imagen. Explicaremos de forma resumida, las diferentes transformaciones que pueden sufrir las imágenes y la forma de evitar estas.

A continuación, una vez dada una breve pincelada sobre el tratamiento digital de la imagen, nos centraremos en el análisis de estas. Explicaremos que es y en qué consiste el análisis de imágenes y daremos a conocer los diferentes pasos en los que se divide esta técnica, explicando cada uno de ellos brevemente.

El siguiente paso será adentrarnos en una técnica de análisis de imágenes. Desarrollaremos las técnicas basadas en la correlación de imágenes, más concretamente la técnica de correspondencia por áreas, ya que es la que definiremos, explicaremos y nos servirá de ayuda tanto teóricamente como prácticamente, para la realización de un programa sencillo que nos permita evaluar y detectar cambios entre dos imágenes.

Para finalizar comentaremos como funciona nuestro programa y aportaremos una serie de mejoras que se podrían introducir en nuestro proyecto.

2. Estudios previos

2.1 Teoría tratamiento digital de la imagen (TDI)

En los siguientes apartados se hace un repaso de los conceptos básicos sobre Tratamiento Digital de Imágenes.

2.1.1 Definición de la imagen digital

En primer lugar, definiremos que es una imagen. Una imagen es la representación de una forma o de una entidad determinada. Para trabajar con ellas y realizarle todas las transformaciones que creamos oportunas definimos la imagen matemáticamente con la siguiente función:

$$f(x,y) \quad (\text{Ecuación 1})$$

Siendo x e y , las coordenadas espaciales en el plano, y f , el brillo o nivel de gris en cualquier par de las coordenadas anteriores.

Para nuestro proyecto, como herramientas vamos a trabajar con imágenes digitales monocromas. Sobre ellas llevaremos a cabo las técnicas y procesos que describiremos posteriormente.



Ilustración 1. Imagen con sus respectivos píxeles

Muestreo y cuantificación

Una imagen puede ser continua tanto respecto a sus coordenadas x e y , como a su amplitud. Para convertirla a forma digital, hay que digitalizarla en ambos sentidos.

Por ello podemos definir la imagen digital como una matriz cuyos índices de fila y columna identifican un punto de la imagen y el valor indica el nivel de gris en ese punto. A los elementos que componen esta matriz se les denomina píxeles.

$$f(x, y) \sim \begin{bmatrix} f(0,0)f(0,1).....f(0,M-1) \\ f(1,0)f(1,1).....f(1,M-1) \\ \\ \\ f(N-1,0)f(N-1,1)....f(N-1,M-1) \end{bmatrix} \quad (\text{Ecuación 2})$$

El proceso de seleccionar un conjunto de valores finitos, discretos de una señal continua se denomina muestreo (sampling) y los valores seleccionados se denominan muestras (samples). Las muestras seleccionadas continúan siendo continuas, es decir, tienen un valor real. El proceso de discretización de estos valores se denomina cuantificación y es el segundo paso en el proceso de digitalización de una señal. El tercer y último paso es la codificación de los resultados obtenidos.

Para concluir podemos asegurar, visto lo anterior, que la calidad de una imagen digital se determina en gran manera por el número de muestras y niveles de gris utilizados en el muestreo y cuantificación.

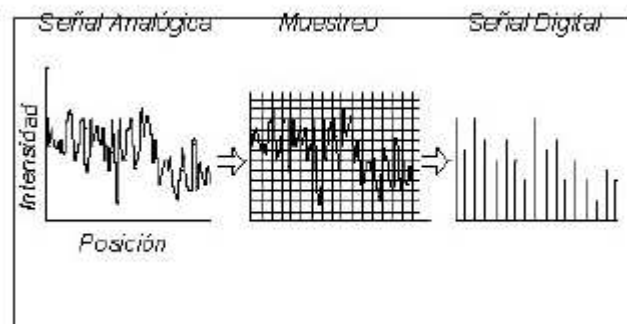


Ilustración 2. Muestreo y cuantificación.

Parámetros de control de la calidad de una imagen

La calidad de una imagen se mide en función de dos tipos de parámetros: los relativos a la extensión y los relativos a la intensidad o profundidad, conociendo esta última como la capacidad de distinguir diferencias de intensidad y de color en un mismo punto.

Resolución

Es la capacidad de un detector para discriminar detalles, concretamente, es la capacidad para diferenciar dos puntos muy próximos antes de que se fundan en uno solo. Dado que las imágenes digitales están formadas por puntos discontinuos que se perciben como un continuo a una determinada distancia, este parámetro tiene una vital importancia.

Numero de grises

La calidad de la imagen depende, como segundo factor principal, del número de valores de intensidad que es capaz de representar. Existe diferencia si trabajamos con una imagen acromática o cromática. Entendemos por imagen acromática, a la compuesta por una escala

Análisis de imágenes basado en correlación de imágenes

tonal de grises cuyo rango varía de 2, tal como un dibujo en blanco y negro (1 bit por pixel (bpp)). El parámetro bpp representa la información con la que contamos y permite saber el número de valores que nos proporcionará una imagen y que será igual a 2^n donde n es el valor de bpp.

En el caso de las imágenes cromáticas los parámetros que se tienen en cuenta son los 256 niveles de gris a través de tres canales RGB (Red, Green, Blue), rojo, verde y azul.

Perturbaciones características de las imágenes y métodos de corrección

Las perturbaciones típicas que se producen en las imágenes son el ruido, las interferencias, las manchas, las franjas, los bordes de colores, las pérdidas o alteraciones de tono, intensidad o saturación. Estas perturbaciones tienen su origen en fallos del dispositivo que impiden una respuesta adecuada. Las perturbaciones más características de las señales digitales son el contorneado, la pixelación y el aliasing. La más común en las señales analógicas es el ruido aunque también se produce en señales digitales.

Pixelación y contorneado

La pixelación y el contorneado son dos perturbaciones de las señales digitales que se producen por insuficiencia de la resolución en el primer caso y por insuficiencia del rango de intensidades en el segundo. En la siguiente figura podemos ver un ejemplo de pixelación y contorneado. En la ilustración se muestra un esquema que correspondería a una imagen determinada (izquierda) de la que se han suprimido la mitad de las muestras (centro) lo que produciría un efecto de pixelación, o la mitad de los valores (derecha) lo que produciría un efecto de contorneado.

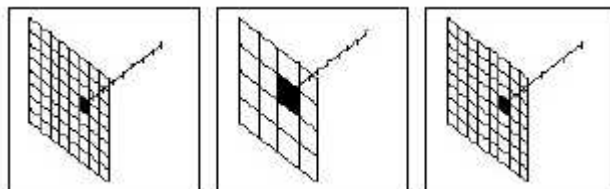


Ilustración 3. Pérdida de calidad de una imagen (a) por insuficiente muestreo (b) y cuantificación (c).

Ruido

Es una señal no deseada que se introduce en un sistema de transmisión. Puede tener diversos orígenes pero generalmente se consideran como errores de medición superpuestos a la imagen. Por ejemplo, el ruido de fondo es la agitación de electrones presentes en el canal antes del paso de los electrones propios de la señal. La única forma de evitar el ruido es disminuir la capacidad del canal.

La unidad de medida del ruido son los decibelios (dB) referidos a números S/N. En general todos los dispositivos introducen cierto grado de ruido a la imagen, tanto si se ha generado internamente o se capta por algún dispositivo. Se distinguen dos grandes tipos de ruido:

Análisis de imágenes basado en correlación de imágenes

- El ruido determinado o no aleatorio se genera por un sistema concreto y tiene siempre un aspecto similar: franjas o bandas o nubes de puntos de determinadas características de tamaño y orientación. Se pueden filtrar y eliminar si se conoce su patrón de recurrencia.
- El ruido indeterminado o aleatorio se genera por diversas causas que no se pueden identificar o corregir pues su distribución es aleatoria y cambia con el tiempo. La única forma de eliminar este ruido es introducir filtros a la imagen.

Aliasing

Es la aparición de fenómenos extraños que alteran notablemente el aspecto de la imagen original, suprimiendo información relevante o alterando la existente hasta volverla equívoca. Estos fenómenos se producen en el momento de la reconstrucción, por pérdida o modificación, debido a que algunos de los componentes de la señal original estaban comprendidos en frecuencias superiores al límite de Nyquist y ha sido reconstruido en frecuencias más bajas.

Antialiasing

Son el conjunto de técnicas que se utilizan para prevenir o disimular los defectos comprendidos genéricamente bajo la denominación de aliasing.

2.1.2 Sistema de procesamiento de imágenes

Cuando hablamos de tratamiento digital de las imágenes abarcamos dos áreas principales de aplicación. En primer lugar, el procesamiento de imágenes con la finalidad de mejorar la información de modo que sea capaz de ser asimilada por observadores humanos, y el procesamiento y análisis de imágenes, con la finalidad de hacer que sean reconocidas por computadores especializados.

Por procesamiento se entiende al conjunto de técnicas y métodos encaminados a mejorar una imagen, con el fin de la independencia de su contenido, facilitar su posterior interpretación. Ejemplos de procesamiento son el tratamiento de radiografías para realzar zonas de interés, el procesamiento de fotografías en mal estado para facilitar el reconocimiento de personas, etc.

Por análisis de imágenes se alude al conjunto de técnicas y métodos encaminados a facilitar la extracción de la información contenida en ellas. La finalidad principal sería llegar a hacer eficaz la visión por computador, es decir, conseguir que una máquina dotada de sensores pueda reconocer fragmentos más o menos grandes del espacio y los objetos que le rodean y ajustar su conducta de acuerdo con esta información. El proceso implica la detección, la identificación y la comprensión de patrones que son relevantes en una aplicación determinada. Un ejemplo de análisis de imágenes puede ser el uso de robots en cadenas de producción que pueden reconocer si un producto tiene defectos y debe ser retirado de la cadena, o el uso de robots en procesos de fabricación para reconocer formas características.

A continuación, vamos a presentar de forma genérica, las etapas fundamentales de un sistema de procesado de imágenes y una pequeña descripción de cada una de las etapas.

2.1.2.1 Adquisición de la imagen

Consiste en la adquisición mediante un conjunto de dispositivos de la imagen digital. Estos dispositivos concretamente son un sensor de imágenes y un digitalizador. El primero es un dispositivo físico sensible a una determinada banda del espectro de energía electromagnética (banda de rayos X, ultravioleta, visible o infrarrojo) y que produzca una señal eléctrica de salida proporcional al nivel de energía detectado. El segundo, el digitalizador, es un utensilio que se encarga de convertir la señal de salida del sistema sensible a forma digital.

2.1.2.2 Almacenamiento

Consiste en proporcionar la capacidad de almacenamiento adecuada para la adquisición de la imagen. Es un reto importante en el diseño de los sistemas de tratamiento de imágenes, ya que por ejemplo una imagen de 8 bit necesita un millón de bytes de memoria. Dentro del almacenamiento tenemos tres categorías:

- Almacenamiento a corto plazo: Para esto podemos utilizar la memoria del ordenador, o bien externalizar esta operación utilizando tarjetas de memoria especializadas denominadas memorias temporales, que almacenan una o más imágenes a las que puede accederse con rapidez.
- Almacenamiento en línea: Para reutilizarla rápidamente.
- Almacenamiento en archivo: Caracterizado por un acceso poco frecuente y un almacenamiento pasivo.

2.1.2.3 Preprocesamiento

Durante esta fase se llevan a cabo una serie de acciones que tienen por finalidad facilitar el trabajo posterior; acciones tales como eliminar ruidos o calibrar adecuadamente los rangos monocromáticos y cromáticos, el contraste o la definición de las diferentes áreas.

2.1.2.4 Segmentación

Se encarga de separar la imagen en sus partes constituyentes, con lo que entramos en el dominio del análisis y, con ello, en el núcleo de los problemas propios del procesamiento de imágenes. La segmentación supone la detección de puntos, líneas y bordes. A partir de este momento, se busca identificar bordes continuos, fronteras y regiones. La segmentación depende por lo general de la aplicación que dicta las propiedades visuales de los elementos de interés cuya detección se busca. El método más general se basa en la detección de discontinuidades y el problema más general deriva de que las discontinuidades resultan de bordes significativos.

2.1.2.5 Representación y descripción

Como resultado del proceso anterior, la imagen se habrá diferenciado entre píxeles correspondientes a bordes y píxeles correspondientes a regiones. Para que estos

Análisis de imágenes basado en correlación de imágenes

procesos resulten más adecuados se dependerá del contexto. Si lo que se busca es diferenciar figuras de fondos, la segmentación en bordes resultará más adecuada, sin embargo en el caso de identificar propiedades materiales será preferible identificar regiones. La representación consiste fundamentalmente en una selección de características que se realzan y se procesan para obtener datos adicionales dados generalmente en forma numérica o mediante especificaciones acerca de la topología de la imagen.

2.1.2.6 Reconocimiento e interpretación

Una vez realizada la descripción de la imagen es posible comparar los datos obtenidos con los datos contenidos en una base de datos y asignar etiquetas identificadoras a los distintos elementos. Este proceso de asignación se denomina reconocimiento. En cambio por interpretación se alude al paso final de asignar un significado a un conjunto de objetos etiquetados.

2.1.3 Transformaciones

2.1.3.1 Interpolación

Al realizar transformaciones afines a las imágenes se presentan cambios en la posición de los píxeles respecto a su posición inicial en el plano imagen y por lo tanto se requieren métodos de interpolación numérica para los valores de intensidad que se asignarán a las nuevas posiciones en el plano imagen.

Relacionado con la imagen digital consiste en la creación de nuevos píxeles a partir de los existentes, de modo que la imagen resultante represente lo mismo que la original, solo que con más píxeles.

Hay tres tipos de interpolación diferente:

1. Interpolación del vecino más cercano

Este es el algoritmo de interpolación más simple posible. La interpolación del vecino más próximo selecciona el valor del píxel más cercano redondeando las coordenadas del punto de interpolación deseado.

Aunque se trata de un método sencillo provoca efectos indeseables en las imágenes por lo que no se aconseja su uso para el suavizado de imágenes.

El proceso es el siguiente: se recorre pixel a pixel la imagen original y para cada pixel crea varios iguales, según la interpolación que se realice, y los introduce en la imagen definitiva.

2. Interpolación bilineal

Es un método algo más sofisticado que el anterior. El algoritmo bilineal interpola a partir de los cuatro píxeles adyacentes que rodean al punto deseado de

Análisis de imágenes basado en correlación de imágenes

interpolación. Se basa en obtener la solución de los coeficientes de la siguiente ecuación:

$$u(x, y) = c_1x + c_2y + c_3xy + c_4 \quad (\text{Ecuación 3})$$

Donde $u(x,y)$ representa la intensidad del píxel, y (x,y) la coordenada donde se redefinirá la intensidad del píxel.

Se obtienen los coeficientes usando la fórmula anterior en los cuatro píxeles de la vecindad que se pretende interpolar.

3. Interpolación bicúbica

Interpolan a partir de los dieciséis píxeles de la imagen original adyacentes al píxel sobre el que se quiere realizar la interpolación. El valor de la intensidad aplicado al píxel responde a:

$$u(x, y) = \sum_{j=0}^3 \sum_{i=0}^3 a_{ji} x^i y^j \quad (\text{Ecuación 4})$$

Donde los dieciséis coeficientes a_{ji} se obtienen de las dieciséis ecuaciones de la forma anterior.

Como podemos observar la ecuación de la interpolación bicúbica se convierte en la bilineal cuando el término del sumatorio es de 0 a 1.

2.1.3.2 Tratamiento estadístico de imágenes

2.1.3.2.1 Brillo medio

En una imagen es la suma de todos los brillos dividido por el número total de píxeles de la imagen.

$$\mu = x = \frac{1}{M \cdot N} \sum_{u=0}^{255} u_i n_i \quad (\text{Ecuación 5})$$

Siendo M y N las dimensiones de la imagen, u_i el valor de brillo de 0 a 255 y n_i es el número de píxeles que tienen ese valor de brillo

2.1.3.2.2 Varianza de una imagen

Valora la separación existente entre los brillos de los píxeles y el valor medio (elevado al cuadrado para no compensar variaciones positivas con negativas).

$$\sigma^2 = \frac{1}{M \cdot N} \sum_{i=1}^M \sum_{j=1}^N [x(i, j) - \mu]^2 \quad (\text{Ecuación 6})$$

2.1.3.2.3 Desviación típica

Es la raíz cuadrada de la varianza.

Desviación típica $\rightarrow \sigma$ (Ecuación 7)

2.1.3.2.4 Histograma

El histograma de una imagen es el gráfico en el que se representan las veces que aparece cada brillo. En el eje de abscisas del gráfico se representa los distintos brillos de la imagen y en el eje de ordenadas aparece la frecuencia relativa con la que cada color aparece en la imagen. El histograma aporta información sobre el brillo y el contraste de la imagen, pudiendo ser utilizado para ajustar estos parámetros.

Un pico en nuestro histograma en el lado izquierdo indica un gran número de píxeles que están oscuros o negros, mientras que un pico en la parte derecha indica un gran número de píxeles luminosos o blancos.



Imagen original

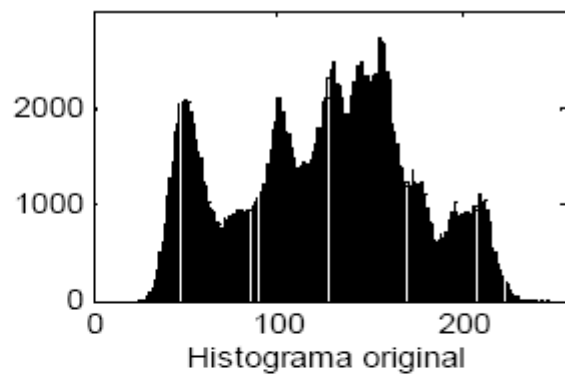


Imagen ecualizada

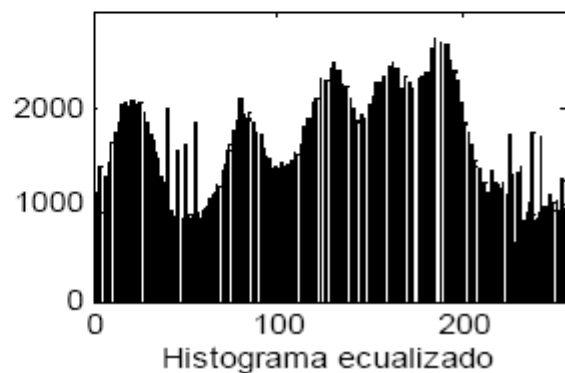


Ilustración 4. Histograma de una imagen monocroma original y ecualizada.

Para convertir el histograma en una función de densidad de probabilidad, podemos normalizarlo provocando que la variable representada haya sido escalada para ajustarse a un rango entre 0 y 1. Para ello las ordenadas se dividen entre el número total de píxeles.

Análisis de imágenes basado en correlación de imágenes

Una forma de manipulación del histograma es su ecualización que reduce automáticamente el contraste en las áreas muy claras o muy oscuras de la imagen. Consiste en una transformación no lineal que considera la distribución acumulativa de la imagen original, para generar una imagen resultante cuyo histograma será aproximadamente uniforme. Obviamente esto no es posible debido a la naturaleza discreta de los datos digitales de una imagen. Sin embargo, se consigue una aproximación al dispersar los picos del histograma de la imagen, dejando intactas las partes más bajas.

2.1.3.3 Operaciones de punto

Las operaciones orientadas al punto transforman a la imagen modificando un píxel a la vez, sin importar el estado de los píxeles vecinos.

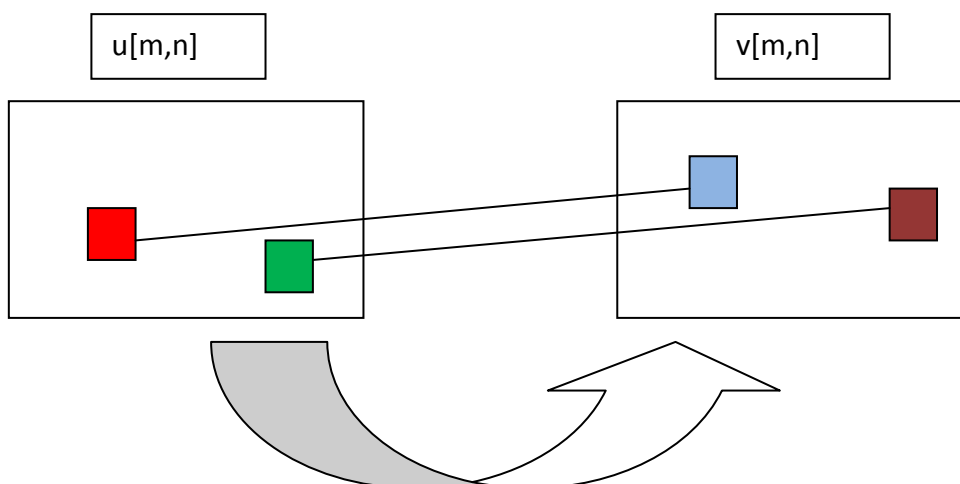


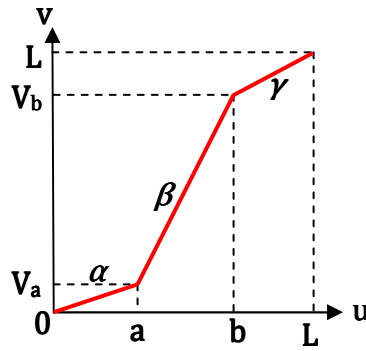
Ilustración 5. Operaciones de punto

1. Stretching

Conocida como estiramiento o ampliación de contraste se aplica en aquellas imágenes con poco contraste. Una imagen con poco contraste es aquella en la que los valores de los niveles de gris de los píxeles que la componen toman valores próximos o iguales, haciendo difícil diferenciar a simple vista los valores entre unos y otros. Esto lo podríamos ver claramente en el histograma, ya que en una pequeña zona de niveles de píxeles habría mucha frecuencia de aparición.

Para solucionar el problema tenemos que conseguir que todos los valores de los píxeles que ocupan un entorno pequeño, pasen a ocupar un entorno más grande del rango dinámico total, lo cual equivale a estirar la parte del histograma donde se concentran los valores de más aparición. La función a utilizar es la siguiente:

$$v = f(u) = \begin{cases} \alpha u + V_0 & 0 \leq u < a \\ \beta(u - a) + v_a & a \leq u < b \\ \gamma(u - b) + v_b & b \leq u < L \end{cases} \quad (\text{Ecuación 8})$$



Gráfica 1

Se puede observar tres intervalos para los valores que toman los píxeles de u , la cual es nuestra imagen original. Si tenemos un intervalo que abarca la mayor frecuencia de aparición, la pendiente de nuestra función deberá ser mayor que la unidad para que así se produzca un estiramiento.

En el caso de que un intervalo se estire, quiere decir que por otro lado, los otros dos intervalos tienen que comprimirse, ya que el rango dinámico se mantiene constante, es decir, siguen existiendo el mismo número de niveles ($L+1$). Esto supone que un número de niveles de la variable u pasan a ocupar un número menor de niveles de v .

En los niveles menos frecuentes, hay una pérdida de información, pero no es importante ya que la mayor parte de la información total se encuentra en el rango de valores más habituales.

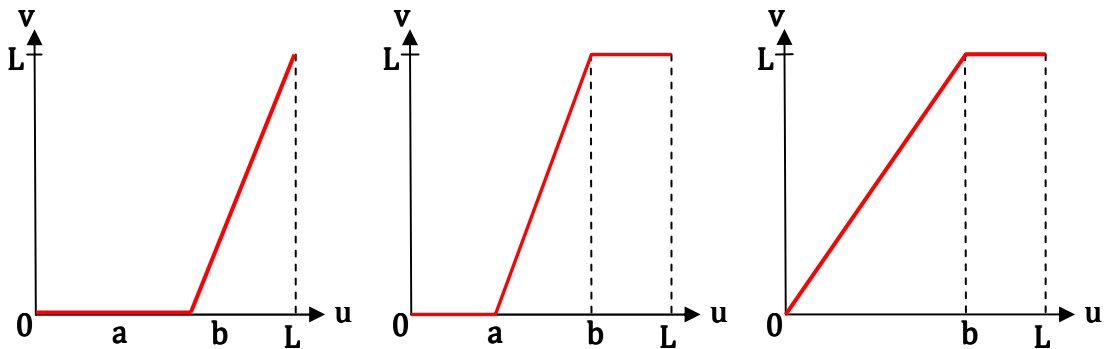
2. Clipping

Se trata de un caso especial de la técnica de ampliación de contraste. Pueden darse tres casos: recorte por la parte inferior, recorte por la parte superior y recorte por ambos lados.

$$v = f(u) = \begin{cases} v = 0 & \text{si } u < a \\ v = \frac{L(u-a)}{L-a} & \text{si } u > a \end{cases} \quad (\text{Ecuación 9})$$

$$v = f(u) = \begin{cases} v = 0 & \text{si } u < a \\ v = \frac{L(u-a)}{L-a} & \text{si } u > a \end{cases} \quad (\text{Ecuación 10})$$

$$v = f(u) = \begin{cases} v = 0 & \text{si } u < a \\ v = \frac{L(u-a)}{L-a} & \text{si } a < u < b \\ v = L & \text{si } u > b \end{cases} \quad (\text{Ecuación 11})$$



Gráfica 2. Relativas a las ecuaciones 9,10 y 11

El clipping sirve para ampliar de una forma drástica el contraste de un número reducido de niveles de gris. Este rango de niveles suele situarse donde se desea resaltar el contraste, permitiendo diferenciar los niveles que están por encima y por debajo de dicho rango. Estos niveles de gris corresponden a los que están en el intervalo que podemos observar en las funciones y su correspondiente ecuación [a,b]. En este caso, los niveles que están por debajo del umbral “a”, pasan a valer un único valor, cero, mientras que los niveles que se encuentran por encima de “b”, valgan el valor máximo, L. Por tanto podemos decir que una operación de clipping corresponde a un stretching donde los valores de niveles fuera del intervalo [a,b] sufren la máxima compresión. Esta técnica es muy buena para reducir el ruido cuando se sabe que los valores de brillo caen en un cierto rango de valores.

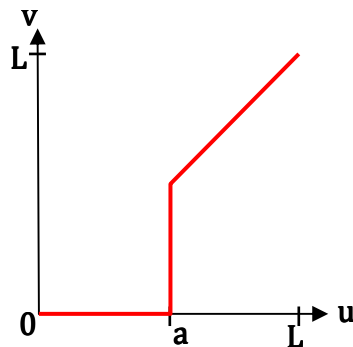
3. Umbralización.

Al igual que el clipping, se trata de un caso particular de ampliación de contraste. Mediante el proceso de umbralización, se divide el histograma de la imagen usando un umbral único. Dependiendo si el tipo de umbralización es superior, inferior o binarización, algunos píxeles se forzaran a blanco, otros a negro y otros se mantendrán.

- Umbralización Inferior

Se respeta el brillo de los píxeles que alcanzan un determinado valor. Los que tienen menor brillo que ese se colocan a blanco o a negro.

$$p(z) = \begin{cases} \frac{a^b z^{b-1} e^{-az}}{(b-1)!} & \text{para } z \geq a \\ 0 & \text{para } z < a \end{cases} \quad (\text{Ecuación 12})$$

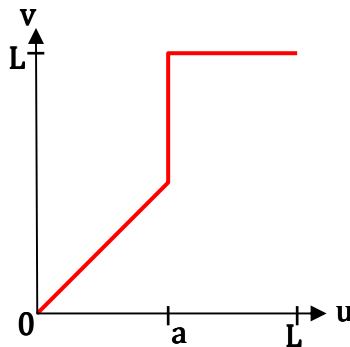


Gráfica 3

- Umbralización Superior

Se respeta el brillo de los píxeles que no alcanzan un determinado valor. Los que tienen mayor brillo que ese se colocan a blanco o a negro.

$$v = f(u) = \begin{cases} v = u & \text{si } u < a \\ v = L & \text{si } u > a \end{cases} \quad (\text{Ecuación 13})$$



Gráfica 4

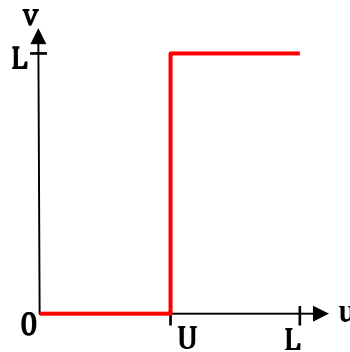
- Binarización

Los píxeles cuyo nivel no superan el umbral se fuerzan a blanco y los que lo superan, a negro. La imagen resultante de esta operación será una imagen binaria. Esta binarización es muy interesante en ciertas aplicaciones ya que puede servir para resaltar las partes brillantes de una imagen, que suelen corresponder a los objetos situados sobre un fondo oscuro. Otra aplicación bastante útil de esta operación, es el tratamiento de imágenes que tienen carácter binario. Un ejemplo de este tipo es un papel escrito, que al pasarlo por un scanner, puede que la imagen obtenida no sea binaria debido a la iluminación del fondo y al ruido

Análisis de imágenes basado en correlación de imágenes

introducido por el sensor. Con la binarización podemos convertir dicha imagen no binaria en una binaria.

$$v = f(u) = \begin{cases} v = 0 & \text{si } u < a \\ v = L & \text{si } u > a \end{cases} \quad (\text{Ecuación 14})$$



Gráfica 5

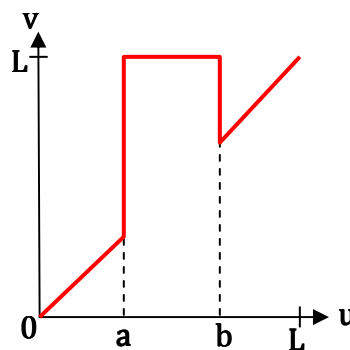
4. Slicing

Con el este proceso se resaltan zonas de la imagen con un brillo determinado. Lo resaltado se puede mandar a blanco o a negro.

Esto se aplica cuando queremos destacar un determinado rango de niveles de gris debido a que contiene información importante de la imagen.

- Conservando el fondo:

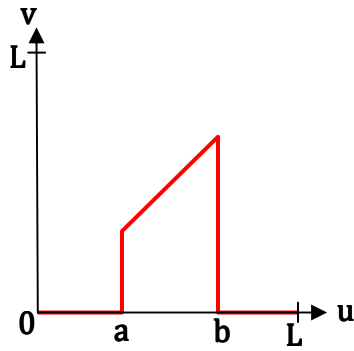
$$v = f(u) = \begin{cases} L & \text{si } a < u < b \\ u & \text{resto} \end{cases} \quad (\text{Ecuación 15})$$



Gráfica 6

- No conservando el fondo:

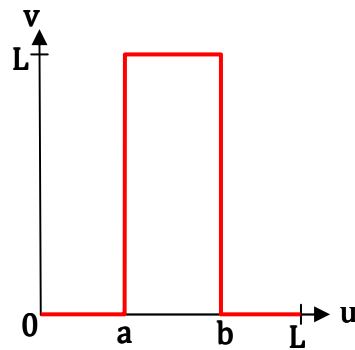
$$v = f(u) = \begin{cases} L & \text{si } a < u < b \\ u & \text{resto} \end{cases} \quad (\text{Ecuación 16})$$



Gráfica 7

- No conservando el fondo, umbralizado:

$$v = f(u) = \begin{cases} L & \text{si } a < u < b \\ u & \text{resto} \end{cases} \quad (\text{Ecuación 17})$$

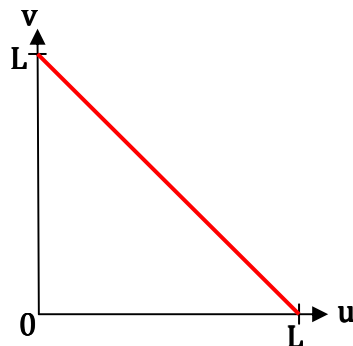


Gráfica 8

5. Complementación

También conocido como brillo negativo, y consiste en sustituir cada píxel de la imagen por su complementario.

$$v = f(u) = L - u \quad (\text{Ecuación 18})$$



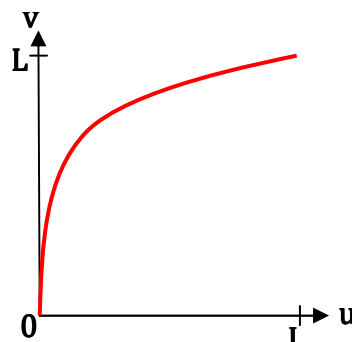
Gráfica 9

6. Operación logaritmo

A menudo el rango de dinámico de los píxeles de una imagen excede la capacidad del dispositivo de visualización, por lo que hay muchos píxeles que no pueden ser visualizados. Para solucionar este problema utilizamos la operación logaritmo que comprime el rango dinámico ajustando la constante para que el resultado sea el esperado.

$$v = c \cdot \log(1 + |u|) \quad (\text{Ecuación 19})$$

Otro ejemplo sería para mejorar el contraste en los niveles de gris más bajos comprimiendo el rango de valores más altos, por lo que este procedimiento es muy útil para mejorar el contraste en imágenes oscuras.



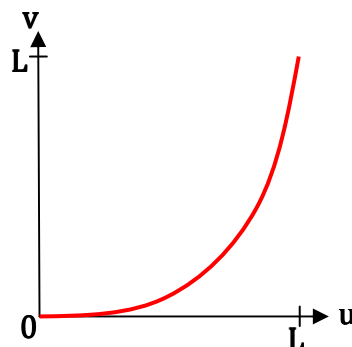
Gráfica 10

7. Operación exponencial

Produce el efecto contrario a la operación logarítmica. Su utilidad es ampliar el rango dinámico del dispositivo de salida y obtener una mejor definición.

La función utilizada para cada píxel es:

$$v = 10^{u/c} - 1 \quad (\text{Ecuación 20})$$



Gráfica 11

Análisis de imágenes basado en correlación de imágenes

Como su función es contraria a la operación logarítmica, la operación exponencial la utilizaremos para mejorar el contraste en imágenes claras.

2.1.3.4 Operaciones aritméticas

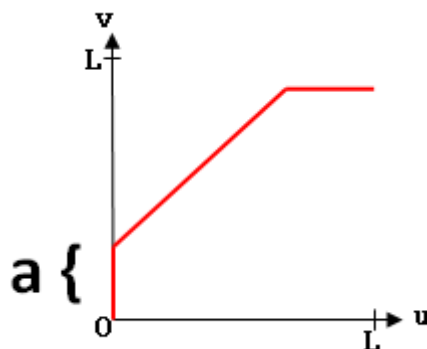
1. Operaciones aritméticas con una imagen

- Operación suma

En esta operación le sumamos al brillo de cada píxel un valor constante, incrementando el brillo de la imagen este determinado valor. Esto lo vemos reflejado en el histograma que se desplaza, tantos píxeles como lo hayamos desplazado la imagen.

$$R(x, y) = A(x, y) + a \quad (\text{Ecuación 21})$$

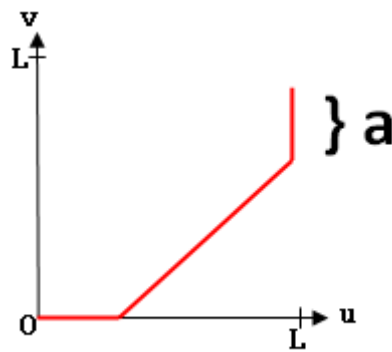
Debido a este desplazamiento pueden existir píxeles cuyo valor de brillo superen el máximo soportado por el dispositivo de salida (overflow), por lo que habrá que recortar estos valores al máximo permitido por el dispositivo. Este hecho se conoce como saturación y supone una pérdida de información con respecto a la imagen original.



Gráfica 12

- Operación resta

En esta operación le restamos el brillo de cada píxel un valor constante, decrementando el brillo de la imagen este determinado valor. Esto lo vemos reflejado en el histograma que se desplaza, tantos píxeles como lo hayamos desplazado la imagen.

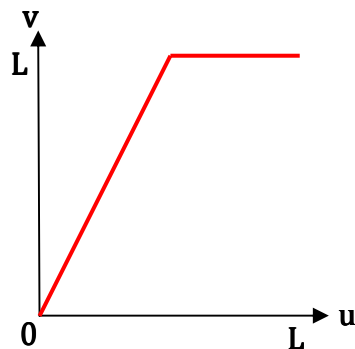


Gráfica 13

- Operación multiplicación

Aumenta la intensidad de la imagen en un valor constante. El efecto de esta operación se muestra en el histograma, en el cual se produce un estiramiento hacia la izquierda.

$$R(x, y) = b * A(x, y) \quad (\text{Ecuación 23})$$



Gráfica 14

Tanto en la suma como en la multiplicación, se aumenta el nivel de gris de los píxeles, pero de forma distinta. En la suma, el valor constante indica el número de niveles de grises a aumentar. Sin embargo, en la multiplicación b significa el factor a multiplicar. Por ejemplo si:

- $B = 1$ Ningún cambio
- $B = 2$ Se duplica el valor de gris. Los píxeles > 127 se saturan.
- $B = 0.5$ Se encoge a la mitad el histograma.
- $B < 1$ Sería la operación división.

2. Operaciones aritméticas. Combinación de imágenes

- Suma

Sumar dos imágenes da como resultado otra imagen que es una mezcla de ambas.

$$R(x, y) = A(x, y) + B(x, y) \quad (\text{Ecuación 24})$$

Antes de sumar las imágenes tenemos que tener en cuenta la saturación de la imagen resultante, es conveniente reducir previamente el rango de las imágenes a sumar. Para ello se dividen los valores de los píxeles entre el número de imágenes que participan en la suma produciéndose una imagen media de las originales.

$$R(x, y) = \frac{(A(x, y) + B(x, y))}{2} \quad (\text{Ecuación 25})$$

Como cada imagen se produce en instantes diferentes de tiempo, cada imagen tiene diferentes patrones de ruido. Al hacer la media de todas ellas se obtiene como resultado una imagen donde el ruido aleatorio ha sido atenuado. Esto es debido a que esta media refuerza los valores de los píxeles que no varían en ambas imágenes.

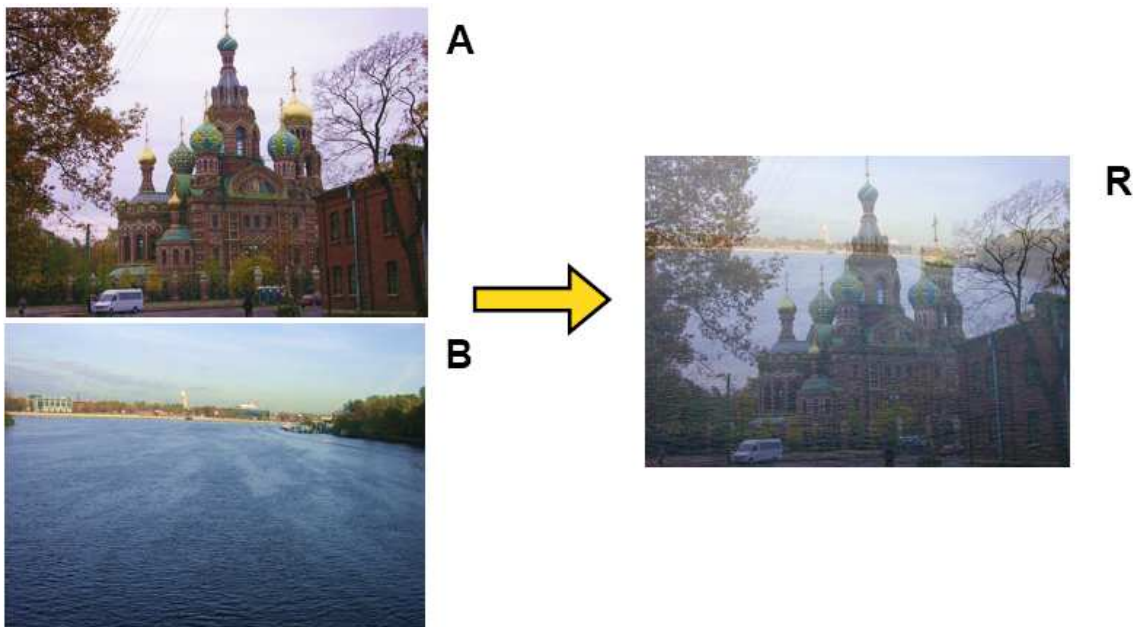


Ilustración 6. Operación aritmética (Suma)

Análisis de imágenes basado en correlación de imágenes

- Resta

Obtenemos una imagen que es la diferencia entre ambas. Esta técnica es de gran aplicación en segmentación y realce. Es una técnica útil para detectar el cambio producido en dos imágenes que han sido captadas en dos instantes de tiempo diferentes.

Como era de esperar puede ocurrir el caso contrario a la suma, en la que se produzca la saturación de los píxeles, pero en este caso en cero. Como en procesamiento se utilizan píxeles con números positivos podemos utilizar un reescalado donde el valor más negativo se corresponde con el 0 y el máximo con el 255 para la escala de grises de 0 a 255. Para ello aplicamos la siguiente función:

$$R(x, y) = \left(\frac{A(x, y)}{2} + 128 \right) - \frac{B(x, y)}{2} \quad (\text{Ecuación 26})$$

Si solo queremos conocer la diferencia entre las imágenes, entonces como solución tomamos el valor absoluto de la resta.

$$R(x, y) = \text{abs}(A(x, y) - B(x, y)) \quad (\text{Ecuación 27})$$

En este caso si los píxeles son negros, las imágenes son iguales, por el contrario cuando más clara es una zona más se diferencian las imágenes.

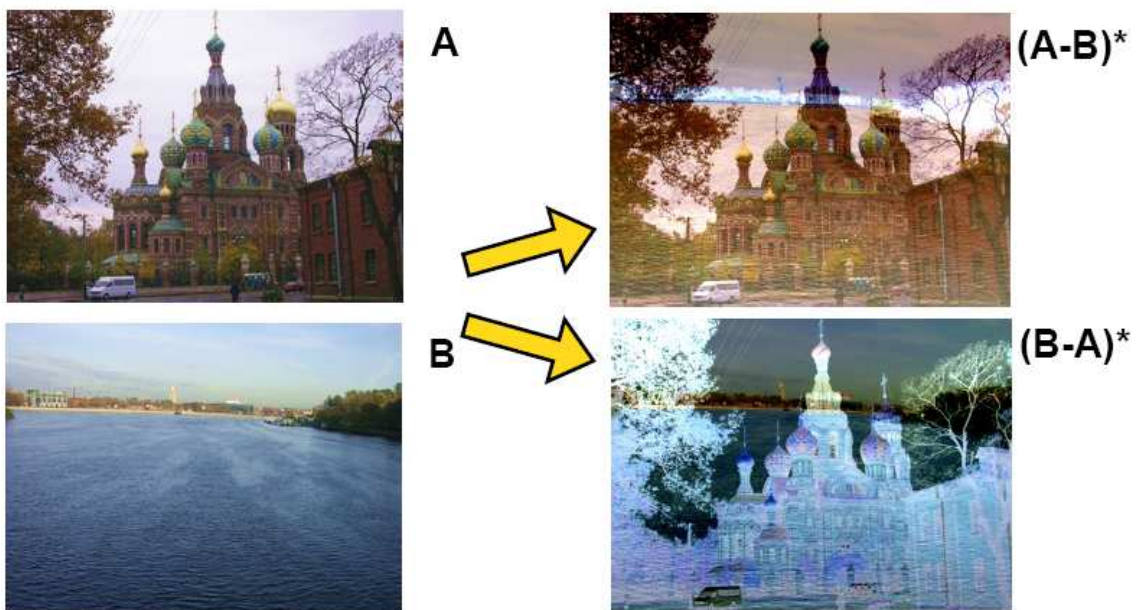


Ilustración 7. Operación aritmética (Resta)

- Producto

En primer lugar a la hora de multiplicar hay que escalar el resultado ya que la multiplicación de dos imágenes, produciría una totalmente saturada.

$$R(x, y) = \frac{A(x, y) * B(x, y)}{255} \quad (\text{Ecuación 28})$$

El producto de imágenes puede ser usado para realizar una transformación de intensidad diferente para cada píxel.

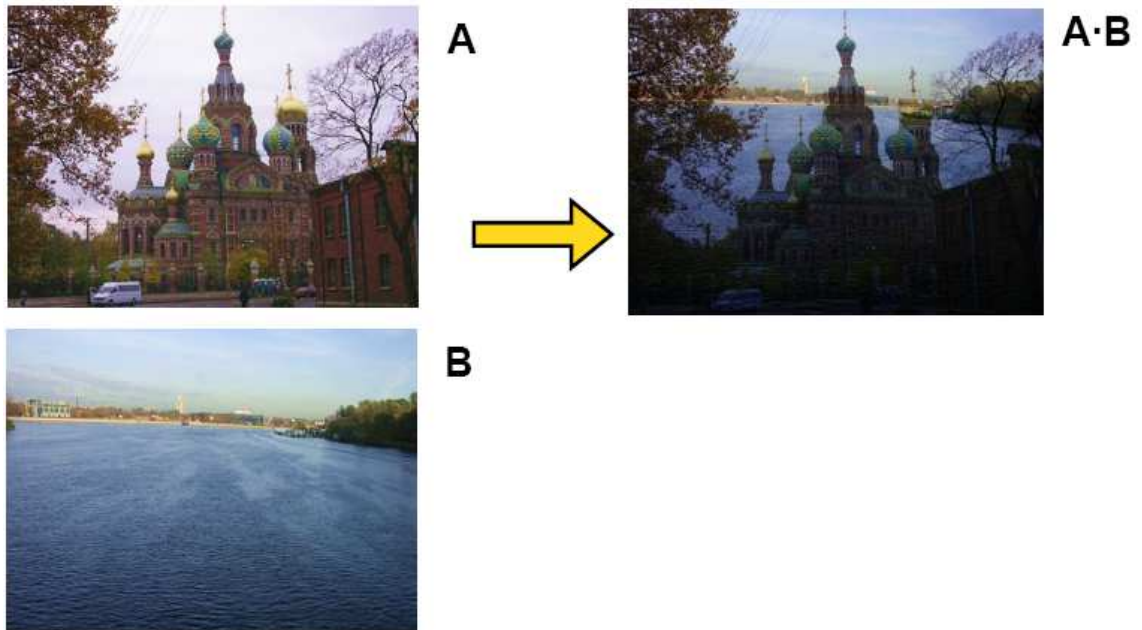


Ilustración 8. Operación aritmética (Multiplicación)

- División

La división es la operación contraria a la multiplicación, y por ello también debemos realizar un escalado pero esta vez multiplicando por 255.

$$R(x, y) = \frac{A(x, y)}{B(x, y)} * 255 \quad (\text{Ecuación 29})$$

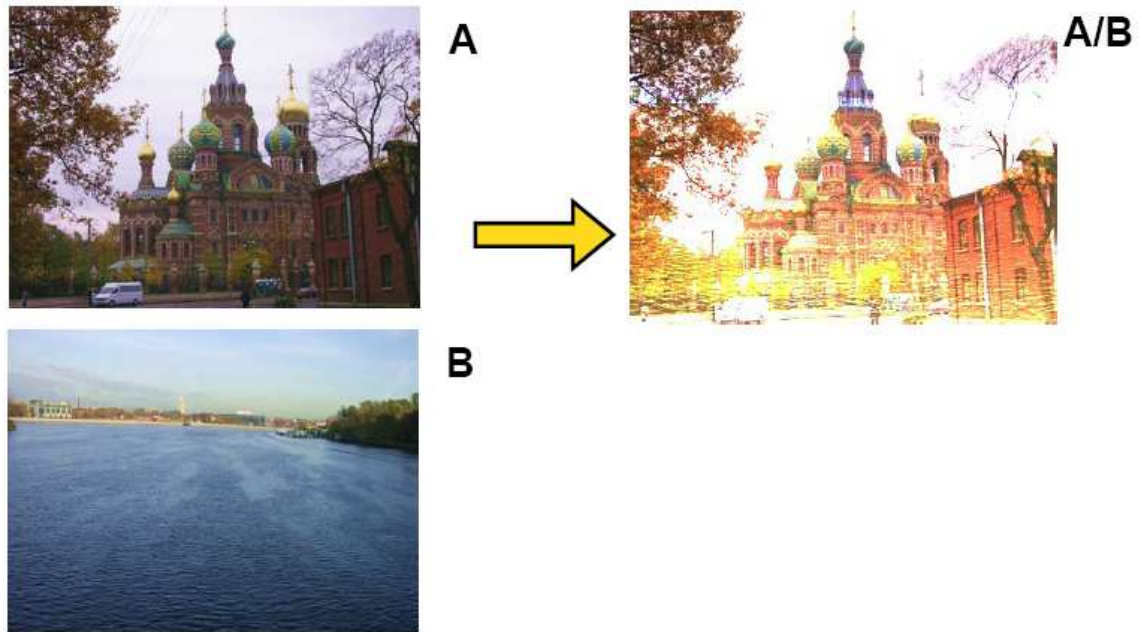


Ilustración 9. Operación aritmética (División)

3. Operaciones lógicas

Estos operadores tienen sentido cuando al menos una de las imágenes es binaria y actúa como máscara de la otra, de manera que el nivel de negro (0) es FALSE y el nivel de blanco (nivel máximo L-1) es TRUE.

Los principales operadores lógicos son:

- AND
- OR
- XOR
- NOT

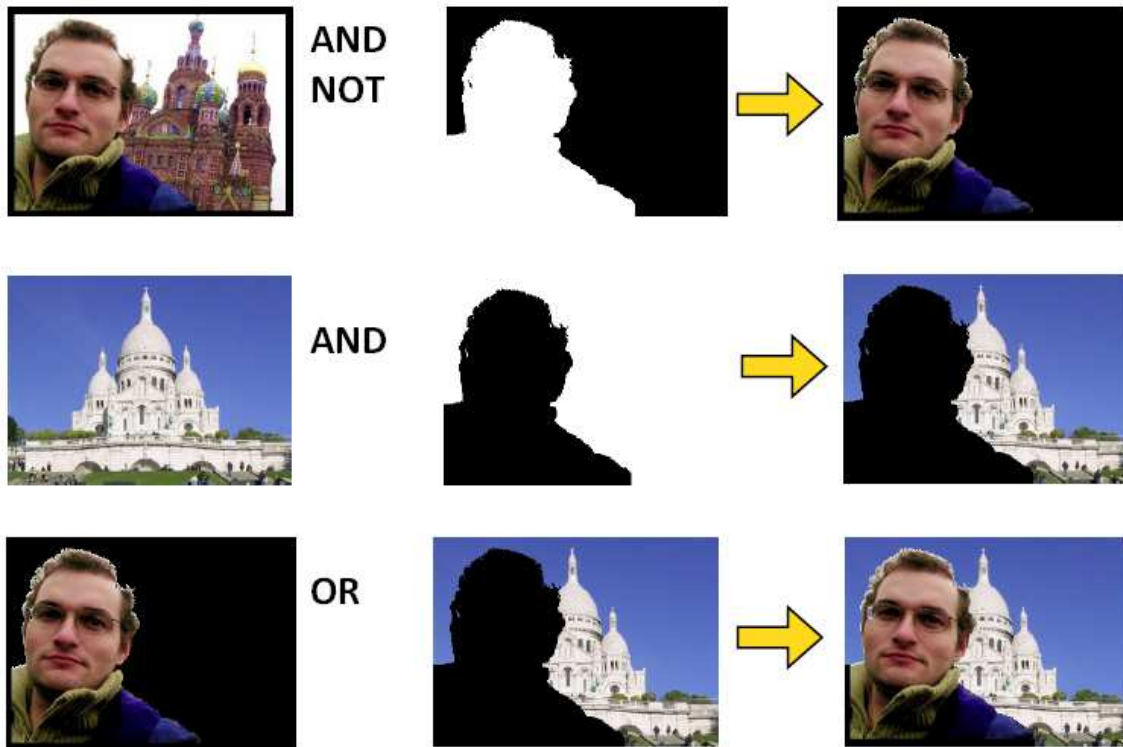


Ilustración 10. Operaciones lógicas

2.1.3.5 Introducción de ruido

Entendemos el ruido en una imagen digital como cualquier valor de un píxel que no se corresponde con la realidad. Siempre que se adquiere una imagen, ésta está contaminada por ruido. Suele ser debido a los equipos utilizados en la captación de las imágenes (ruido de cuantificación de la imagen, efecto de niebla, etc) y al ruido añadido en los tramos de transmisión (posibles interferencias o errores al transmitir los bits de información).

1. Ruido uniforme

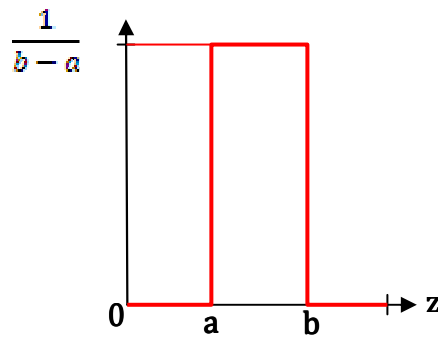
El ruido que afecta a la imagen sigue una distribución uniforme. La probabilidad de tomar cualquier valor de gris dentro de un intervalo definido es constante.

$$p(z) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq z \leq b \\ 0 & \text{en otro caso} \end{cases} \quad (\text{Ecuación 30})$$

Con valores de media y desviación típica:

$$\mu = \frac{a+b}{2} \quad (\text{Ecuación 31})$$

$$\sigma^2 = \frac{(b-a)^2}{12} \quad (\text{Ecuación 32})$$



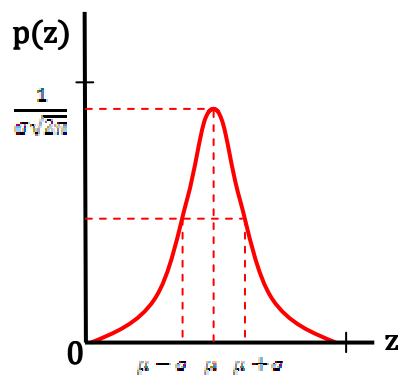
Gráfica 15

No es de los más frecuentes, pero se utiliza mucho en las simulaciones.

2. Ruido gaussiano

La probabilidad de que el ruido valga “z” es:

$$p(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \quad (\text{Ecuación 33})$$



Gráfica 16

Los parámetros que lo definen son la media (μ) y la desviación típica (σ).

Con la distribución anterior, aproximadamente el 70% de las veces el ruido se encuentra comprendido entre los valores de $(\mu-\sigma)$ y $(\mu+\sigma)$ y aproximadamente el 95% de las veces el ruido se encuentra comprendido entre los valores de $(\mu-2\sigma)$ y $(\mu+2\sigma)$.

Se caracteriza por tener un espectro de energía constante para todas las frecuencias.

El ruido gaussiano tiene un efecto general en toda la imagen, es decir, la intensidad de cada píxel de la imagen se ve alterada en cierta medida con respecto a la intensidad en la imagen original.

3. Ruido exponencial

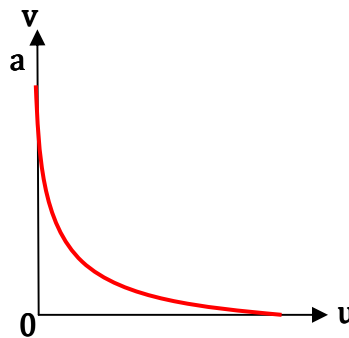
La probabilidad de que el ruido valga “z” es:

$$p(z) = \begin{cases} ae^{-az} & \text{para } z \geq a \\ 0 & \text{para } z < a \end{cases} \quad (\text{Ecuación 34})$$

Con valores de media y desviación típica que vienen dadas por las siguientes funciones:

$$\mu = \frac{1}{a} \quad (\text{Ecuación 35})$$

$$\sigma^2 = \frac{1}{a^2} \quad (\text{Ecuación 36})$$



Gráfica 17

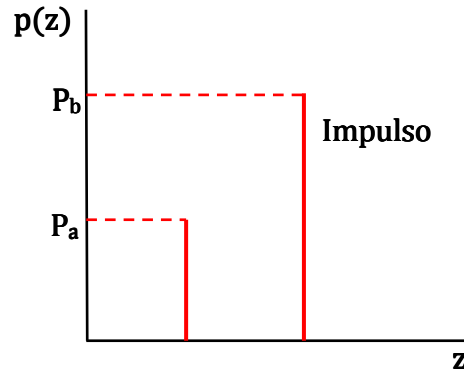
4. Ruido sal y pimienta

Es causado principalmente por fallos en el funcionamiento de los sensores encargados de capturar una imagen o por errores de tiempo cuando se produce el proceso de digitalización de la misma.

Al momento de digitalizar la imagen este tipo de ruido generalmente toma valores extremos en la imagen (cerca de los valores de negro o muy cercanos a los valores de blanco), esto es debido a que los impulsos de ruido pueden ser negativos o positivos. Generalmente suelen suponerse que los valores de a y b se encuentran saturados, ya sea en sus valores máximos o en sus valores mínimos cuando se digitaliza la imagen.

La probabilidad de que el ruido valga “z” es:

$$p(z) = \begin{cases} P_a & \text{para } z = a \\ P_b & \text{para } z = b \\ 0 & \text{en otro caso} \end{cases} \quad (\text{Ecuación 37})$$



Gráfica 18

Si $b > a$, la intensidad aparecerá como un punto blanco en la imagen y a aparecerá como un punto negro.

5. Ruido Rayleigh

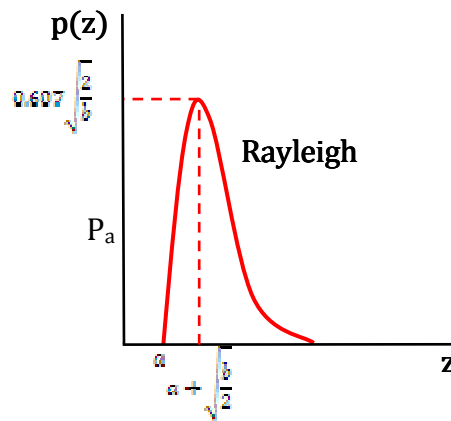
La probabilidad de que el ruido valga " z " es:

$$p(z) = \begin{cases} \frac{2}{b}(z-a)e^{-\frac{(z-a)^2}{b}} & \text{para } z \geq a \\ 0 & \text{para } z < a \end{cases} \quad (\text{Ecuación 38})$$

El valor de la media y la desviación típica viene dada por:

$$\mu = a + \sqrt{\frac{\pi b}{4}} \quad (\text{Ecuación 39})$$

$$\sigma^2 = \frac{b(4-\pi)}{4} \quad (\text{Ecuación 40})$$



Gráfica 19

Este ruido suele generarse en el proceso de obtención de las imágenes.

6. Ruido gamma

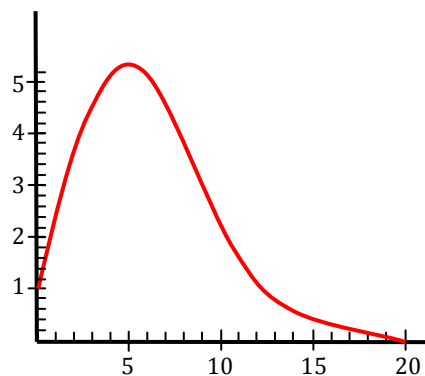
La probabilidad de que el ruido valga “z” es:

$$p(z) = \begin{cases} \frac{a^b z^{b-1} e^{-az}}{(b-1)!} & \text{para } z \geq a \\ 0 & \text{para } z < a \end{cases} \quad (\text{Ecuación 41})$$

Con valores de media y desviación típica definidos por:

$$\mu = \frac{b}{a} \quad (\text{Ecuación 42})$$

$$\sigma^2 = \frac{b}{a^2} \quad (\text{Ecuación 43})$$



Gráfica 20

7. Ruido periódico

El ruido periódico se suele originar debido a una interferencia eléctrica o electromagnética durante la adquisición de la imagen. Debido a sus características, el ruido periódico puede reducirse significativamente mediante filtrado en el dominio de la frecuencia, ya que el ruido periódico produce picos de frecuencia en el espectro de Fourier que son fácilmente identificables.

2.1.3.6 Operaciones geométricas

A diferencia de las operaciones aritméticas o la introducción de ruido, en las operaciones geométricas no se modifican los brillos de los píxeles que componen las imágenes, lo que sí se modifica es la posición donde se van a ubicar.

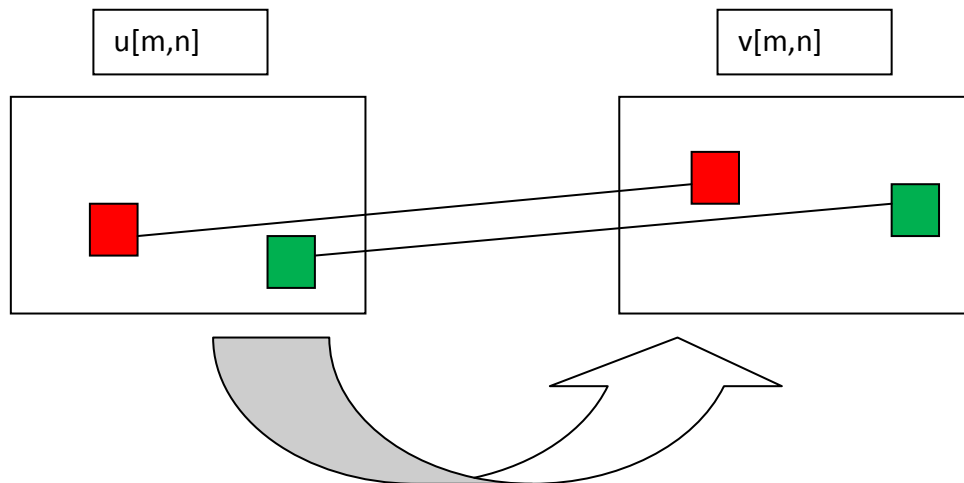


Ilustración 11. Operaciones geométricas

$$v[m',n'] = u[f_1(m,n), f_2(m,n)] \quad (\text{Ecuación 44})$$

El píxel de llegada se busca en algún punto de la imagen de entrada.

1. Simetría

Con la simetría, la imagen mantiene la colocación de los píxeles en un eje mientras los invierte en el otro. Si la inversión se realiza en el eje X la simetría será horizontal y si se hace en el eje Y, vertical.

- En el caso horizontal (Mirror)

Esta consiste en intercambiar las columnas de una imagen.

$$x' = x ; y' = Y_0 - y \quad (\text{Ecuación 45})$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & Y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (\text{Ecuación 46})$$

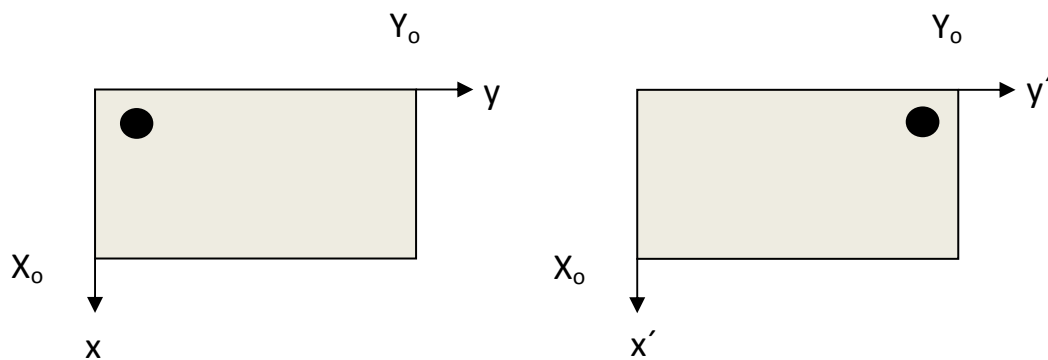


Ilustración 12

Tomando (x, y) como coordenadas en la imagen original e (x', y') como las coordenadas en la imagen simétrica.

- En el caso vertical (flip)

Esta consiste en intercambiar las filas de una imagen.

$$x' = X_0 - x; y' = y \quad (\text{Ecuación 47})$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & X_0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (\text{Ecuación 48})$$

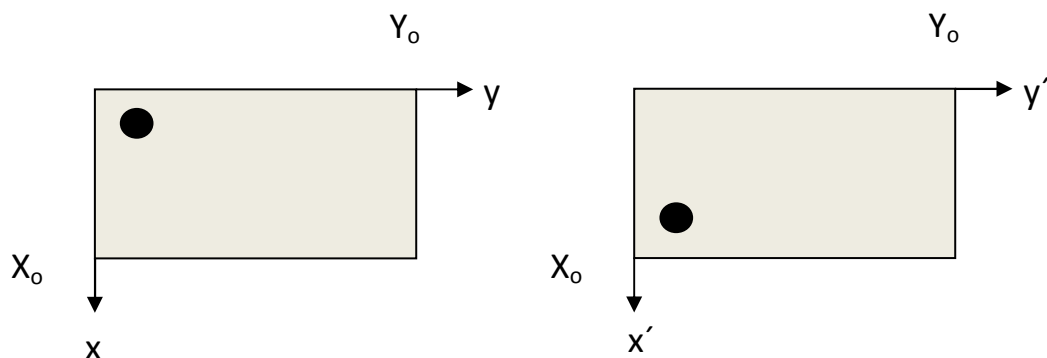


Ilustración 13

Análisis de imágenes basado en correlación de imágenes

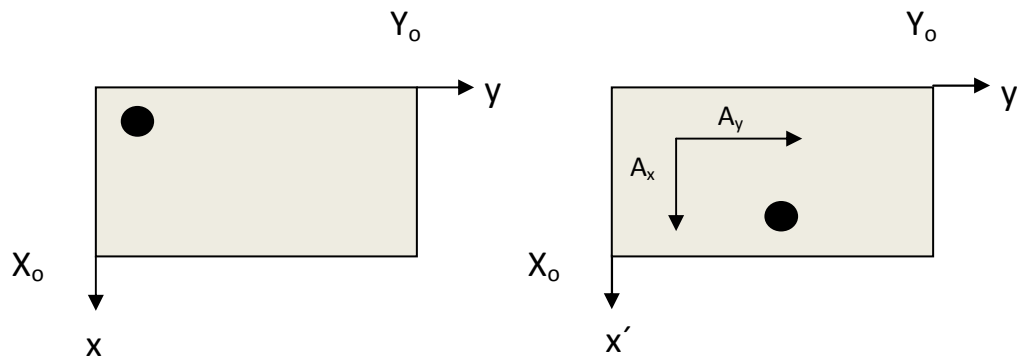
Tomando (x,y) como coordenadas en la imagen original e (x',y') como las coordenadas en la imagen simétrica.

2. Traslaciones

Es una operación geométrica donde una posición determinada de un píxel (x,y) , se desplaza a una nueva posición empleando unos desplazamientos $(\Delta x, \Delta y)$. Las ecuaciones que muestran esta operación son las siguientes:

$$x' = x + \Delta x; y' = y + \Delta y \quad (\text{Ecuación 49})$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta x \\ 0 & 1 & \Delta y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (\text{Ecuación 50})$$



3. Escalado

Es una operación geométrica que permite cambiar la escala de la imagen. Las ecuaciones que describen esta operación son las siguientes:

$$x' = k_1 x; y' = k_2 y \quad (\text{Ecuación 51})$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (\text{Ecuación 52})$$

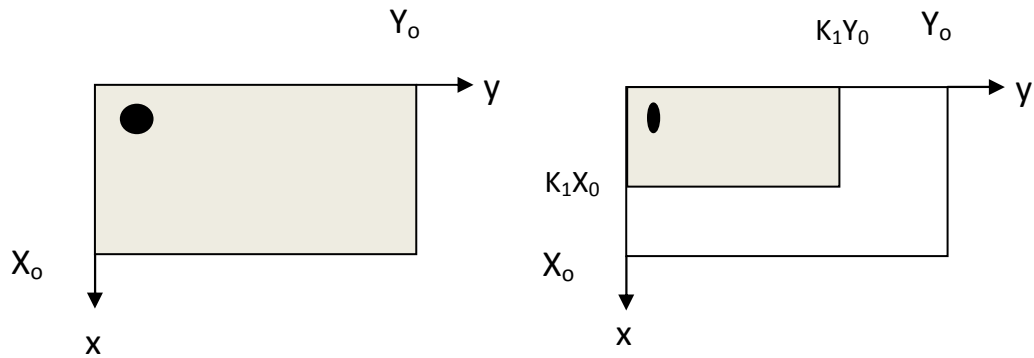


Ilustración 15

4. Rotación.

Consiste en girar la imagen un ángulo definido. Un detalle a tener en cuenta es que la imagen puede cambiar de tamaño respecto a la original con este proceso.

Deducimos la relación entre posiciones que es:

$$\begin{aligned} y'_0 &= y_0 \cos \alpha + x_0 \sin \alpha \\ x'_0 &= x_0 \cos \alpha + y_0 \sin \alpha \end{aligned} \quad (\text{Ecuación 53})$$

Y de estas ecuaciones obtenemos la matriz de rotación:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (\text{Ecuación 54})$$

Si los nuevos ejes giran en sentido horario, la imagen gira en sentido anti horario y viceversa. Para girar la imagen en sentido horario α será negativo.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (\text{Ecuación 55})$$

- Rotación simple $\pm 90^\circ$

Al rotar cualquier imagen 90° , se produce un reacomodo de píxeles. Por este motivo si por ejemplo las dimensiones de la imagen original eran N y M, al efectuar la rotación, en la dirección contraria a las manecillas del reloj respecto al centro de la imagen, se producirá una transposición y el número de columnas será M y el de filas N.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad \alpha = +90^\circ \quad \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

(Ecuación 56)

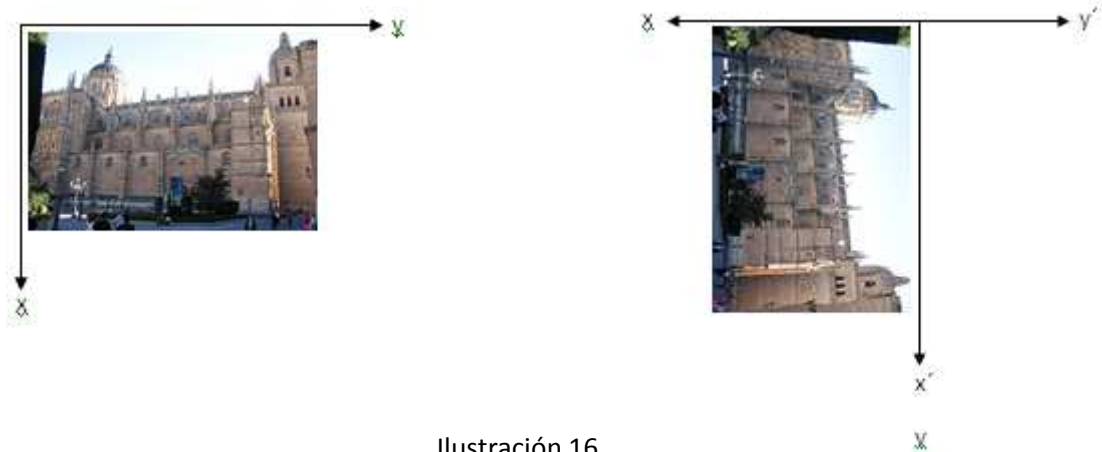


Ilustración 16

- Rotación 180°

Al rotar una imagen 180°, obtenemos una imagen con las mismas dimensiones que la original pero en sentido contrario.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad \alpha = +180^\circ \quad \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

(Ecuación 57)

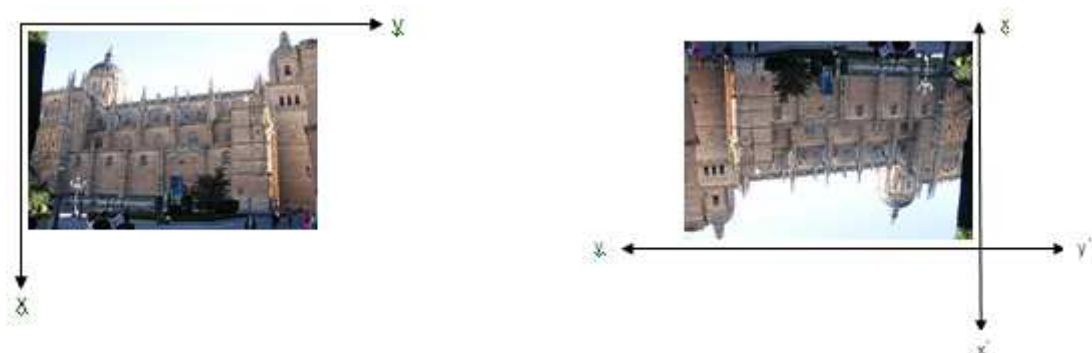


Ilustración 17

- Rotación libre directa

Si queremos rotar una imagen un ángulo Θ utilizamos la misma ecuación descrita anteriormente.

- Rotación libre inversa

Para evitar tener píxeles sin valor asignado, es posible inventarse el valor del píxel, en vez de la posición. En resumen las posiciones son exactas, y los brillos interpolados.

5. Inclinación

La inclinación transforma una región rectangular en un romboide. Nos sirve para simular una perspectiva.

2.1.3.7 Convolución

1. Unidimensional

La convolución de dos secuencias $x[n]$ e $y[n]$, es otra secuencia $z[n]$ dada por la expresión:

$$x[n] = x[n] * y[n] = \sum_{n'=-\infty}^{\infty} x[n'] * y[n-n'] \quad -\infty < n < +\infty \quad (\text{Ecuación 58})$$

Si el tamaño de $x[n]$ es N_1 y el de $y[n]$ es N_2 , el tamaño de $z[n]$ es $N_1 + N_2 - 1$

$$x[n] = x[n] * y[n] = y[n] * x[n] \quad (\text{Ecuación 59})$$

2. Bidimensional

La convolución de dos secuencias bidimensionales $x[m,n]$ e $y[m,n]$ es otra secuencia bidimensional $z[m,n]$ dada por la expresión:

$$z[m,n] = x[m,n] * y[m,n] = \sum_{m'=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} x[m',n'] y[(m-m'),(n-n')] \\ -\infty < n < +\infty \\ -\infty < m < +\infty$$

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{j2\pi \frac{k}{N}n}, k = 0, 1, 2, \dots, (N-1)$$

(Ecuación 60)

Si el tamaño de $x[m,n]$ es $[M_1, N_1]$ y el de $y[m,n]$ es $[M_2, N_2]$, el tamaño de $z[m,n]$ es $[(M_1+M_2-1), (N_1+N_2-1)]$

2.1.3.8 Transformadas de la imagen

1. Transformada de Fourier unidimensional

La transformada de una secuencia $x[n]$ es:

$$X(\Omega) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\Omega n} \quad (\text{Ecuación 61})$$

Donde $\Omega = 2\pi F/N$, siendo N el número de valores de la secuencia $x[n]$ para la variable Ω de periodo N para la variable F .

La transformada inversa será:

$$x[n] = \frac{1}{2\pi} \int_{2\pi} X(\Omega) e^{j\Omega n} d\Omega \quad (\text{Ecuación 62})$$

2. Transformada de Fourier bidimensional

La transformada de Fourier de una imagen $x[m,n]$ es una función continua:

$$X(\Omega_1, \Omega_2) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x[m,n] e^{-j\Omega_1 m} e^{-j\Omega_2 n} \quad (\text{Ecuación 63})$$

Siendo $\Omega_1 = 2\pi F_1/N$ y $\Omega_2 = 2\pi F_2/N$.

La transformada inversa será:

$$x(K) = \frac{1}{(2\pi)^2} \int_{2\pi} \int_{2\pi} X(\Omega_1, \Omega_2) e^{j\Omega_1 k} e^{j\Omega_2 n} d\Omega_1 d\Omega_2 \quad (\text{Ecuación 64})$$

3. Transformada discreta de Fourier unidimensional

La transformada de Fourier de una secuencia discreta genera una función continua.
La transformada discreta de Fourier genera una secuencia:

$$X[K] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi \frac{K}{N} n}, k = 0, 1, 2, \dots, (N-1) \quad (\text{Ecuación 65})$$

La transformada inversa será:

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{j2\pi \frac{k}{N}n}, n = 0, 1, 2, \dots \quad (\text{Ecuación 66})$$

4. Transformada discreta de Fourier Bidimensional

Para la secuencia bidimensional, $x[m,n]$ su transformada discreta de Fourier es:

$$X[k,l] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x[m,n] e^{(-j2\pi \frac{k}{M}M)} e^{(-j2\pi \frac{l}{N}N)} \quad (\text{Ecuación 67})$$
$$0 < k < (M-1), 0 < l < (N-1)$$

Y la transformada inversa será:

$$x[m,n] = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} X[k,l] e^{-j2\pi \frac{k}{M}m} e^{-j2\pi \frac{l}{N}n} \quad (\text{Ecuación 68})$$

5. Propiedades de la transformada de Fourier

- Linealidad

El espectro de la suma lineal de imágenes es igual a la suma lineal de los espectros.

Tenemos que:

$$TF\{x_1[m,n]\} = X_1[k,l] \quad (\text{Ecuación 69})$$

$$TF\{x_2[m,n]\} = X_2[k,l] \quad (\text{Ecuación 70})$$

Si:

$$x[m,n] = ax_1[m,n] + bx_2[m,n] \quad (\text{Ecuación 71})$$

Se cumple que:

$$X[k,l] = aX_1[k,l] + bX_2[k,l] \quad (\text{Ecuación 72})$$

- Translación

Teniendo que:

$$TF \{x_1[m, n]\} = X_1[k, l] \quad (\text{Ecuación 73})$$

Si se desplaza x_1 :

$$x_2[m, n] = x_1[(m - m_0), (n - n_0)] \quad (\text{Ecuación 74})$$

Se cumple que:

$$X_2[k, l] = X_1[k, l] e^{j2\pi m_0 k} e^{j2\pi n_0 l} \quad (\text{Ecuación 75})$$

- Separabilidad

Mediante esta propiedad podemos calcular la transformada discreta de Fourier de una función bidimensional como una multiplicación de dos transformadas de Fourier discretas, calculando primero una TFD sobre la variable de uno de los ejes y al resultado multiplicarle por la TFD de la segunda imagen.

$$T(x) = T(x_1) * T(x_2) \quad (\text{Ecuación 76})$$

- Periodicidad

La transformada de una secuencia $x[m, n]$ de tamaño $M \times N$ es una función continua y periódica. La transformada discreta de Fourier son muestras de la anterior, pero solo se considera un periodo. El resto es repetido.

- Simetría conjugada

La transformada de Fourier cumple:

$$X[k, l] = X^*[-k, -l] X(\Omega) = \sum_{-\infty}^{\infty} x[n] e^{-j\Omega n} \quad (\text{Ecuación 77})$$

- Centrado de la transformada

Los valores significativos suelen estar en los extremos. Al multiplicar la imagen por $(-1)^{m+n}$ antes de la transformación, implica el centrado de la transformada:

Si:

$$TF\{f[m,n]\} = F(k,l) \quad (\text{Ecuación 78})$$

$$TF\{f[m,n](-1)^{m+n}\} = F\left(k - \frac{M}{2}, l - \frac{N}{2}\right) \quad (\text{Ecuación 79})$$

- Rotación

Si rotamos una imagen su espectro también rota.

- Convolución

Si tenemos:

$$TF\{x_1[m,n]\} = X_1[k,l] \quad (\text{Ecuación 80})$$

$$TF\{x_2[m,n]\} = X_2[k,l] \quad (\text{Ecuación 81})$$

Entonces:

$$x_1[m,n] * x_2[m,n] = X_1[k,l] X_2[k,l] \quad (\text{Ecuación 82})$$

- Producto

Si tenemos:

$$TF\{x_1[m,n]\} = X_1[k,l] \quad (\text{Ecuación 83})$$

$$TF\{x_2[m,n]\} = X_2[k,l] \quad (\text{Ecuación 84})$$

Entonces:

$$x_1[m,n] \cdot x_2[m,n] = X_1[k,l] * X_2[k,l] \quad (\text{Ecuación 85})$$

Esta convolución ha de ser circular porque el espectro es periódico no limitado.

- Conservación de la energía

$$\sum \sum |x|^2 = \sum \sum |X|^2 \quad (\text{Ecuación 86})$$

$$\sum \sum xx^* = \sum \sum XX^* \quad (\text{Ecuación 87})$$

- Escalado en campo continuo

Si:

$$TF\{f(x, y)\} \text{ es } F(w_x, w_y) \quad (\text{Ecuación 88})$$

Entonces:

$$TF\{f(x, y)\} = \frac{1}{|ab|} F\left(\frac{w_x}{a}, \frac{w_y}{b}\right) \quad (\text{Ecuación 89})$$

- Escalado en campo discreto

$$x_k \rightarrow X[kn] \quad (\text{Ecuación 90})$$

2.1.3.9 Operaciones área de filtros

En las operaciones de punto, cada píxel de salida depende sólo de un píxel de entrada, es decir, no teníamos en cuenta la relación de vecindad entre píxeles.

En las operaciones de área (filtrado espacial), el valor de un píxel depende de la vecindad local de ese píxel.

Matemáticamente podemos verlo en las siguientes funciones:

- En operaciones de punto:

$$v[m, n] = f(u[m, n]) \text{ ó } v[m, n] = f(u_1[m, n], u_2[m, n]) \quad (\text{Ecuación 91})$$

- En operaciones de área:

$$v[m, n] = f(u[m-k, n-k], \dots, u[m, n], \dots, u[m+k, n+k]) \quad (\text{Ecuación 92})$$

Análisis de imágenes basado en correlación de imágenes

Un tipo interesante de operaciones de área son las convoluciones discretas, que son transformaciones en las que el valor del píxel resultante es una combinación lineal de los valores de los píxeles vecinos en la imagen. La matriz de los coeficientes de esta combinación lineal la denominamos máscara o núcleo de convolución.

Matemáticamente lo podríamos reflejar así:

- Sea a una máscara de convolución.
- Imagen de entrada u .
- Imagen de salida v .
- Algoritmo

Calcular la convolución $v=a \circledast u$ para cada píxel $[m,n]$ de la imagen u haciendo:

$$v[m,n] = \sum \sum a[k,l] u[m-k, n-l] \quad (\text{Ecuación 93})$$

Uno de los dilemas que nos surgen a la hora de convolucionar una imagen con una máscara es qué hacer con los bordes. Esta situación puede ser resuelta de estos cuatro modos diferentes:

1. Asignar un cero en el resultado a los píxeles donde no cabe la máscara.
2. Suponer que los píxeles que se salen tienen un valor constante (normalmente cero).
3. Modificar el valor en los píxeles que no caben (variar el multiplicador)
4. Suponer que la imagen se pliega por los extremos.

Según apliquemos distintos operadores de convolución es posible obtener diferentes efectos en nuestra imagen:

- Suavizado o difuminado de la imagen, reducir contrastes grandes en la imagen.
- Perfilado: resaltar los contrastes. Es la operación opuesta al suavizado.
- Bordes: detectar zonas de variación de la imagen.
- Detección de cierto tipo de características, tales como esquinas, segmentos...

Así como el uso de las dos primeras es más habitual en restauración y mejora de imágenes las dos últimas suelen usarse más en análisis de imágenes.

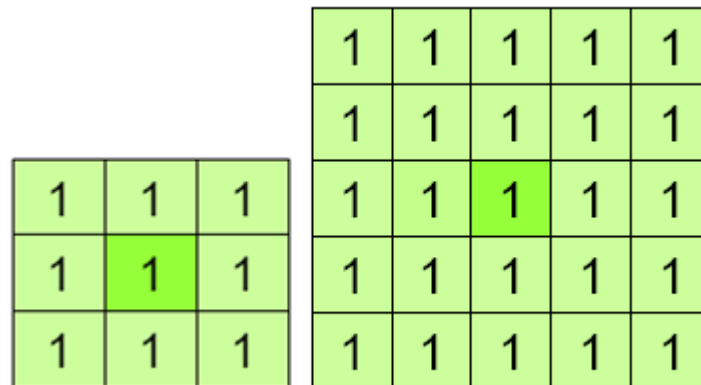
1. Filtros suavizantes lineales

Cumplen una doble acción ya que difuminan la imagen (comportamiento filtro paso bajo) y a su vez disminuyen el ruido en imágenes ruidosas.

Análisis de imágenes basado en correlación de imágenes

Las máscaras de estos filtros se caracterizan por tener todos sus coeficientes positivos, no como las máscaras de los filtros paso alto que presentan tanto coeficientes positivos como negativos.

- El filtro de media (aritmética) es el operador de suavizado más simple, este considera la media de los píxeles en un entorno $w \times h$ centrado en un píxel (x,y) .



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

Ilustración 18

La descripción práctica de este proceso sería:

1. Superponemos la máscara sobre cada punto de la imagen.
2. Sumamos los productos de los coeficientes de la máscara y los valores de los puntos de la imagen situados bajo esta.

Según lo dicho hasta ahora podemos entender que el efecto de este filtro será de suavizado, ya que los píxeles vecinos se parecerán debido a la mezcla que se produce entre ellos. Al aplicar este filtro podremos observar como la imagen se hará borrosa y las transiciones fuertes se disolverán parcialmente.

También debemos decir que cuanto mayor es el tamaño de la máscara, mayor es el efecto de difuminado de la imagen.

Las ventajas son que reduce el número de operaciones, además respecto al ruido, la potencia de este se reduce en un factor igual al número de píxeles que abarca la máscara, así que cuanto mayor sea el tamaño de la máscara, mejor será la relación señal-ruido.

- Suavizado direccional

Como el filtro de media distorsiona la información de los bordes, como solución para reducir el empañamiento se puede usar un filtro de promediado direccional, que consiste en calcular las medias de regiones longitudinales a lo

Análisis de imágenes basado en correlación de imágenes

largo de todas las direcciones posibles. De todas ellas, elegimos las más parecidas al valor del punto en la imagen de entrada.

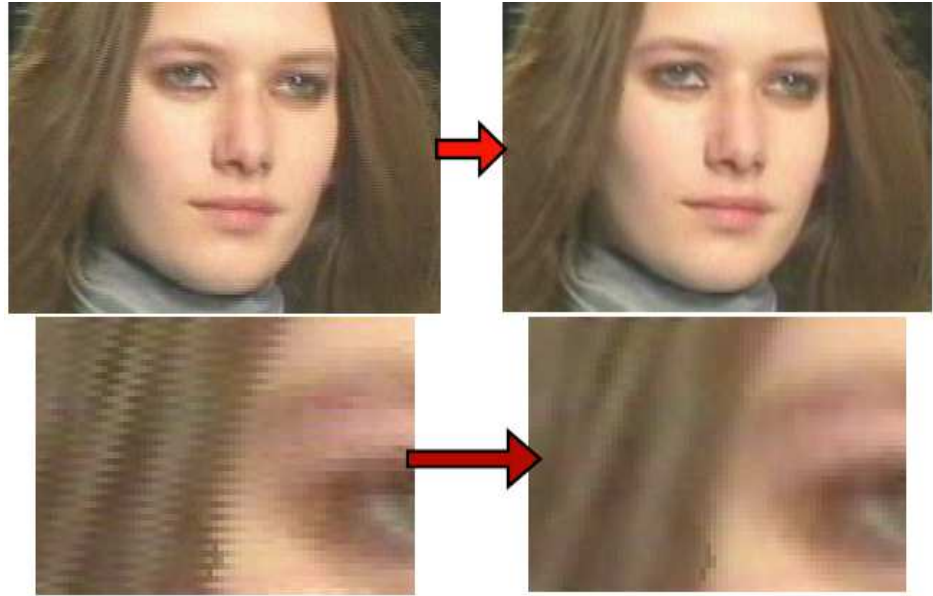


Ilustración 19. Suavizado direccional

- Suavizado gaussiano

Consiste en aplicar un suavizado gaussiano donde se hace la media ponderada, donde los pesos toman la forma de una campana de Gauss.

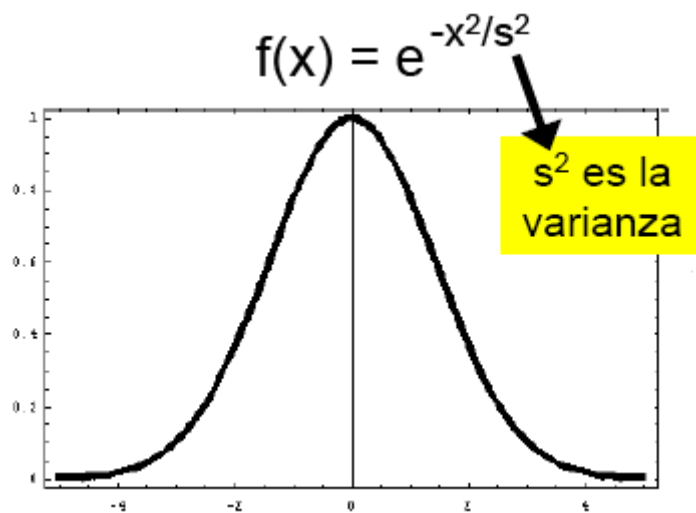


Ilustración 20. Suavizado gaussiano

La varianza, s^2 , indica el nivel de suavizado. Si la varianza es grande la campana será más ancha, por tanto el suavizado será mayor. Si por el contrario la varianza es pequeña la campana será más estrecha y tendremos menos suavizado.

Campana de gauss discreta

Para el cálculo de la máscara gaussiana (1D), usemos las filas del triángulo de Pascal de la forma siguiente:

1/2·	1	1					
1/4·	1	2	1				
1/8·	1	3	3	1			
1/16·	1	4	6	4	1		
1/32·	1	5	10	10	5	1	
1/64·	1	6	15	20	15	6	1

Ilustración 21. Campana de Gauss

Normalmente, el suavizado gaussiano se aplica en dos dimensiones. Los pesos de la máscara dependen de la distancia al píxel central.

Campana de Gauss 2D

$$f(x, y) = e^{-(x^2 + y^2)/x^2} \quad (\text{Ecuación 94})$$

1/16·	1	2	1
	2	4	2
	1	2	1

Ilustración 22

Esto se puede simplificar teniendo en cuenta que el filtro gaussiano es separable, y podemos obtener un suavizado 2D aplicando dos máscaras gaussianas bidimensionales, una horizontal y otra vertical.

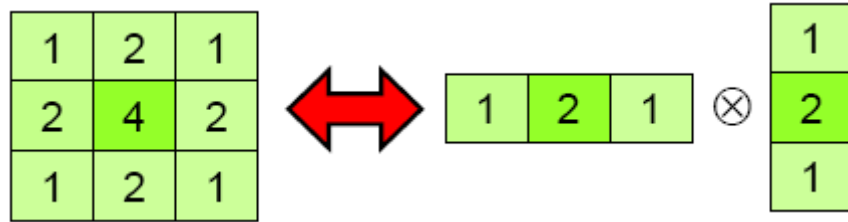


Ilustración 23

En comparación con el filtro de media para conseguir un mismo grado de suavizado la mascara gaussiana debe ser de mayor tamaño. A favor del efecto del suavizado gaussiano podemos decir que su efecto es más natural que el del filtro de media, y por esto suele ser más habitual en procesamiento y análisis de imágenes.

2. Filtros suavizantes no lineales

En las convoluciones, f es una combinación lineal cualquiera. Pero también puede ser interesante usar otras funciones no lineales.

- Media geométrica

$$v[m, n] = \sqrt[4]{u[m-1, n-1] \cdot u[m-1, n] \cdot u[m, n-1] \cdot u[m, n]} \quad (\text{Ecuación 95})$$



Ilustración 24. Media geométrica

Su efecto es similar a la media geométrica. Sin embargo todas las transformaciones no lineales no son útiles e interesantes. Las que más se utilizan son la de máximo, mínimo y media.

- Filtro de máximo

$$v[m, n] = \max\{u[m-k, n-k], \dots, u[m, n], \dots, u[m+k, n+k]\} \quad (\text{Ecuación 96})$$

Donde k es el radio y el tamaño es $2k+1$

El resultado de aplicar este filtro es un cierto efecto de difuminado y aclarado de la imagen. Desaparecen los detalles más oscuros.

Si el tamaño es muy grande pueden ocurrir dos efectos. En primer lugar puede producirse un efecto de cuadriculado, como el máximo se aplica en una zona cuadrada, los píxeles muy claros generan un cuadrado uniforme alrededor. Y en segundo lugar, pueden aparecer colores falsos, al aplicar el filtro puede aparecer colores que no existen en la imagen original.

- Filtro de mínimo

$$v[m, n] = \min\{u[m-k, n-k], \dots, u[m, n], \dots, u[m+k, n+k]\} \quad (\text{Ecuación 97})$$

Donde k es el radio y el tamaño es $2k+1$.

Su efecto es parecido al máximo, pero tomando los valores menores, es decir, los más oscuros.

- Filtro de mediana

Esta es una técnica alternativa cuando el objetivo que queremos alcanzar es más la reducción del ruido que el difuminado de la imagen.

Consiste en sustituir el valor del nivel de gris de cada punto de la imagen de entrada por el valor mediano de los puntos que están incluidos dentro de una ventana entorno.

$$v[m, n] = \text{mediana}\{u[m-k, n-l]\}, \text{ para } (k, l) \in W \quad (\text{Ecuación 98})$$

Sobre este tipo de filtro debemos tener en cuenta que:

1. Este filtro no es lineal:

$$\text{mediana}\{ax(m, n) + by(m, n)\} \neq a \cdot \text{mediana}\{x(m, n)\} + b \cdot \text{mediana}\{y(m, n)\} \quad (\text{Ecuación 99})$$

2. La supresión del ruido la realiza bien si es binario (compuesta de fuertes componentes de pico).
3. Los resultados son pobres si el ruido tiene distribución Gaussiana, o si el número de píxeles de la ventana afectados por el ruido es más de la mitad.

La mediana produce un efecto suavizado, aunque más “abrupto” en los bordes que la media y el suavizado gaussiano. Pero el verdadero interés es la eliminación de ruido potencial.

3. Filtros agudizadores

Los filtros agudizadores como su propio nombre indica provocan el efecto de agudizar los detalles finos (altas frecuencias) de la imagen. El objetivo de este realce es destacar los detalles finos de una imagen o intensificar detalles que han sido difuminados, bien sea por error o bien por efecto natural de modo de adquisición de la imagen. Son el efecto contrario a los filtros de suavizado.

Estos filtros destacan las variaciones en la imagen, bordes. Un borde es la frontera entre dos regiones con propiedades de nivel de gris relativamente distintas.

En el apartado detección de bordes del tema análisis de imágenes, hablaremos más concretamente de los operadores utilizados para la detección de estos.

3. Análisis de imágenes

Se define como el estudio de las características de la imagen, con el fin de extraer cualquier tipo de información adicional contenida en la imagen y que a simple vista es poco evidente. Esta información puede ser de muy diversos tipos tales como medidas sobre la imagen, reconocimiento de formas, etc.

El análisis de imágenes abarca toda esa serie de procesos que permiten extraer información de una imagen. Estos procesos los podemos resumir en los siguientes:

1. Extracción de características globales de la imagen

- Detección de bordes, líneas y esquinas
- Texturas.
- Detección del movimiento.

2. Segmentación de imágenes.

- Basada en Umbralización.
- Basada en extracción de contornos.
- Orientada a regiones.

3. Transformaciones morfológicas.

4. Representación de contornos y regiones mediante información extraída de la imagen segmentada.

3.1 Extracción de características de una imagen

Atendiendo al medio de extracción se van a clasificar en dos grandes grupos:

- Características espaciales de la imagen.
- Características de transformadas.

3.1.1 Características especiales de una imagen

Las características espaciales de un objeto pueden extraerse a partir de los niveles de gris que presenta el objeto en la imagen (características de amplitud) o por medio de la distribución espacial de píxeles de dicho objeto, en función del resto de los de la imagen (características del Histograma).

3.1.1.1 Características de amplitud

Caracterizamos el objeto mediante sus características físicas. Por ejemplo en una imagen obtenida con rayos X, la amplitud de los niveles de gris de ésta, representa la absorción de las diferentes partes del cuerpo, lo que permite distinguir los tejidos, de las partes óseas, o bien distinguir zonas enfermas de los tejidos, de zonas que no lo están.

Análisis de imágenes basado en correlación de imágenes

Un objeto caracterizado por su amplitud puede ser fácilmente extraído de la imagen, por ejemplo, determinando un umbral del nivel de gris, a partir del cual discernir el objeto del fondo y así poder separarlos.

3.1.1.2 Características del histograma

Se basan en la observación del histograma de una determinada región de la imagen. El histograma como hemos comentado anteriormente representa el número de píxeles que poseen un determinado nivel de gris.

Para justificar la obtención de estas características, se va a dividir cada valor del histograma por el número total de píxeles, obteniendo así el histograma normalizado, que representa la probabilidad de que un determinado nivel de gris aparezca en la imagen. Se define esta probabilidad como:

$$\rho_u(x) = \text{Prob}[u = x] = \frac{\text{numero de píxels con nivel de gris } x}{\text{numero de píxels total de la región}} \quad (\text{Ecuación 100})$$
$$x = 0, 1, \dots, L-1$$

El valor de $L-1$ representa el nivel de gris máximo posible y viene dado por el máximo valor de cuantificación asignado a la escala de grises.

A partir del histograma obtenemos características de una variable aleatoria, como la dispersión, la media, la varianza etc. Otras características importantes son la mediana y la moda.

El proceso sería el siguiente: Se elige un tamaño de ventana para la aplicación del cálculo y dependiendo del efecto deseado, se sustituye el valor de cada píxel por el valor de la media, moda o varianza de la ventana centrada en él, obteniendo así otra imagen con el efecto deseado o el valor de la característica extraída.

El histograma es utilizado, en la umbralización de imágenes, para extraer el valor umbral y así binarizar la imagen para su segmentación.

3.1.2 Características de la transformada

Son aquellas características obtenidas a partir de la transformada de una determinada imagen. En muchas ocasiones una transformación a un dominio concreto puede proporcionar gran información acerca de la imagen.

Como la imagen es una función discreta, para obtener estas características se utiliza la transformada discreta de Fourier descrita en apartados anteriores. A partir de ella se pueden calcular las respuestas en frecuencia que presentan determinadas formas. Generalmente las altas frecuencias se emplean para detectar contornos, y las líneas aisladas que aparecen en la imagen transformada aportan información sobre la orientación del objeto en la imagen.

Análisis de imágenes basado en correlación de imágenes

La utilidad de estas transformadas es identificar objetos y orientaciones en una imagen desconocida, conociendo las transformadas de Fourier de ciertas figuras como cuadrados, rectángulos, etc.

3.1.3 Detección de bordes, texturas y movimientos

Los bordes constituyen una de las características más importantes que se pueden encontrar en una imagen ya que delimitan los objetos que en ella se encuentran. Los bordes se definen como cambios locales significativos de luminancia en la imagen, que generalmente están asociados a una discontinuidad. Debido a su localización pueden dividirse en:

- Bordes con transiciones abruptas, en los que hay grandes cambios de luminancia entre píxeles adyacentes.
- Bordes con transiciones progresivas, en los que los cambios de luminancia son progresivos y la variación entre píxeles adyacentes no es grande.

También se pueden clasificar según su discontinuidad en:

- Bordes de salto, donde la luminancia cambia de valor a un lado y a otro de la discontinuidad.
- Bordes de línea, donde la luminancia cambia volviendo en poco espacio a su valor inicial

Las técnicas empleadas en la detección de bordes tienen como objetivo localizar en la imagen los píxeles en los que se produce una variación de luminancia, empleando operadores derivativos. Estos son básicamente dos: Operadores de primera derivada o Gradiente y Operadores de segunda derivada o Laplacianos. Los primeros buscan grandes picos en la variación mientras que los segundos, buscan pasos de respuesta positiva a negativa y viceversa o sea el paso por cero.

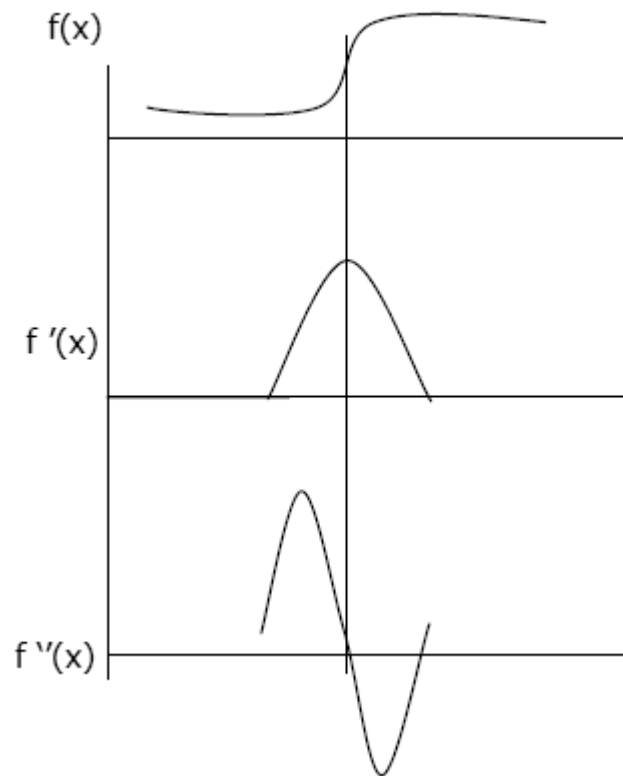


Ilustración 25. Primera y segunda derivada unidimensional de f.

3.1.3.1 Operadores tipo gradiente

Es esencialmente una operación de detección de cambios locales significativos de luminancia. Estos cambios están asociados con un pico local de la primera derivada (ver figura anterior). El equivalente en una imagen bidimensional es el gradiente. Localizando los puntos donde el gradiente es elevado, se sabrá donde se han producido las máximas variaciones de luminancia, por tanto donde estarán los bordes. Se define el operador gradiente aplicado a una imagen como:

$$\nabla f(x, y) = [G_x, G_y] = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right] \quad (\text{Ecuación 101})$$

Indica que el gradiente queda determinado conociendo las derivadas parciales de la direcciones ortogonales x e y. Su módulo y dirección vendrán dado por:

$$|\nabla f| = \sqrt{G_x^2 + G_y^2} \quad (\text{Ecuación 102})$$

$$\angle \nabla f = \tan^{-1} \left(\frac{G_y}{G_x} \right) \quad (\text{Ecuación 103})$$

La dirección del gradiente es perpendicular al borde. Para evitar el cálculo computacional, el módulo se suele aproximar con:

$$|\nabla f| = |G_x| + |G_y| \quad (\text{Ecuación 104})$$

Como la imagen es una función discreta, se aproxima el concepto de derivada a espacios discretos. Esta aproximación no es más que la diferencia finita entre píxeles vecinos. Por ello la expresión del gradiente será:

$$\nabla f(x, y) = [G_x, G_y] = \left[\frac{\Delta f}{\Delta x}, \frac{\Delta f}{\Delta y} \right] \quad (\text{Ecuación 105})$$

Esto mismo se puede expresar a partir de sus máscaras de convolución, una que realiza la derivada en la dirección x, y la otra en la dirección y:

$$H_x = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \quad H_y = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$$

Ilustración 26. Máscaras de convolución

Estas máscaras son muy sensibles al ruido por lo que se han desarrollado otros operadores con máscaras de mayor orden que veremos a continuación.

El proceso del cálculo de la imagen gradiente se resume en la siguiente figura:

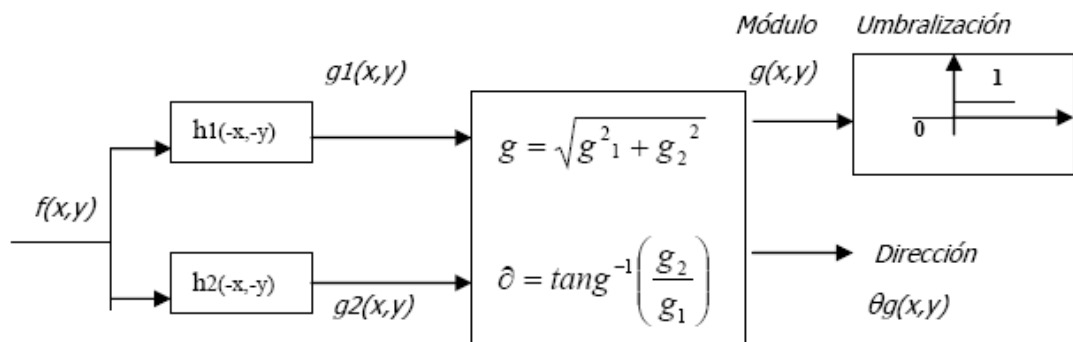


Ilustración 27. Proceso cálculo de la imagen gradiente

El resultado obtenido es una imagen $g(x,y)$ cuyos puntos valen el módulo del gradiente de la imagen original. Dada esta imagen se suele considerar que un píxel forma parte de un borde si su valor está entre el 10% de los valores más altos. El valor umbral a partir del cual un píxel se considera borde es calculado observando el histograma de la imagen gradiente. El resultado es una imagen binaria $I(x,y)$, que representará los bordes que aparecen en la imagen original y que se puede expresar:

$$I(x, y) = \begin{cases} 1 & \text{si } g(x,y) > \text{umbral} \\ 0 & \text{en otro caso} \end{cases} \quad (\text{Ecuación 106})$$

3.1.3.1.1 Operador de Roberts

Es uno de los más antiguos. Al usar máscaras de tamaño 2x2 utiliza pocos píxeles para aproximar el gradiente haciéndolo muy sensible al ruido.

$$H_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad H_y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Ilustración 28. Operadores Roberts

3.1.3.1.2 Operador de Prewitt

Este operador utiliza máscaras de 3x3:

$$H_x = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \quad H_y = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Ilustración 29. Operadores de Prewitt

3.1.3.1.3 Operador de Sobel

Igual que el operador de Prewitt, con la excepción que le da más peso a los píxeles más cercanos al centro de las máscaras. Las máscaras serán las siguientes:

$$H_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad H_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

Ilustración 30. Operador de Sobel

3.1.3.1.4 Operador Isotrópico (de Frei-Chen)

Se diferencia del anterior en los valores cercanos al centro donde aparece un valor intermedio. Véanse las máscaras:

$$H_x = \begin{pmatrix} -1 & 0 & 1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ -1 & 0 & 1 \end{pmatrix} \quad H_y = \begin{pmatrix} -1 & -\sqrt{2} & -1 \\ 0 & 0 & 0 \\ 1 & \sqrt{2} & 1 \end{pmatrix}$$

Ilustración 31. Operador de Frei-Chen

3.1.3.1.5 Operadores tipo compás

Son operadores que miden el gradiente en cualquier dirección (en intervalos de 45°), y la única ventaja que presentan respecto a los otros es que para calcular el gradiente sólo se necesita usar una matriz de una determinada dirección en vez de dos como en los operadores gradiente generales. Los operadores compás para una plantilla 3x3 son:

$\uparrow (N)$	$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$	$\Leftarrow \uparrow (NO)$	$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix}$
$\Leftarrow (O)$	$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$	$\Leftarrow \Downarrow (SO)$	$\begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}$
$\Downarrow (S)$	$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$	$\Downarrow \Rightarrow (SE)$	$\begin{bmatrix} -1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$
$\Rightarrow (E)$	$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$	$\uparrow \Rightarrow (NE)$	$\begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$

Ilustración 32. Operadores tipo Compás

Estos operadores permiten calcular la variación de luminancia de los píxeles de una imagen en cualquier dirección, con lo que, para determinados casos son más efectivos que los cálculos convencionales, pues evitan el tener que calcular la dirección del gradiente y su módulo a partir de las componentes ortogonales calculadas. Su desventaja es que para un punto hay que calcular el gradiente en todas las direcciones.

La forma de manejar estos operadores es la siguiente: Se calculan los gradientes en las ocho posibles direcciones, y se toma como valor del gradiente el valor del máximo gradiente direccional. Matemáticamente se puede expresar de la siguiente forma:

Se llama $g_k(x,y)$ al gradiente compás en la dirección $\Theta = \pi/2 + k\pi/4$, donde $k=0,...,7$. El gradiente en la posición (x,y) es:

$$g(x, y) = \max \{ g_k(x, y) \} \quad (\text{Ecuación 107})$$

3.1.3.2 Operadores tipo laplaciano

El laplaciano de una función bidimensional $f(x,y)$, es el resultado de aplicar un operador derivativo de segundo orden y se puede escribir así:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (\text{Ecuación 108})$$

Como se hizo con el gradiente, hay que hacer una aproximación discreta de la segunda derivada, utilizando las ecuaciones en diferencias centradas en la posición (i,j) :

$$\frac{\partial^2 f}{\partial x^2} = f(i+1, j) - 2f(i, j) + f(i-1, j) \quad (\text{Ecuación 109})$$

$$\frac{\partial^2 f}{\partial y^2} = f(i, j+2) - 2f(i, j+1) + f(i, j) \quad (\text{Ecuación 110})$$

Combinando estas dos ecuaciones en un único operador se obtendría la máscara 1 de la siguiente figura:

1) $\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	2) $\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	3) $\begin{bmatrix} -1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$
--	--	---

Ilustración 33. Operador de tipo Laplaciano

Se observa en la figura anterior, que el valor central toma el valor positivo de la suma de todos los píxeles que le rodean, por lo que la suma de los coeficientes de la máscara es cero, luego si se aplica a una región de píxeles constante, el resultado será nulo.

Detecta un borde cuando la salida del operador realiza una transición por cero. Es un operador muy sensible al ruido por lo que se utiliza como paso intermedio para el operador Marr-Hildred.

3.1.3.3 Operadores gradientes estocásticos

Al igual que los operadores laplacianos y de gradiente, son operadores que se basan en el diseño de máscaras de extracción de bordes, teniendo en cuenta la presencia de

Análisis de imágenes basado en correlación de imágenes

ruido. Para entenderlo mejor, se va a considerar un modelo de borde cuya transición se produce en un solo píxel.

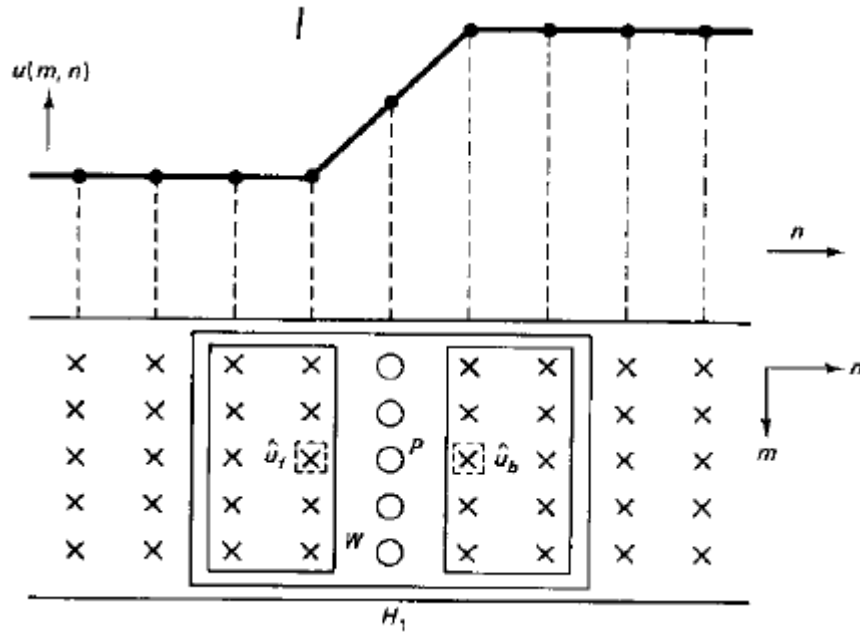


Ilustración 34. Gradientes estocásticos

Para detectar un borde en una localización P se calcula el gradiente horizontal de la siguiente forma:

$$g_1(i, j) = f_f(i, j-1) - f_b(i, j+1) \quad (\text{Ecuación 111})$$

Donde los valores de $f_f(i, j)$ y $f_b(i, j)$ son las aproximaciones anterior y posterior respectivamente de $f(i, j)$, considerando el ruido de dos ventanas de tamaño W cada una, a ambos lados del píxel donde se calcula la existencia de borde. Estas estimaciones pueden ser calculadas de la siguiente forma:

$$f_f(i, j) = \sum_{(k, l) \in W} a(k, l) v(i-k, j-l) \quad (\text{Ecuación 118})$$

Donde $v(i, j)$ es la imagen, $f(i, j)$ sumada con ruido aleatorio y $a(k, l)$ son los coeficientes del filtro. Además, debe especificarse la relación señal ruido (SNR) que se pretende obtener. Esto puede ser expresado a través de máscaras que dependan de la SNR pretendida. En la figura siguiente vemos algunas de estas máscaras:

	SNR = 1	SNR = 9
3 × 3	0.97 0 -0.97 1.00 0 -1.00 0.97 0 -0.97	0.776 0 -0.776 1.00 0 -1.00 0.776 0 -0.776
5 × 5	0.802 0.836 0 -0.836 -0.802 0.845 0.897 0 -0.897 -0.845 0.870 1.00 0 -1.00 -0.870 0.845 0.897 0 -0.897 -0.845 0.802 0.836 0 -0.836 -0.802	0.267 0.364 0 -0.364 -0.267 0.373 0.562 0 -0.562 -0.373 0.463 1.00 0 -1.00 -0.463 0.373 0.562 0 -0.562 -0.373 0.267 0.364 0 -0.364 -0.267

Ilustración 35. Operadores gradientes estocásticos

3.1.3.4 Operadores derivados de Gausianas

Se trata de un operador detector de bordes, al igual que los tres anteriores.

3.1.3.4.1 Laplaciana de Gausiana

También llamado operador Marr-Hildred o de Cruce por cero. Parte del principio de que para detectar un borde de salto, la primera derivada de la función imagen debe tener un máximo en ese punto, y a la vez la segunda derivada debe presentar un cruce por cero en ese mismo punto. Como el operador Laplaciano es muy sensible al ruido se debe suavizar previamente la imagen antes de aplicar la segunda derivada. El filtro empleado debe cumplir dos condiciones:

- El filtro debe ser local, y tomar información únicamente de los puntos cercanos al que se va a analizar
- Debe ser paso banda, para reducir el número de frecuencias en las que se puede producirse el cambio.

La distribución Gausiana optimiza ambas condiciones. El filtro de suavizado gaussiano viene dado por:

$$G(x, y) = ce^{-\frac{x^2+y^2}{2\sigma^2}} \quad (\text{Ecuación 112})$$

Donde c es una constante de normalización y σ es la desviación típica, x e y son la coordenadas de la imagen.

Primero se convoluciona la imagen con la Gausiana $G(x, y)$ y a continuación se le aplica al resultado la Laplaciana:

$$H(x, y) = \nabla^2(G(x, y) * f(x, y)) \quad (\text{Ecuación 113})$$

Este es el operador Marr-Hildreth o LoG.

La ventaja de este operador, frente a los demás operadores de bordes, es que se tiene en cuenta una mayor área de píxeles lejanos, influencia que viene determinada por σ .

A continuación, se van a ver unas máscaras que aproximan la forma gaussiana, para distintas σ .

$\sigma = 0,391 \text{ pixels}$

$$\begin{bmatrix} 1 & 4 & 1 \\ 4 & 12 & 4 \\ 1 & 4 & 1 \end{bmatrix}$$

$\sigma = 0,625 \text{ pixels}$

$$\begin{bmatrix} 1 & 2 & 3 & 2 & 1 \\ 2 & 7 & 11 & 7 & 2 \\ 3 & 11 & 17 & 11 & 3 \\ 2 & 7 & 11 & 7 & 2 \\ 1 & 2 & 3 & 2 & 1 \end{bmatrix}$$

$\sigma = 1,0 \text{ pixels}^*$

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 6 & 7 \\ 1 & 3 & 6 & 9 & 11 \\ 1 & 3 & 7 & 11 & 12 \end{bmatrix}$$

Ilustración 36. Operadores de Marr-Hildreth

3.1.3.4.2 Detector de Canny

Su objetivo es la detección de todos los bordes, consiguiendo que la distancia entre el pixel obtenido como borde y el borde real debe ser lo más pequeña posible.

Estas condiciones se expresan matemáticamente de la siguiente manera:

$$SNR = \frac{A \left| \int_{-W}^0 f(x) dx \right|}{n_0 \sqrt{\int_{-W}^W f^2(x) dx}} \quad (\text{Ecuación 114})$$

$$\text{Localización} = \frac{A |f(0)|}{n_0 \sqrt{\int_{-W}^W f^2(x) dx}} \quad (\text{Ecuación 115})$$

$$\text{Distancia} = \pi \left(\frac{\int_{-\infty}^{\infty} f^2(x) dx}{\int_{-\infty}^{\infty} f'^2(x) dx} \right)^{1/2} \quad (\text{Ecuación 116})$$

Con el detector de Canny se busca la optimización del producto de la relación señal ruido por la localización, así como el cumplimiento de la tercera condición. El detector que cumple con lo anteriormente expuesto es la derivada primera de una Gaussiana, es decir un operador que es combinación de un filtro Gaussiano con una aproximación de gradiente.

3.1.3.5 Medida de la calidad de un detector de bordes

El comportamiento de los operadores se diferencia en el comportamiento que tienen cuando hay presencia de ruido. Evidentemente aquellos que toman en consideración el ruido como son los operadores gaussianos y los estocásticos ofrecerán mejores resultados.

Si se llama n_0 al número total de píxeles que forman el borde y n_1 al número de píxeles que son erróneamente considerados como borde, debido a la presencia de ruido en la imagen, el error de detección es:

$$P = \frac{n_1}{n_0} \quad (\text{Ecuación 117})$$

Otra forma de medir el error de detección es la siguiente:

$$P = \frac{1}{\max(N_1, N_D)} \sum_{i=1}^{N_D} \frac{1}{1 + \alpha d_i^2} \quad (\text{Ecuación 118})$$

Donde:

d_i = menor distancia del borde i -ésimo al borde ideal.

α = constante positiva.

N_i = número ideal de píxeles de borde.

N_d = número total de píxeles detectados.

3.1.3.6 Detección de líneas

La detección de líneas se puede considerar como una extensión de la detección de bordes, ya que para ello se emplean cuatro máscaras en distintas direcciones. Las máscaras podrían ser las siguientes:

$E - O$	$NE - NO$	$N - S$	$NO - NE$
$\begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}$	$\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$

Ilustración 37. Detección de líneas

La primera máscara detectaría líneas horizontales, la segunda, líneas que forman un ángulo de 45° y la tercera y la cuarta, líneas verticales y con ángulo de -45° respectivamente.

3.1.3.7 Detección de esquinas

Existen cuatro métodos para su detección:

- El primero, se basa en detectar los bordes de los objetos y buscar aquellos puntos en los que la curvatura del borde varíe bruscamente.
- Método de Tomasi y Kanade consistente en determinar partes de la imagen que tienen un elevado gradiente horizontal y vertical al mismo tiempo.
- Método de Kitchen y Rosenfeld está basado en la variación de la dirección del gradiente
- Método Susan, que parte de una imagen con varios niveles de gris.

3.1.4 Texturas

Las texturas se definen como patrones visuales homogéneos que se observan en ciertos tipos de materiales como pueden ser madera, tela, piedras.... Dependen del grado de resolución de la imagen, puesto que una imagen puede ofrecer un aspecto homogéneo a cierta distancia y sin embargo no parecerlo o tener otra textura.

La extracción de esta característica va a tener importancia en la segmentación de imágenes mediante texturas así como en la descripción de regiones. El análisis de texturas se puede realizar por dos procedimientos:

- Análisis Estadístico, donde se analizan los estadísticos de primer orden o superiores, de los niveles de gris o de otra propiedad de una determinada zona de la imagen.
- Análisis Frecuencial, donde se realiza el estudio a partir de la transformada de Fourier.

3.1.4.1 Análisis estadístico de texturas

3.1.4.1.1 Estadísticos de primer orden

Como en el apartado de características del histograma, a partir de él podemos obtener propiedades que nos dan información sobre la textura de la imagen. Algunas de estas son la media, que nos da una estimación del nivel de gris de la textura o la desviación típica que indica la media de la dispersión respecto al valor de la media.

3.1.4.1.2 Estadísticos de segundo orden

Hay dos tipos de estadísticos de segundo orden empleados, los obtenidos a partir de Matrices de coocurrencia y los Estadísticos de las diferencias.

3.1.4.1.2.1 Matrices de coocurrencia

Se obtienen calculando la probabilidad condicional P_{δ} , la cual indica que dos propiedades aparezcan separadas por una distancia δ . Para el caso particular de los niveles de gris, la matriz de coocurrencia estaría formada por los elementos $P_r, \Theta(i,j)$ describiendo la frecuencia de aparición de dos píxeles con niveles de gris i,j en una determinada ventana, separados una distancia r en la dirección Θ . Como encontrar matrices de coocurrencia para cada δ es prohibitivo, se suele limitar el cálculo a una serie de ángulos $0^\circ, 45^\circ, 90^\circ, 135^\circ$ y al valor de $r=1$ pixel.

3.1.4.1.2.2 Estadísticos de las diferencias

Se obtienen a partir de la distribución de probabilidad $P_{\delta}(k)$ de los valores pertenecientes a los píxeles intermedios, entre lo que están separados por una distancia δ . Esta distribución se define de la siguiente forma:

$$P_{\delta}(k) = \sum_{i=1}^L \sum_{j=1}^L P_{\delta}(i, j) \quad (Ecuación 119)$$

$$|i - j| = k$$

3.1.4.2 Análisis frecuencial de texturas

Se realiza a partir de la transformada de Fourier. Teniendo en cuenta que la transformada de Fourier de la imagen $f(x,y)$ es $F(u,v)$ su módulo será:

$$P(u, v) = |F(u, v)|^2 \quad (Ecuación 120)$$

Tomando coordenadas polares se obtienen dos distribuciones, la primera indicará el tamaño de la textura dominante y la segunda la dirección de esta:

$$P(r) = 2 \sum_{\theta=0}^{L/2} P(r, \theta) \quad (Ecuación 121)$$

$$P(\theta) = 2 \sum_{r=0}^{L/2} P(r, \theta) \quad (Ecuación 122)$$

3.1.5 Detección de movimiento

Los movimientos detectados pueden ser debidos a movimiento de la cámara, movimientos de los objetos, a cambios en la iluminación o a cambios en la forma o tamaño de los objetos.

Suponiendo la cámara inmóvil, la iluminación constante y los objetos rígidos, uno de los métodos más simples para la detección de movimiento es el basado en la diferencia de imágenes.

Se parte de una secuencia de imágenes $f(x,y,t)$ y la imagen que indicará el campo de velocidades entre los instantes t_0 y t_1 será:

$$F_{t_0 t_1}(x, y) = \begin{cases} 1 & \text{si } |f(x, y, t_0) - f(x, y, t_1)| \geq T \\ 0 & \text{en otro caso} \end{cases} \quad (\text{Ecuación 123})$$

Los métodos presentados tienen la ventaja de ser sencillos, sin embargo tienen varios inconvenientes: la cámara debe estar fija y la imagen resultado indica zonas en las que existe movimiento, pero no cuantifica ni la magnitud ni la dirección.

Existen dos grandes grupos de métodos de obtención del campo de velocidades, que solucionan los problemas anteriores.

- El primero busca una serie de características en la primera imagen para encontrar su correspondencia en la segunda.
- El segundo grupo de métodos pretende la obtención de un mapa denso del campo de velocidades (flujo óptico). Es decir, buscar en un entorno, la zona más parecida en la segunda imagen, dependiendo de sus variaciones espaciales y temporales.

3.2 Segmentación de imágenes

Una vez extraídas las características de una imagen, el siguiente paso es la partición de la imagen para su posterior reconocimiento e interpretación. Por tanto podemos decir que la segmentación consiste en separar los objetos localizados en un entorno, sobre la base de una o varias de las características obtenidas en el proceso de extracción.

La segmentación se basa en tres propiedades:

- Similitud: Los píxeles de un elemento tienen valores parecidos de alguna propiedad.
- Discontinuidad: Los objetos destacan del entorno, teniendo unos bordes definidos.
- Conectividad: Los píxeles pertenecientes a un mismo objeto deben ser contiguos y estar agrupados.

Se puede decir que las técnicas de segmentación se basan en el cumplimiento de una de estas propiedades o combinación de ellas, dando lugar a la búsqueda de partes uniformes de la imagen o de lo contrario, partes donde se produce un cambio.

En función de lo anterior los métodos de segmentación se agruparan en tres grupos:

- Segmentación basada en píxeles(o en umbralización)
- Segmentación basada en bordes.
- Segmentación basada en Regiones.

3.2.1 Segmentación basada en píxeles

Tiene en cuenta el nivel de gris de un píxel, para decidir si él mismo pertenece o no al objeto de interés. Para ellos, se debe encontrar el rango de niveles de gris que caracterizan dicho objeto, teniendo en cuenta la similitud, lo que requiere entonces la obtención y el análisis del histograma de la imagen.

El objetivo de éste método, es el de encontrar de una manera óptima los valores característicos de la imagen que establecen la separación del objeto de interés, con respecto a las regiones que no pertenecen al mismo, es decir encontrar los valores de niveles de gris más apropiados a tomar como umbrales.

Este método también es conocido como la segmentación basada en Umbralización, ya que se convierte una imagen con varios niveles de gris a una nueva con solo dos (imagen binaria), de manera que los objetos quedan separados del fondo. La umbralización se basa en que los píxeles de un determinado objeto tienen el mismo nivel de gris.

A continuación mostramos la función que define la imagen umbralizada según el caso de que se trate:

Cuando el valor buscado es mayor que el umbral,

$$g(x, y) = \begin{cases} 1 & \text{si } T \geq f(x, y) \\ 0 & \text{en otro caso} \end{cases} \quad (\text{Ecuación 124})$$

Cuando es menor que el umbral,

$$g(x, y) = \begin{cases} 1 & \text{si } T \leq f(x, y) \\ 0 & \text{en otro caso} \end{cases} \quad (\text{Ecuación 125})$$

Cuando los objetos pertenecen a un intervalo:

$$g(x, y) = \begin{cases} 1 & \text{si } T_a \leq f(x, y) \leq T_b \\ 0 & \text{en otro caso} \end{cases} \quad (\text{Ecuación 126})$$

3.2.1.1 Mínimo entre máximos

Si la imagen constará del objeto y del fondo, el histograma en el caso ideal tendría dos únicos valores de gris. Como en la vida real existirá ruido, se tendrá algo parecido a picos, por tanto puede suponerse que el umbral será el valor mínimo entre esos dos máximos. En el histograma se debe buscar el nivel de gris que tiene un valor máximo para los dos máximos, y después se busca el valor mínimo entre ambos valores.

3.2.1.2 Medidas iterativas

Suponiendo que el histograma puede asemejarse a dos gaussianas con igual desviación típica, tomamos como el punto de cruce el valor intermedio de las dos medias. Al desconocer su valor se hará una suposición y por un método iterativo se irá acercando al valor real. Para ello se supone que la media del objeto es el valor más alto y el de fondo, el más bajo (o viceversa), se toma por tanto el valor de cruce como el nivel de gris medio de la imagen. A partir de este valor se divide el histograma en dos partes de las que se obtienen sus medias. Se elige el nuevo valor y se vuelven a obtener las dos medias. Si son distintas se obtiene un nuevo valor para la umbralización, si no, se para.

3.2.1.3 Método de Otsu para la determinación del umbral

Al igual que las medidas iterativas, este método también supone que el histograma es la suma de dos gaussianas, aunque ahora las desviaciones típicas pueden ser distintas. El umbral debe minimizar la suma ponderada de cada una de las varianzas de los objetos presentes, ya que se supone que conforme las gaussianas se aproximan más al histograma real, las desviaciones serán menores.

Para determinar los coeficientes se toman las probabilidades de cada uno de los objetos. Así si el umbral es t se tiene que las probabilidades son:

$$q_1(t) = \sum_{i=1}^t P(i) \quad (\text{Ecuación 127})$$

$$q_2(t) = \sum_{i=t+1}^I P(i) \quad (\text{Ecuación 128})$$

Las medias:

$$\mu_1(t) = \sum_{i=1}^t i \frac{P(i)}{q_1(t)} \quad (\text{Ecuación 129})$$

$$\mu_2(t) = \sum_{i=t+1}^I i \frac{P(i)}{q_2(t)} \quad (\text{Ecuación 130})$$

Y las varianzas:

$$\sigma_1^2 = \sum_{i=1}^t (i - \mu_1(t))^2 \frac{P(i)}{q_1(i)} \quad (\text{Ecuación 131})$$

$$\sigma_2^2 = \sum_{i=t+1}^I (i - \mu_2(t))^2 \frac{P(i)}{q_2(i)} \quad (\text{Ecuación 132})$$

Siendo entonces la varianza ponderada:

$$\sigma_W^2 = q_1(t)\sigma_1^2 + q_2(t)\sigma_2^2 \quad (\text{Ecuación 133})$$

Se recorre todo el rango de los niveles de gris calculándose el valor anterior y se elige aquel umbral que minimice esta varianza.

3.2.1.4 Método de Kittle- Illingworth

Se parte de nuevo de dos gaussianas de medias μ_1, μ_2 y desviaciones σ_1, σ_2 con unas proporciones q_1 y q_2 . Este método se basa en minimizar la siguiente expresión:

$$J = \sum_{i=1}^I P(i) \log \frac{P(i)}{f(i)} \quad (\text{Ecuación 134})$$

Siendo:

$$f(i) = \frac{q_1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{i-\mu_1}{\sigma_1}\right)^2} + \frac{q_2}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{i-\mu_2}{\sigma_2}\right)^2} \quad (\text{Ecuación 135})$$

Donde $P(i)$ representa la probabilidad de que aparezca un valor de nivel de gris en el histograma de la imagen y $f(i)$ es la probabilidad de que aparezca un valor de nivel de gris suponiendo que el histograma es la suma de dos gaussianas.

J representa el error al tomar $f(i)$ en vez del histograma $P(i)$ y por tanto hay que encontrar el valor que lo minimice.

3.2.1.5 Uso de la entropía. Método de Pum

La entropía se define como:

$$H = -\sum p_i \log(p_i) \quad (\text{Ecuación 136})$$

El método pretende maximizar la suma de dos entropías cada una correspondiente a las dos gaussianas. Lo que equivale a maximizar:

$$f(t) = \frac{H_t}{H_T} \frac{\log P_t}{\log(\max(P_0, P_1, \dots, P_t))} + \left[1 - \frac{H_t}{H_T}\right] \frac{\log(1 - P_t)}{\log(\max(P_{t+1}, P_{t+2}, \dots, P_N))} \quad (\text{Ecuación 137})$$

Siendo:

$$H_T = -\sum_{i=0}^N P_i \log(p_i) \quad (\text{Ecuación 138})$$

$$H_t = -\sum_{i=0}^t p_i \log(p_i) \quad (\text{Ecuación 139})$$

$$P_t = \sum_{i=0}^t P_i \quad (\text{Ecuación 140})$$

3.2.1.6 Uso de la entropía. Método de Kapur

Se define la probabilidad de que un píxel pertenezca a un nivel de gris de la siguiente forma:

$$\frac{p(i)}{P_T} \quad \frac{p(i)}{1 - P_T} \quad (\text{Ecuación 141})$$

Con lo que las entropías quedarían de la siguiente manera:

$$H_t = -\sum_{i=0}^t \frac{P_i}{P_T} \log\left(\frac{P_i}{P_T}\right) \quad (\text{Ecuación 142})$$

$$H_T = -\sum_{i=0}^N \frac{P_i}{1 - P_T} \log\left(\frac{P_i}{1 - P_T}\right) \quad (\text{Ecuación 143})$$

Al igual que el método anterior hay que maximizar la suma de las dos entropías:

$$H_t + H_T \quad (\text{Ecuación 144})$$

3.2.1.7 Umbralización recursiva

En el caso de que la imagen tuviera más de dos niveles de grises dominantes, para detectar los objetos se emplearía cualquiera de los métodos vistos anteriormente realizando la separación del histograma en dos. Posteriormente se analizaría cada uno de los histogramas individualmente, para ver si es necesario aplicar una nueva umbralización o no. El proceso se repite hasta que no se necesita realizar ninguna umbralización más.

3.2.1.8 Umbralización basada en el color

El color de cada uno de los puntos de la imagen está definido por tres componentes RGB. Para identificar los puntos de un determinado color en la imagen se comprueba que los niveles de sus tres componentes corresponden a los del color buscado. La función que expresa una umbralización por color es la siguiente:

$$g(x, y) = \begin{cases} 1 & R_a \leq f_r(x, y) \leq R_b \\ 1 & G_a \leq f_v(x, y) \leq G_b \\ 1 & B_a \leq f_a(x, y) \leq B_b \\ 0 & \text{en cualquier caso} \end{cases} \quad (\text{Ecuación 145})$$

Donde $f_r(x, y)$, $f_g(x, y)$ y $f_b(x, y)$ son las funciones que dan los niveles de rojo, verde y azul de cada uno de los puntos de la imagen. Se haría la umbralización para cada una de ellas y a partir de los resultados individuales, operados a través de la operación lógica AND se obtendrían los píxeles que cumplen la condición para ese color ya que los píxeles que interesan son lo que han pasado la umbralización para las tres componentes.

Al cambiar la iluminación existe un problema en el espacio RGB ya que cambian los niveles de cada componente. Para solucionar este problema trabajaremos relacionando las componentes ya que si una componente sube o disminuye por efecto de la iluminación, las otras lo harán en la misma proporción, por lo que sus cocientes serán constantes. La nueva función será de la siguiente forma:

$$g(x, y) = \begin{cases} 1 & R_a \leq f_r(x, y) \leq R_b \\ 1 & G'_a \leq \frac{f_g(x, y)}{f_r(x, y)} \leq G'_b \\ 1 & B'_a \leq \frac{f_b(x, y)}{f_r(x, y)} \leq B'_b \\ 0 & \text{en cualquier caso} \end{cases} \quad (\text{Ecuación 146})$$

3.2.1.9 Histogramas laterales

Esta técnica utiliza los histogramas laterales, en los que se suman los valores de los píxeles por columnas y por filas. Serán útiles para determinar en qué zona se encuentran los objetos.

3.2.1.10 Umbralización Adaptativa

Las imágenes en las que la iluminación no sea constante, presentan el problema de que puntos que pertenezcan al mismo objeto, presenten distintos valores en sus niveles de gris. Para separar el objeto del fondo, se divide la imagen en varias subimágenes de dimensiones $n \times m_i$ donde se supone que los cambios de la iluminación no son fuertes. A cada una de las subimágenes se aplicaría alguno de los métodos anteriores para que el umbral se vaya adaptando a los cambios de iluminación.

3.2.1.11 Etiquetado

Con el etiquetado conseguimos la separación de los objetos que mediante la umbralización no hayan podido ser distinguidos entre sí, ya que todos tienen el mismo nivel de gris. Consiste en ir asignando etiquetas a los píxeles, mirando si los contiguos tienen el mismo valor o no. Si lo tienen la etiqueta será igual y si no será diferente.

3.2.2 Segmentación basada en bordes

Se basan en realizar la búsqueda del valor máximo del gradiente sobre cada línea que forma la imagen. Cuando un máximo es encontrado, un algoritmo de trazado trata de seguir el máximo del gradiente alrededor del objeto, hasta encontrar de nuevo el punto inicial, para luego buscar el próximo máximo en el gradiente.

Parten de los bordes encontrados en la imagen por los detectores, ya que estos últimos por si solos no pueden unir los bordes, puesto que están afectados por el ruido de la imagen y además hay partes que no son detectadas.

Existen diferentes métodos para la segmentación que difieren en la estrategia para la construcción del contorno, así como en la cantidad de información previa que se incorpora en el método. A continuación mostramos alguno de estos métodos.

3.2.2.1 Conectividad

La extracción de contornos puede ser entendida como la unión de los bordes que se encuentren conectados. Un píxel se dice que está conectado cuando tiene alguna propiedad común con los píxeles de su entorno.

3.2.2.2 Seguimiento del contorno

Los métodos de seguimiento del contorno, trazan los contornos ordenando los puntos de tipo borde sucesivos. Un ejemplo de algoritmo es el siguiente:

1. Comenzar dentro de una región cualquiera; por ejemplo, por el primer píxel encontrado en la región al hacer un rastreo por filas.
2. Girar a la izquierda y pasar al píxel siguiente, si esta dentro de esta misma región, en caso contrario girar a la derecha y pasar al píxel siguiente.
3. Continuar hasta que se llegue al punto del que se partió.

3.2.2.3 Unión de bordes y búsqueda heurística en Grafos

Se puede entender el contorno como un camino, que se forma a través de un grafo construido uniendo los elementos bordes. Suponiendo que un grafo con nodos en las posiciones x_k , $k=1,2,..$ se forma desde el nodo A al B; una función de evaluación $\phi(x_k)$ da el valor del camino de A a B cuando pasa por el nodo x_k . Los algoritmos de búsqueda heurística, examina los sucesores del nodo de salida A y seleccionan el que maximiza la función $\phi(x)$. El nodo seleccionado se convierte ahora en el nodo de salida y se repite el proceso hasta alcanzar B. La sucesión de nodos seleccionada constituye el contorno.

3.2.2.4 Programación dinámica

El método anterior, solamente nos aseguraba el camino “óptimo” entre dos nodos consecutivos, pero no que este fuera el camino óptimo global. Para resolver este problema aparecen los métodos de programación dinámica. Estos métodos se basan en el principio de optimización de Bellman, el cual dice que “el camino óptimo entre dos puntos dados de un camino, es también óptimo entre dos puntos cualquiera que se encuentre en él”.

Para aplicar esta idea a la extracción de contornos, se va a suponer que se ha convertido el mapa de los puntos que son bordes de una imagen, en un grafo de N nodos. La función de evaluación será la siguiente:

$$S(x_1, x_2, \dots, N) = \sum_{K=1}^N |g(x_k)| - \alpha \sum_{K=2}^N |\theta(x_k) - \theta(x_{k-1})| - \beta \sum_{K=2}^N d(x_k, x_{k-1}) \quad (\text{Ecuación 147})$$

Donde x_k , $K=1, \dots, N$ representan los nodos, es decir los píxeles pertenecientes al borde; $d(x,y)$ es la distancia entre dos nodos x e y . $|g(x_k)|$, $\theta(x_k)$ son el módulo y el ángulo del gradiente de nodo en la posición x_k ; y los parámetros α y β son constantes no negativas. Teniendo en cuenta esta definición, se dice que el camino que conecta los N nodos, es óptimo cuando la función S sea máxima.

$$\phi(x_N, N) = \max\{S(x_1, \dots, x_N, N)\} \quad (\text{Ecuación 148})$$

3.2.2.5 Relajación de bordes

Este método tiene en cuenta los bordes vecinos para la elaboración del contorno. Todas las propiedades de una imagen junto con la posibilidad de existencia de más bordes, se evalúan iterativamente con más precisión hasta que el contexto de los bordes queda claro. Se basa en la fuerza de los bordes en una vecindad local especificada, aumentando o disminuyendo la confianza de cada borde. Un borde débil posicionado entre dos bordes fuertes, proporciona un ejemplo de contexto, es altamente probable que este borde débil entre ellos sea parte de un contorno. Si, por otra parte, un borde, incluso uno fuerte no tiene contexto que lo soporte, puede que no forme parte de ningún contorno.

3.2.2.6 La transformada de Hough

Se basa en la información que suministra toda la imagen. La transformada de Hough hace una transformación de los puntos de la imagen a un espacio de parámetros de dimensión n (adecuada a la forma geométrica buscada) a partir del cual, en función del valor de los parámetros más repetido, determina la forma geométrica en cuestión. En el espacio transformado, en cada punto se acumulará un valor entero, correspondiente al número de votos, que indicará la presencia de esa forma geométrica en la imagen.

3.2.3 Segmentación orientada a regiones

3.2.3.1 Segmentación de regiones por clasificación

Se utilizan los niveles de gris presentes en la imagen para obtener una partición del espacio. Se asocia a cada píxel, la clase de nivel de gris a la cual pertenece. Las regiones son definidas por los conjuntos de píxeles conexos pertenecientes a una misma clase. La segmentación por clasificación utiliza el histograma como utensilio para la clasificación de los distintos brillos. Es un método eficaz, si la clasificación permite definir, las diferentes regiones homogéneas de la imagen.

3.2.3.2 Segmentación por crecimiento de regiones

Las regiones deben ser homogéneas maximales, donde esto último significa que los criterios de homogeneidad a partir de los cuales se forman las regiones no se seguirán cumpliendo tras la unión de una región con alguna de sus adyacentes.

3.2.3.3 Unión de regiones

Considera a cada píxel como una región. Obviamente estas regiones serán homogéneas pero no necesariamente maximales por lo que el proceso deberá repetirse hasta que las regiones sean maximales. El método se puede describir:

- Definir una segmentación inicial que cumpla el criterio de homogeneidad.
- Definir un criterio para unir regiones adyacentes.

- Unir las regiones adyacentes, si cumplen el criterio de unión. Para cuando no puedan unirse dos regiones sin romper el criterio de homogeneidad.

El proceso por tanto se puede resumir como que la descripción de una región se compara con la de otra adyacente. Si coinciden, se unen las regiones y se computa el nuevo descriptor de la región. En caso contrario, las regiones se marcan como que no pueden unirse y el proceso continúa.

3.2.3.4 Unión de regiones por agregación de píxeles

El crecimiento de regiones es un procedimiento que agrupa píxeles o subregiones en regiones mayores. Las regiones crecen mediante la agregación de píxeles similares respecto a la propiedad utilizada para la segmentación. El punto de partida es un conjunto de píxeles semilla, es decir los píxeles sobre los cuales se evalúan los demás, y a partir de los cuales se hace crecer la región, añadiendo a dichos píxeles semilla, aquellos vecinos que tienen propiedades similares (nivel de gris, color y textura).

3.2.3.5 Separación y unión de regiones

Consiste en subdividir la imagen inicialmente en un conjunto de regiones disjuntas y luego unir las o separarlas en un intento de que todas ellas:

1. Cubran la imagen completa.
2. Cada una de ellas esté conectada.
3. Sus intersecciones dos a dos, sea el conjunto vacío.
4. Que algún criterio de homogeneidad sea cumplido en cada región.
5. Que la unión de dos cualesquiera de ellas no cumpla algún criterio de homogeneidad.

3.2.4 Segmentación basada en texturas

Utiliza las propiedades explicadas en el apartado 3.2, para realizar la partición de las imágenes. Pudiendo optar por una aproximación estadística con sus correspondientes propiedades diferenciadoras entre unas partes y otras, o por una aproximación por frecuencias, definiendo el contenido local en frecuencias de una determinada región.

3.3 Transformaciones morfológicas

Debido a la segmentación, obtenemos de la imagen los elementos constituyentes más importantes. Estos son objetos que presentan defectos inherentes al proceso de segmentación, como ruido o píxeles más clasificados, por lo que se hace necesario un posterior procesamiento para solventar este problema. Este proceso consiste en la aplicación de transformaciones morfológicas que son aquellas que modifican la forma o estructura de dichos objetos, de forma que se simplifican los datos de la imagen, y se preserven las características esenciales eliminando los aspectos irrelevantes. A continuación vamos a ver las

Análisis de imágenes basado en correlación de imágenes

transformaciones aplicadas tanto a imágenes binarias como a imágenes con varios niveles de gris.

Para la realización de estas transformaciones vamos a recordar ciertas operaciones básicas sobre conjuntos. Para ello nos basaremos en dos conjuntos A y B en un espacio E, con elementos a y b pertenecientes a A y B respectivamente.

- Inclusión. A está incluido en B, $A \subset B$
- Complemento: A^c contiene los elementos que no pertenecen a A.
- Unión: son todos los elementos pertenecientes a uno de los conjuntos.

$$A \cup B = \{x \mid x \in A \text{ ó } x \in B\} \quad (\text{Ecuación 149})$$

- Intersección: son los elementos comunes a ambos conjuntos.

$$A \cap B = \{x \mid x \in A \text{ y } x \in B\} \quad (\text{Ecuación 150})$$

- Reflexión:

$$\hat{A} = \{x \mid x = -a, \text{ para todo } a \in A\} \quad (\text{Ecuación 151})$$

- Diferencia entre A y B

$$A - B = A \cap B^c \quad (\text{Ecuación 152})$$

- Traslación: Un conjunto A es trasladado un vector v cuando cada uno de los elementos de A sufre esa traslación.

$$A_v = \{x \mid x = a + v, \text{ para todo } a \in A\} \quad (\text{Ecuación 153})$$

Dos operaciones basadas en la traslación, son la suma y resta de Minkowski:

$$\text{Suma de Minkowski: } A \oplus B = \bigcup_{x \in B} A_x \quad (\text{Ecuación 154})$$

$$\text{Resta de Minkowski: } A \ominus B = \bigcap_{x \in B} A_x \quad (\text{Ecuación 155})$$

Donde la suma es la unión, y la resta la intersección del conjunto A y el B trasladado su origen a cada elemento de A

3.3.1 Transformaciones morfológicas en imágenes binarias

En una imagen binaria, los píxeles con valor 1 se corresponden con píxeles de la imagen, mientras que los píxeles con valor 0 serán considerados como puntos pertenecientes al fondo.

Este tipo de transformaciones requiere la utilización de un elemento estructural, con cierta forma y tamaño, con el que se irá probando la imagen y cuantificando el modo en que está contenido en la misma.

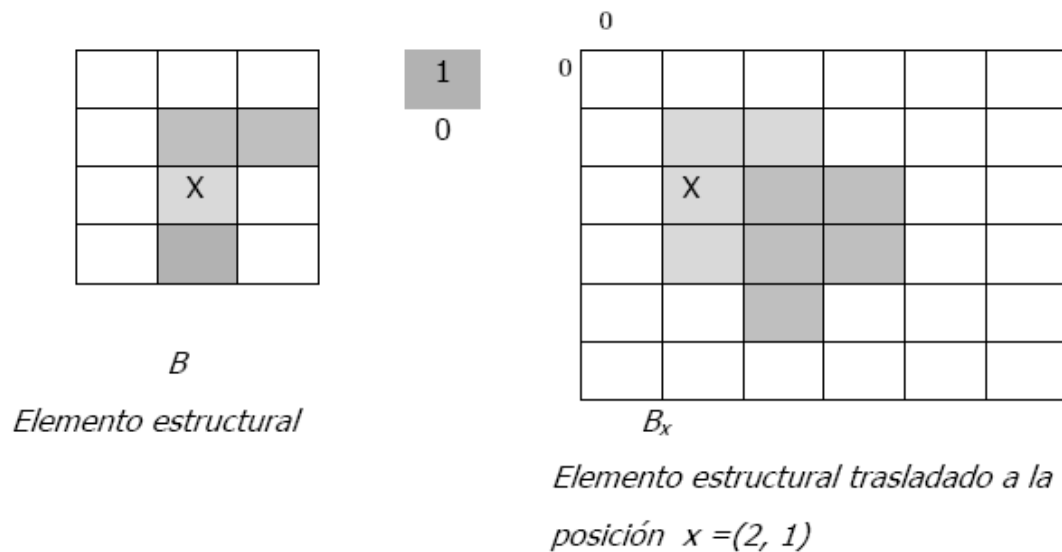


Ilustración 38

La forma de operar con un elemento estructural es la siguiente:

- Elegir un elemento estructural adecuado
- Desplazar B por todas las posiciones del espacio E en el que se encuentra el conjunto a estudiar A : $B_x \equiv B_{(x,y)}$
- En cada posición comprobar si la unión, la intersección o la inclusión de A con B , está dentro de A .
- Los puntos con respuesta afirmativa formarán parte de un nuevo conjunto que será en este caso la imagen transformada.

Considerando el conjunto A como la imagen sobre la que se realiza la transformación y el conjunto B como el elemento estructural describiremos las principales operaciones morfológicas.

3.3.1.1 Erosión

Consiste en una reducción de algún elemento de la imagen original, es decir la degradación de uno de los campos 0 ó 1 pertenecientes a la imagen binaria. Un punto del campo a , seguirá perteneciendo al mismo si esta rodeado de puntos iguales a él, sino pasará a formar parte del otro campo.

Se define la erosión de A por B como: $(A, B) = \{x \mid B_x \subset A\}$

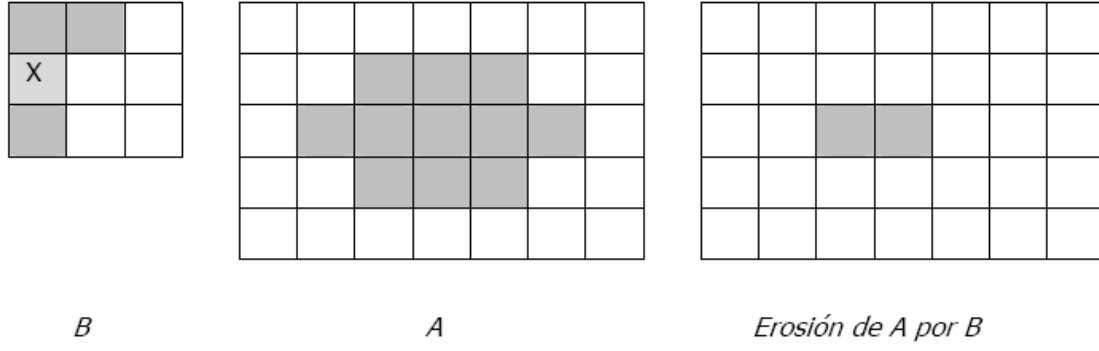


Ilustración 39

Para realizar la erosión habría que recorrer la imagen píxel a píxel, y en cada píxel a su vez recorrer todo el elemento estructural, por ello se dará la definición de la erosión, en función de la Resta de Minkowski, más fácil de implementar:

$$A \ominus B = \bigcap_{b \in B} A_b \quad (\text{Ecuación 156})$$

Con esta definición se considera la erosión como la intersección de todas las traslaciones de A por los puntos $-b$ pertenecientes a la reflexión de B (B).

3.3.1.1.1 Propiedades

- Antixtensiva: si el elemento se encuentra dentro de un elemento estructural la erosión tiene el efecto de encoger la imagen de entrada, obteniendo una imagen incluida dentro de la original. De no ser así se puede llegar a un resultado que no tiene nada que ver con la imagen original.

$$\text{Si } B \text{ contiene al origen, } A \ominus \bar{B} \subseteq A \quad (\text{Ecuación 157})$$

- Invariante a traslaciones

$$A_x \ominus \bar{B} = (A \ominus B)_x \quad (\text{Ecuación 158})$$

$$A \ominus \bar{B}_x = (A \ominus B)_{-x} \quad (\text{Ecuación 159})$$

- Creciente:

$$A \subseteq C \rightarrow A \ominus \bar{B} \subseteq C \ominus \bar{B} \quad (\text{Ecuación 160})$$

- La intersección es distributiva con respecto a la erosión:

$$(A \cap C) \ominus \bar{B} = (A \ominus \bar{B}) \cap (C \ominus \bar{B}) \quad (\text{Ecuación 161})$$

- La unión no es distributiva con respecto a la erosión:

$$(A \cup C) \ominus \bar{B} \supseteq (A \ominus \bar{B}) \cup (C \ominus \bar{B}) \quad (\text{Ecuación 162})$$

3.3.1.2 Dilatación

Es el crecimiento progresivo de uno de los campos. Si un elemento del campo que crece tiene un vecino del campo contrario, este último se convierte al campo que se dilata. Si el vecino es del mismo campo el elemento no se altera.

Se define la dilatación de A por B como: $(A, B) = \{x \mid B_x \cap A \neq \emptyset\}$

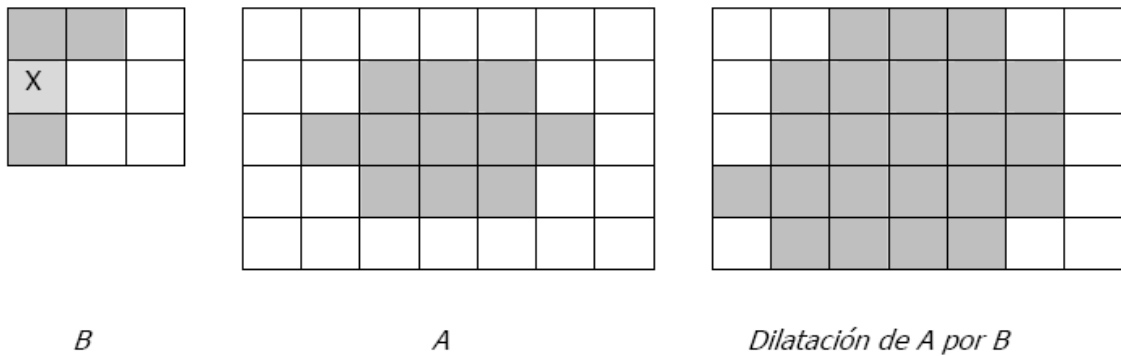


Ilustración 40

La dilatación se expresa en función de la Suma de Minkowski. De esta forma, queda expresada como la unión de traslaciones de A por los elementos de B.

$$A \oplus B = \bigcup_{x \in B} A_x \quad (\text{Ecuación 162})$$

3.3.1.2.1 Propiedades

- Conmutativa:

$$A \oplus B = B \oplus A \quad (\text{Ecuación 163})$$

- Asociativa:

$$A \oplus (B \oplus C) = (A \oplus B) \oplus C \quad (\text{Ecuación 164})$$

- Invariante a traslaciones:

$$A_x \oplus \bar{B} = (A \oplus B)_x \quad (\text{Ecuación 165})$$

- Una traslación en la imagen puede compensarse en el elemento estructural, tomando este último trasladado en la dirección opuesta:

$$A_x \oplus B_{-x} = A \oplus B \quad (\text{Ecuación 166})$$

- Creciente:

$$A \subseteq C \rightarrow A \oplus B \subseteq C \oplus B \quad (\text{Ecuación 167})$$

- Extensiva: Si el origen se encuentra dentro del elemento estructural, la dilatación agranda la imagen de entrada, obteniendo una imagen que incluye a la original. De no ser así, la imagen resultante puede no incluir a la imagen original.

$$\text{Si } B \text{ contiene al origen } A \oplus B \supseteq A \quad (\text{Ecuación 168})$$

- La unión es distributiva con respecto a la dilatación.

$$(A \cup B) \oplus C = (A \oplus C) \cup (B \oplus C) \quad (\text{Ecuación 169})$$

$$(C \cup B) \oplus A = (A \oplus B) \cup (A \oplus C) \quad (\text{Ecuación 170})$$

- La intersección no es distributiva con respecto a la dilatación, pero cumple:

$$(A \cap B) \oplus C \subseteq (A \oplus C) \cap (B \oplus C) \quad (\text{Ecuación 170})$$

$$A \oplus (B \cap C) \subseteq (A \oplus B) \cap (A \oplus C) \quad (\text{Ecuación 171})$$

- Dualidad entre erosión y dilatación: Son duales cuando la negación de una formulación empleada en el primero es igual a la misma formulación empleando el segundo operador en la variable negada. Se podría resumir, en que lo que un conjunto hace al objeto el otro lo hace al fondo.

$$(A \ominus B)^c = A^c \oplus \bar{B} \quad (\text{Ecuación 171})$$

$$(A \oplus B)^c = A^c \ominus \bar{B} \quad (\text{Ecuación 172})$$

3.3.1.3 Acierta o falla (hit or miss)

Se aplica de la siguiente forma:

A partir del elemento estructural B se forma el conjunto B_x , que es ese mismo elemento estructural desplazado para todo elemento $x \in A$. Se considera B_x formado por dos subconjuntos B_{x1} y B_{x2} , correspondiendo el primero a los elementos del objeto y el segundo a los del fondo. Un punto pertenece a Acierta o Falla, si y sólo si B_{x1} está incluido en A y B_{x2} está incluido en A^c , es decir:

$$A \otimes B = \{x \mid B_{x1} \subset A; B_{x2} \subset A^c\} \quad (\text{Ecuación 173})$$

Esta transformación indica donde coincide exactamente el elemento estructural B en el conjunto A.

3.3.1.4 Apertura (Opening)

Consiste en una erosión, seguida de una dilatación, empleando el mismo elemento estructural.

$$A \circ B = (A \ominus \bar{B}) \oplus B \quad (\text{Ecuación 174})$$

El efecto será el suavizado de los contornos de los objetos, rotura de enlaces delgados, y descomposición de objetos, entre otros efectos.

3.3.1.5 Cierre (Closing)

Consiste en una dilatación, seguida de una erosión.

$$A \bullet B = (A \oplus B) \ominus \bar{B} \quad (\text{Ecuación 175})$$

Tiende a suavizar el contorno de la imagen de entrada, rellenando roturas y pequeños agujeros y completando espacios del contorno.

Las operaciones morfológicas que vamos a comentar posteriormente, son transformaciones basadas en las operaciones básicas vistas anteriormente y que se utilizan principalmente para la detección de bordes o preprocesamiento.

3.3.1.6 Extracción del contorno

Se obtiene el contorno de un conjunto A, denotándose $\beta(A)$, erosionando A por B y realizando posteriormente la diferencia entre A y su erosión:

$$\beta(A) = A - (A \ominus B) \quad (\text{Ecuación 176})$$

Así obtendremos el contorno interior del conjunto A. En cambio utilizando la dilatación se puede obtener el contorno exterior de la siguiente forma:

$$P(A) = (A \oplus B) - A \quad (\text{Ecuación 177})$$

3.3.1.7 Relleno de regiones

Utilizando operaciones morfológicas se quiere rellenar el interior de un contorno.

3.3.1.8 Extracción de componentes conexas

Se pretende extraer las componentes conexas contenidas en un conjunto A, es decir, obtener los elementos conectados entre si de dicho conjunto. Para conseguir nuestro objetivo seguimos un proceso iterativo basado en la dilatación.

3.3.1.9 Envoltente convexa (convex hull)

Es la región más pequeña que contiene al objeto, de forma que dos puntos cualesquiera pertenecientes a dicha envoltente pueden conectarse a través de una línea recta, perteneciendo a la región todos los puntos de ella. Para conseguir la envoltente convexa de un conjunto hay que utilizar un procedimiento, basado en transformaciones acierta o falla iterativas.

3.3.1.10 Adelgazamiento (Thinning)

El adelgazamiento de un conjunto A por un elemento estructural B se define en función de la transformación acierta o falla de la siguiente forma:

$$A \oslash B = A - (A \otimes B) = A \cap (A \otimes B)^c \quad (\text{Ecuación 178})$$

3.3.1.11 Engrosamiento (Thickening)

El engrosamiento es el dual morfológico del adelgazamiento y se define de la forma siguiente:

$$A(\bullet)B = A \cup (A \otimes B) \quad (\text{Ecuación 179})$$

3.3.1.12 Esqueletización (Skeletonization)

Por esqueletización se entiende la obtención del esqueleto de una región (objeto), que no es más que una representación en forma de grafo de dicha región. Una de sus características es que pequeñas variaciones de la región se traducen en grandes variaciones en su esqueleto, lo que resulta particularmente útil para detectar pequeños defectos en piezas y en sistemas para el control de calidad. El esqueleto de un conjunto A, denotado S(A) se puede escribir como:

$$S(A) = \bigcup_{k=0}^L S_k(A) \quad (\text{Ecuación 180})$$

Con:

$$S_k(A) = \bigcup_{k=0}^K \{ (A \ominus k B) - [(A \ominus k B) \circ B] \} \quad (\text{Ecuación 181})$$

3.3.1.13 Eliminación de ramas (Pruning)

Tras un proceso de adelgazamiento o de esqueletización, el esqueleto suele presentar unas ramificaciones acompañando al tronco principal, que no aportan información acerca del objeto y además suelen dificultar su identificación, por ello se hace necesario un proceso de poda (pruning) o eliminación de ramas.

3.3.1.14 Eliminación de ruido

Tras la segmentación de una imagen habrá píxeles pertenecientes al objeto que en realidad pertenezcan al fondo, y viceversa. Con erosiones y dilataciones se puede solventar el problema eliminando los huecos de los objetos y los puntos ruidosos del fondo. La expresión que explica este proceso es la siguiente:

$$X_4 = (((A \ominus B) \oplus B) \oplus B) \ominus B \quad (\text{Ecuación 182})$$

Lo que equivale a una apertura seguida de un cierre:

$$X_4 = (A \circ B) \bullet B \quad (\text{Ecuación 183})$$

3.3.2 Transformaciones morfológicas en imágenes con varios niveles de gris

Se trata de extender las transformaciones morfológicas vistas anteriormente a imágenes con distintos niveles de gris.

3.3.2.1 Dilatación

La dilatación de una imagen con varios niveles de gris $f(x,y)$ por un elemento estructural $b(x,y)$, que ahora es una función, se define:

$$(f \oplus b)(s,t) = \max \{ f(s-x, t-y) + b(x,y) \mid (s-x, t-y) \in D_f; (x,y) \in D_b \} \quad (\text{Ecuación 184})$$

Donde D_f es el dominio de la imagen $f(x,y)$ y D_b el del elemento estructural. Con los dominios se asegura que las dos imágenes coincidan en algún punto.

3.3.2.2 Erosión

La erosión para imágenes con varios niveles de gris se define como:

$$(f \ominus b)(s,t) = \max \{ f(s+x, t+y) - b(x,y) \mid (s+x, t+y) \in D_f; (x,y) \in D_b \} \quad (\text{Ecuación 185})$$

Donde D_f es igualmente el dominio de la imagen $f(x,y)$ y D_b el del elemento estructural.

3.3.2.3 Apertura y cierre

Son las mismas que para imágenes binarias.

La apertura de f por b :

$$f \circ b = (f \ominus b) \oplus b \quad (\text{Ecuación 186})$$

Igualmente el cierre de f por b :

$$(f \bullet b)^c = f^c \circ \bar{b} \quad (\text{Ecuación 187})$$

3.4 Representación y descripción de contornos y regiones

Una vez que la imagen ha sido segmentada en regiones y en algunos casos corregidos los errores cometidos en dicha segmentación mediante transformaciones morfológicas, es necesario representar y describir dichos píxeles, mediante unas características que se obtendrán a partir de esas regiones, para que puedan ser procesados posteriormente.

Una región se puede representar de dos maneras, representado su contorno o representando el interior de la región. El que se represente de una forma o de otra va a depender del tipo de características que se quiera extraer. En general si se trata de características geométricas se representará el contorno y si son características físicas se representará la región.

3.4.1 Representación del contorno

3.4.1.1 Códigos cadena

Los códigos cadena representan contornos mediante una sucesión de segmentos conectados de una longitud y dirección dada.

3.4.1.2 Curvas Φ -S

Este método es similar al de los códigos cadena. Pero en este caso se codifica el ángulo Φ de los vectores tangentes en cada punto del contorno, en función de la longitud de arco S de contorno. El ángulo Φ está referenciado al ángulo de la tangente en el punto origen del contorno (ángulo cero), de forma que la curva Φ -S empieza y termina en cero. Tiene las siguientes ventajas:

- Es invariante a traslaciones.
- Las limitaciones en los ángulos posibles sólo están condicionadas por el método de segmentación utilizado.
- Es invariante a rotaciones.

3.4.1.3 Aproximaciones poligonales

Estos métodos aproximan el contorno mediante un polígono, es decir, representan tramos curvos mediante tramos lineales. De esta forma hacen corresponder segmentos de línea al contorno, almacenando sólo los parámetros de esos segmentos de línea en lugar de los puntos discretos del contorno.

3.4.1.4 Método de firmas

En general una firma es una representación uni-dimensional de un contorno bi-dimensional y se puede generar de varias formas.

3.4.2 Representación de regiones

La forma de un objeto puede representarse directamente por la región que ocupa en la imagen, de manera que se construye una imagen binaria con la siguiente expresión:

$$f(x, y) = \begin{cases} 1 & \text{si } (x, y) \text{ pertenece a la región} \\ 0 & \text{en caso contrario} \end{cases} \quad (\text{Ecuación 188})$$

Para representar dicha región de forma más efectiva existen una serie de métodos, de los cuales vamos a ver algunos a continuación.

3.4.2.1 Códigos de longitud variable

Una región o imagen binaria puede ser vista como una sucesión en la que se alternan series de ceros y unos. Si se considera que el objeto, codificado con unos, ocupa en la mayoría de los casos un espacio en la imagen inferior al que ocupa el fondo. Si se codifica exclusivamente las cadenas de unos consecutivos se minimizarán el espacio de memoria ocupado por la imagen. De esta forma, se codificará el comienzo de las cadenas de unos y la longitud de esas cadenas.

3.4.2.2 Proyecciones

Una región puede representarse por sus proyecciones. En general, la proyección de una imagen sobre una recta consiste en obtener para cada punto de ella, el número de píxeles que pertenecen a la perpendicular que pasa por ese punto.

3.4.3 Descriptores de contornos

Son un conjunto de números producidos para describir la forma de un objeto. Estos números se calculan mediante el cálculo de algunas características en particular. Es posible que la forma del objeto no se pueda describir de forma completa a partir de los descriptores, pero estos deben ser suficientemente diferentes para poder distinguir objetos con distintas formas.

3.4.3.1 Longitud de un contorno

Se obtiene contando el número de píxeles del contorno. Si el contorno viene codificado con códigos cadena, se sumarán los componentes horizontales, los verticales y los diagonales multiplicados por $2^{1/2}$.

3.4.3.2 Diámetro de un contorno

Se define como:

$$Diam(C) = \text{máx}[D(p_i, p_j)] \quad (\text{Ecuación 189})$$

Donde D es la distancia entre dos puntos del contorno C y p_i y p_j , dos puntos pertenecientes al contorno. La línea que une los dos puntos del diámetro llamada eje mayor del contorno también es un descriptor útil.

3.4.3.3 Esquinas

La curvatura a lo largo del borde de un objeto se define:

$$|k(t)|^2 = \left(\frac{d^2y}{dt^2}\right)^2 + \left(\frac{d^2x}{dt^2}\right)^2 \quad (\text{Ecuación 190})$$

Son declarados esquinas aquellos puntos donde la función $k(t)$ toma valores elevados, quedando así el objeto definido por sus esquinas.

3.4.3.4 Descriptores de Fourier

Consiste en el cálculo de la transformada de Fourier de un conjunto de píxeles. Considerando el significado de la transformada, si hay contenido en alta frecuencia significa que hay cambios rápidos en la coordenada correspondiente y que si hay contenido en baja frecuencia significa que el contorno varía de forma suave. Dicho de otra forma, las componentes de baja frecuencia capturan la forma general del contorno y las de alta frecuencia capturan los detalles. Esto implica que si se realiza un filtrado paso bajo de los descriptores de Fourier de un contorno sería equivalente a suavizar este contorno.

3.4.4 Descriptores de regiones

Se encargan de caracterizar una región resultante de una segmentación, cumpliendo los mismos requisitos que los descriptores de contornos en cuanto a que deben ser invariantes.

3.4.4.1 Área

El área de una región vendrá dada por la siguiente ecuación:

$$A = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \quad (\text{Ecuación 191})$$

Donde $f(x,y)$ es la imagen binaria.

3.4.4.2 Perímetro

El perímetro P está constituido por el número de píxeles exteriores de la región, es decir los del contorno.

3.4.4.3 Compactidad

Refleja el empaquetamiento de la región y se define como el perímetro al cuadrado partido por el área:

$$C = \frac{P^2}{A} \quad (\text{Ecuación 192})$$

Al ser un descriptor sin dimensiones es insensible a cambios de escala y salvo errores de redondeo debido a la rotación, es invariante ante la rotación.

3.4.4.4 Descriptores topológicos

Indican de una forma sencilla alguna idea sobre la forma de la región, por ejemplo:

- El número de agujeros en la región (H).
- El número de componentes conectados, es decir partes separadas que forman la región (C).
- El número de Euler, que es la diferencia entre los dos anteriores ($E=C-H$)

3.4.4.5 Texturas

Algunas regiones pueden ser descritas como en el apartado referente a texturas, donde fueron obtenidas a partir de métodos estadísticos o frecuenciales.

4. Correlación Digital de Imágenes (DIC)

4.1 Introducción

La correspondencia de elementos homólogos pertenecientes a diferentes imágenes se encuentra en continuo desarrollo por parte de disciplinas como el Análisis de Imagen, la Visión Computacional y la Fotogrametría. Sin embargo, la dificultad depende de los diferentes campos, desde los casos más favorables como los de la Fotogrametría tradicional estereoscópica, hasta casos de gran dificultad de cómputo como los implementados en aplicaciones de Fotogrametría de Rango Cercano o Visión Computacional donde existen grandes diferencias en las condiciones geométricas y radiométricas de las imágenes. Por tanto, podemos afirmar que el problema de la correspondencia de imágenes ó “matching” es una línea de investigación en continua evolución.

Dedicaremos este tema a conocer más en profundidad la técnica de correspondencia de imágenes: desde la asignación de la relación entre dos puntos, hasta los acercamientos que combinan varias estrategias con objeto de establecer una correspondencia entre un punto o elemento de la imagen y una categoría cartográfica más o menos abstracta.

4.2 Antecedentes.

Remontándonos a los primeros trabajos relacionados con este tema, la correspondencia de imágenes o “matching de imágenes” tiene una larga historia. Los primeros experimentos datan de los años cincuenta, especialmente con Hobrough (1959) que planteó una solución de naturaleza analógica, en la que los correladores, realizados mediante hardware comparaban los niveles de gris de dos imágenes. Wild Heerbrugg, entonces el fabricante líder en equipos fotogramétricos, presentó un correlador en el congreso de la ISPRS en 1968, que a pesar de su innovación no tuvo mucho éxito, ya que no era muy flexible ni exacto.

Desde los primeros años de los setenta hasta mediados de los ochenta, la investigación relativa a la correspondencia de imágenes se centró en las técnicas de la correlación digital. Kreiling (1976) fue uno de los primeros pioneros en aplicar correlación digital para generar Modelos Digitales de Elevación (MDE) y ortofotos digitales. Los intentos continuaron añadiendo técnicas de correlación a los instrumentos (Helava, 1978) y (Hoborough, 1978), sin embargo, no se encontró una solución universal a pesar del esfuerzo realizado. Todo esto produjo que los investigadores de muy diversas disciplinas no dieran crédito a la facilidad con que una persona encontraba puntos homólogos y la dificultad tan enorme de que la misma tarea la realizase un ordenador. Hoy, se sabe que la visión estereoscópica natural no utiliza la correlación de niveles de gris tal y como fue puesto de manifiesto por Grimson (1981) y Horn (1983). Como puede observarse, esto pone de manifiesto la subestimación de la complejidad de las tareas visuales que las personas resuelven con gran facilidad.

4.3 Conceptos básicos.

No existe una terminología estandarizada en lo que se refiere a la correspondencia de imágenes, y en ocasiones esto produce una fuerte barrera entre las diferentes disciplinas que participan en la correlación. Debido a esto, a continuación repasaremos algunos conceptos básicos en la correspondencia de imágenes:

Análisis de imágenes basado en correlación de imágenes

- Entidad homóloga o conjugada

Es un término más genérico que punto homólogo. Se denominan entidades conjugadas a las imágenes del espacio objeto, que incluyen puntos, líneas y superficies.

- Entidad de correspondencia

Es la parte de la imagen (primitivas), que se compara con primitivas de otras imágenes para buscar entidades homólogas. Las primitivas incluyen niveles de gris, elementos extraídos y descripciones simbólicas.

- Índice de semejanza

Es una medida cuantitativa de lo buena que es la correspondencia de las entidades entre sí. En su forma más simple puede ser el coeficiente de correlación cruzado o la desviación típica en la correspondencia por mínimos cuadrados. Cuanto mayor es este valor, mayor es la correspondencia entre imágenes.

- Método de correspondencia

Calcula el índice de semejanza entre entidades homólogas. Habitualmente la descripción del método se añade tras el nombre de la entidad de correspondencia, por ejemplo correspondencia por áreas, correspondencia basada en características y correspondencia relacional (simbólica).

- Estrategia de correspondencia

Se utiliza para referirse al esquema en el cual nos vamos a basar, para obtener la solución del problema de correspondencia de imágenes. Abarca el análisis del entorno de correspondencia, la selección del método para la misma y el correspondiente control de calidad de la ejecución.

A continuación, vamos a relacionar los conceptos anteriores con los tres métodos de correspondencia más conocidos.

La correspondencia por áreas se asocia con la correspondencia entre niveles de gris, es decir se compara la distribución del nivel de gris en pequeñas áreas de la imagen, denominadas teselas o facetas de imagen, y la semejanza se calcula mediante técnicas de correlación o de mínimos cuadrados.

La correspondencia basada en características se usa fundamentalmente en la Visión Computacional. Con objeto de determinar las características homólogas se comparan los bordes u otros elementos derivados de las imágenes originales. La semejanza, por ejemplo en la forma, carácter y fuerza de los bordes se mide mediante una función de costo.

El tercer método, correspondencia simbólica, se refiere a aquellos métodos que comparan descripciones simbólicas de las imágenes y miden la semejanza mediante una función de

Análisis de imágenes basado en correlación de imágenes

costo. Las descripciones simbólicas pueden referirse a niveles de gris, o a características derivadas. Pueden realizarse en forma de grafos, árboles, redes semánticas, por mencionar sólo unas pocas posibilidades. La correspondencia simbólica no se basa estrictamente en las propiedades geométricas de semejanza, en contraste con los otros métodos. Compara las propiedades topológicas, en lugar de utilizar la forma o la posición como criterio de semejanza.

Es importante tener presente que el término correspondencia de imágenes hace mención no sólo a un par de imágenes sino también a la posibilidad de trabajar con múltiples imágenes.

En términos generales, la exposición de un problema de correspondencia de imágenes puede plantearse de la siguiente forma:

1. Seleccionar en una imagen la entidad de correspondencia.
2. Buscar en la otra imagen su entidad conjugada (correspondiente).
3. Calcular la posición de la entidad correspondiente en el espacio objeto.
4. Valorar la calidad de la correspondencia.

Obviamente la segunda fase es la más difícil de resolver y la que propicia un contexto de investigación más intenso a muy diversos niveles.

4.4 Dificultades en la correspondencia de imágenes.

A la hora de la localización automática de puntos homólogos pueden surgir varios problemas:

- Vastedad de la búsqueda

Que se produce cuando se calcula el índice de semejanza entre las entidades de correspondencia para toda la imagen. Por ejemplo, en el caso de imágenes de alta resolución 8 megapíxeles (4000 x 2000 píxeles) el tiempo de computo puede dispararse. De ahí la necesidad de incorporar restricciones geométricas y estrategias que permitan reducir dicho espacio de manera eficiente y robusta.

- Ambigüedad

Que se produce si la entidad de correspondencia no es suficientemente característica (única). La solución pasará por la selección de más entidades de correspondencia únicas.

- Mal condicionamiento

Una imagen se caracteriza por su debilidad geométrica y radiométrica. Esta debilidad se concreta en la posibilidad de que al buscar el punto homólogo exista.

- Una solución inexistente

Puede no haber punto homólogo debido a ocultaciones, desplazamientos por relieve, diferencias radiométricas, etc...

Análisis de imágenes basado en correlación de imágenes

- Una solución múltiple

Puede haber más de una solución (aparentes) en el caso de patrones repetitivos asociadas a cultivos, edificaciones, etc...

- Una solución inestable

La solución puede quedar afectada por condiciones geométricas o radiométricas de manera que en diversas imágenes pueden detectarse puntos homólogos distintos (para un mismo punto de referencia).

Particularmente, Kasser en su estudio de la correspondencia de imágenes identifica la siguiente serie de problemas en la imagen, que debilitan la capacidad de encontrar puntos homólogos:

- Zonas con pobreza radiométrica o patrones repetitivos. En este segundo caso, pueden emplearse máscaras más grandes, no alineadas con la dirección de modulación del patrón repetitivo, o imágenes complementarias.
- Zonas ocultas. Por ejemplo, en zonas urbanas, el punto homólogo puede no ser visible y sin embargo, la máscara de correlación dar un máximo. Es necesario el empleo de más imágenes y / o ampliar el recubrimiento.
- Pérdidas de recubrimiento por relieve o pérdida del caso normal. Pendientes pronunciadas y / o líneas de ruptura que dan lugar a deformaciones de las máscaras de correlación.
- Iluminación diferencial debido a los cambios en la posición relativa de los vectores fuente de iluminación, normal a la superficie, eje de toma. Pueden aparecer zonas saturadas radiométricamente debido a comportamientos especulares de la superficie.
- Objetos móviles, sombras desplazadas en el caso de imágenes de satélite con estereoscopia transversal respecto de la órbita.

La introducción de estas restricciones es un paso importante para convertir la correspondencia de imágenes en un problema bien definido. Las restricciones pueden ser vistas como una información adicional acerca de los parámetros, y en este sentido la información adicional es 'conocimiento' acerca del proceso de correspondencia.

En resumen, el problema de la correspondencia de imágenes es doble: en primer lugar se debe obtener una aproximación lo suficientemente buena y posteriormente descubrir el justo término entre las hipótesis planteadas y las restricciones impuestas a la solución por un lado y mantener la generalidad por otro.

4.5 Distorsiones geométricas de las entidades de correspondencia

Diferencia de escala entre las dos imágenes. Se produce cuando la correlación se lleva a cabo con un par de imágenes con diferente escala. Al proyectarse las dos teselas sobre la superficie plana, los píxeles interiores ya no resultan conjugados puesto que aparecen referidos a distintas posiciones dentro del espacio objeto.

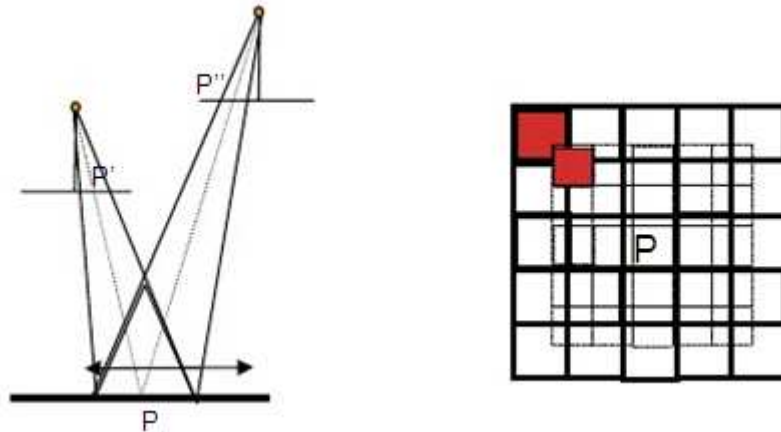


Ilustración 41

Diferencia de ángulos de rotación entre las dos imágenes. Efecto causado por la diferencias de rotación entre las imágenes, debido a que una de las imágenes rota alrededor de un eje y la otra alrededor de otro distinto.

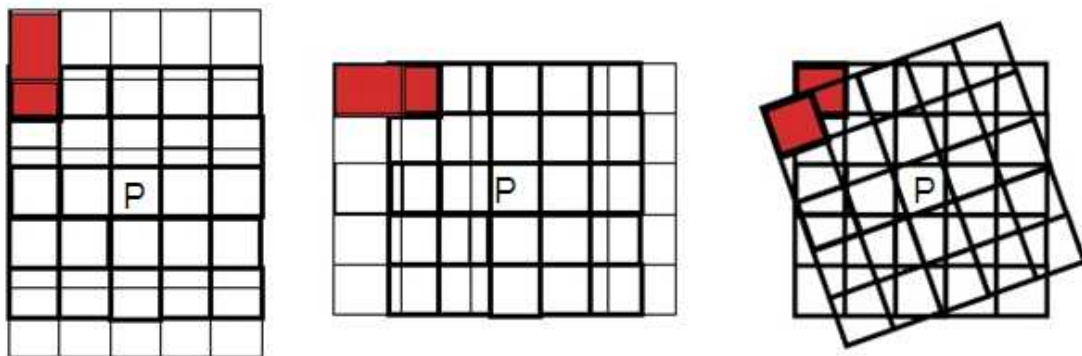


Ilustración 42

Distorsión geométrica causada por la inclinación. Como se observa en la figura una inclinación de la superficie provoca un acortamiento o reducción de la misma. Al inclinarse más la superficie, llega a alcanzarse un ángulo crítico para el que no se ve la imagen derecha, llegándose a la situación de oclusión. Observando la figura se evidencia otra consecuencia producida por la inclinación de la superficie: la resolución decrece en la dirección perpendicular al eje de rotación.

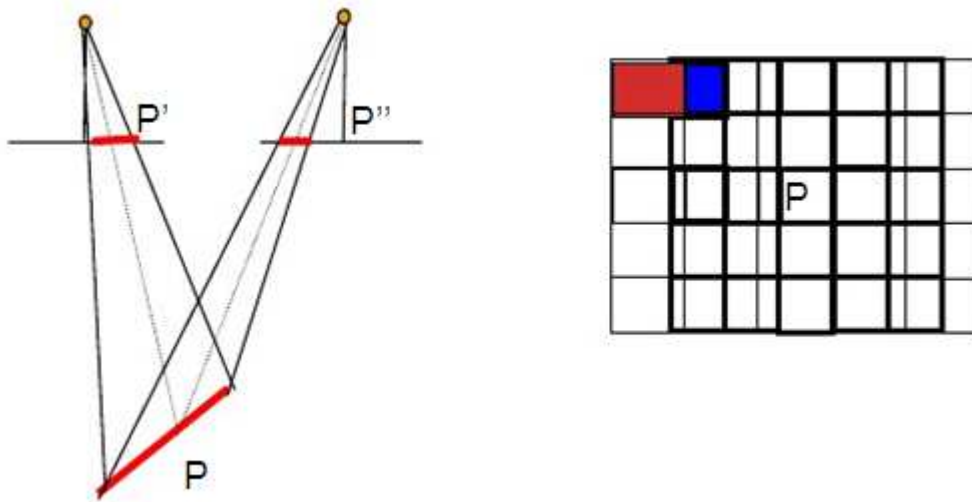


Ilustración 43

Distorsión geométrica debida al relieve. Como se vio antes, las teselas imágenes de posiciones homólogas se proyectan sobre la tesela superficial 'montañosa'. Como podemos observar en la siguiente figura, los píxeles individuales no son conjugados ni siquiera en el caso de un par estereoscópico estrictamente vertical.

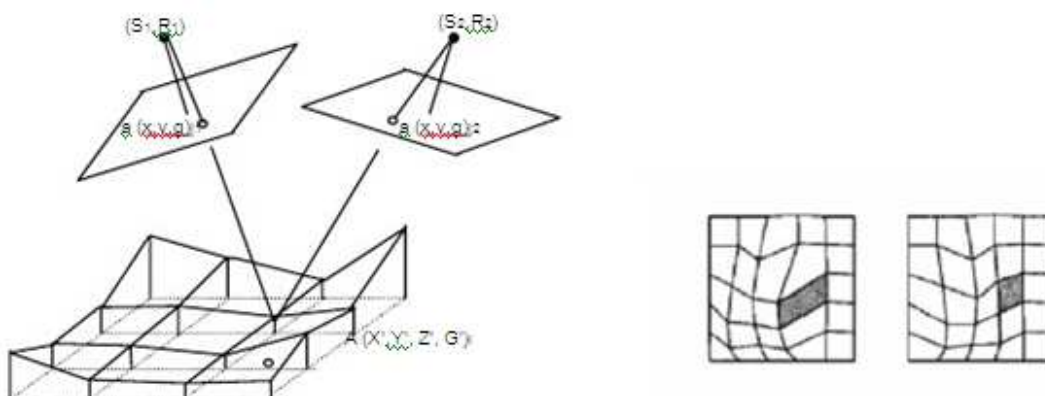


Ilustración 44

Se han analizado el efecto de las distorsiones geométricas sobre las teselas de la imagen que se usan como entidades en la correspondencia por áreas. Generalmente, el método de correspondencia por áreas no se ve afectado por las distorsiones de relieve, sin embargo son muchos más sensibles a la mayoría de distorsiones geométricas. Por ejemplo, la correspondencia de bordes es tremendamente sensible a las distorsiones geométricas. Cualquier variación en la posición del punto de vista provocará que el borde sufra desplazamiento e incluso deformación lo que complicará su correspondencia.

4.6 Estrategias de correspondencia por áreas

En la correspondencia por áreas las entidades son los niveles de gris. La idea de este método es comparar la distribución de niveles de gris de una pequeña subimagen, denominada tesela imagen, con su correspondiente en la otra imagen. A continuación introduciremos la terminología a usar.

Análisis de imágenes basado en correlación de imágenes

Se conoce como patrón o ventana patrón, a la tesela imagen que permanece en una posición fija en una de las imágenes. Se entiende por ventana de búsqueda la zona del espacio de búsqueda dentro de la que se comparan las teselas imagen (llamadas ventanas de correspondencia) con el patrón, comparación que se realiza utilizando distintos criterios para calcular el índice de semejanza. Los dos más conocidos son la correlación cruzada y la correspondencia por mínimos cuadrados. La posición y tamaño de la ventana de búsqueda puede determinarse, acotarse, en función de alguna restricción geométrica.

La correspondencia entre dos píxeles se establece a partir de la correlación (grado de semejanza) de los niveles de gris correspondientes a los píxeles pertenecientes a sendas máscaras (matrices) centradas sobre los píxeles candidatos.

A partir de un píxel que actúa como referencia (dato) en una imagen, cuya máscara no se mueve, se busca en la otra imagen o imágenes, el píxel o píxeles, cuya máscara ofrece el mayor nivel de correlación.

Tiene un carácter geométrico, es decir, no es invariante a la geometría de la máscara. La técnica funciona adecuadamente siempre y cuando exista una regularidad geométrica entre las dos zonas de la imagen.

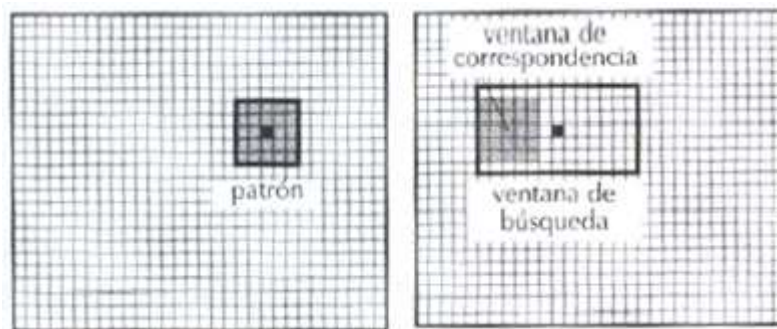


Ilustración 45

Independientemente del cálculo utilizado para el índice de semejanza, a continuación vamos a tratar una serie de conceptos que debemos tener claros.

- Posición del patrón

El aspecto de la elección del centro del patrón puede parecer superficial, sin embargo no lo es. En teoría el centro del patrón podría situarse dentro de un área que fuera la mitad del tamaño del patrón y menor que la imagen, pero un análisis más cuidadoso indicaría que hay que ser más selectivo a la hora de aceptar dichos límites. La correspondencia por áreas puede fallar en determinadas condiciones. Sirvan como ejemplos: la colocación del patrón en zonas que presentan oclusión en la otra imagen, la selección de un área de baja relación señal ruido, la elección de un área con patrones repetidos, la ubicación en un área con líneas de ruptura...

Análisis de imágenes basado en correlación de imágenes

- Tamaño del patrón

El tamaño del patrón y de la ventana de correspondencia son dos parámetros importantes. Al aumentar el tamaño, normalmente aumenta la unicidad de la función de niveles de gris, pero también lo hacen los problemas de distorsión geométrica por lo que se debe encontrarse una solución de compromiso, por ejemplo calculando un índice de unicidad para distintos tamaños del patrón, que sirva para controlar una ubicación útil del mismo.

- Posición y tamaño de la ventana de búsqueda

La posición de la ventana de búsqueda es crucial, ya que la correspondencia por áreas necesita aproximaciones muy buenas. Su tamaño no obstante, no juega un papel importante porque la aproximación limita el tamaño a unos pocos píxeles. Una estrategia de correspondencia basada en el método jerárquico asegura la obtención de buenas aproximaciones.

- Criterios de tolerancia

Deben analizarse los factores que aparecen en la medida de la semejanza, entre el patrón y la ventana de correspondencia. El criterio de aceptación cambia, incluso dentro de la misma imagen, por lo que los valores umbral u otros deberían definirse localmente, umbrales adaptativos.

- Control de calidad

El control de la calidad incluye una valoración de la precisión y de la exactitud de las posiciones homólogas o conjugadas. Además debe estudiarse la consistencia de los puntos correspondientes, incluyendo su adecuación a las expectativas o el conocimiento sobre el espacio objeto.

4.6.1 Correlación

La idea es medir el grado de semejanza del patrón con la ventana de correspondencia calculando el factor de correlación cruzado.

El coeficiente de correlación ρ se define como:

$$\rho = \frac{\sigma_{LR}}{\sigma_L \sigma_R} \quad (\text{Ecuación 193})$$

Si ρ está normalizado se verifica $-1 \leq \rho \leq 1$. En la ecuación anterior los términos se definen como:

σ_{LR} , covarianza de las teselas de imagen L (izquierda) y R (derecha).

σ_L , desviación típica o estándar de la tesela de imagen L (patrón).

σ_R , desviación típica de la tesela de la imagen R (ventana de correspondencia).

Análisis de imágenes basado en correlación de imágenes

Si se introducen las funciones imagen $g_L(x,y)$ $g_R(x,y)$ para las teselas de imagen izquierda y derecha (en este caso el patrón y la ventana de correspondencia) y se notan en la forma de g_L , g_R se obtienen las siguientes ecuaciones de definición:

$$\bar{g}_L = \frac{\sum_{i=1}^n \sum_{j=1}^m g_L(x_i, y_j)}{n \cdot m} \quad (\text{Ecuación 194})$$

$$\bar{g}_R = \frac{\sum_{i=1}^n \sum_{j=1}^m g_R(x_i, y_j)}{n \cdot m} \quad (\text{Ecuación 195})$$

$$\sigma_L = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (g_L(x_i, y_j) - \bar{g}_L)^2}{n \cdot m - 1}} \quad (\text{Ecuación 196})$$

$$\sigma_R = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (g_R(x_i, y_j) - \bar{g}_R)^2}{n \cdot m - 1}} \quad (\text{Ecuación 197})$$

$$\sigma_{LR} = \frac{\sum_{i=1}^n \sum_{j=1}^m ((g_R(x_i, y_j) - \bar{g}_R)^2 (g_L(x_i, y_j) - \bar{g}_L)^2)}{n \cdot m - 1} \quad (\text{Ecuación 198})$$

El factor de correlación cruzado se determina dentro del espacio de búsqueda para cada posición f,c de la ventana de correspondencia. El problema siguiente es determinar qué posición u,v proporciona el máximo factor de correlación.

El factor de correlación máximo cruzado normalizado ρ toma valores en el intervalo [-1, 1]. Se obtiene un valor igual a la unidad si el patrón y la ventana de correlación son idénticos. Si no existe correlación entre las dos teselas de imagen, es decir si no hay coincidencia en absoluto, entonces $\rho=0$. El valor $\rho=-1$ indicaría una correlación inversa, como en el caso de la diapositiva y el negativo de la misma imagen.

Si se reduce la ventana de búsqueda en base a alguna restricción geométrica, los valores de los factores de correlación pueden ser dibujados entonces en un gráfico como el de la ilustración 46. El máximo se halla ajustando un polinomio, por ejemplo una parábola, a los valores de correlación. Se hace notar que el máximo raramente coincide con los valores enteros de la dirección o posición del píxel.

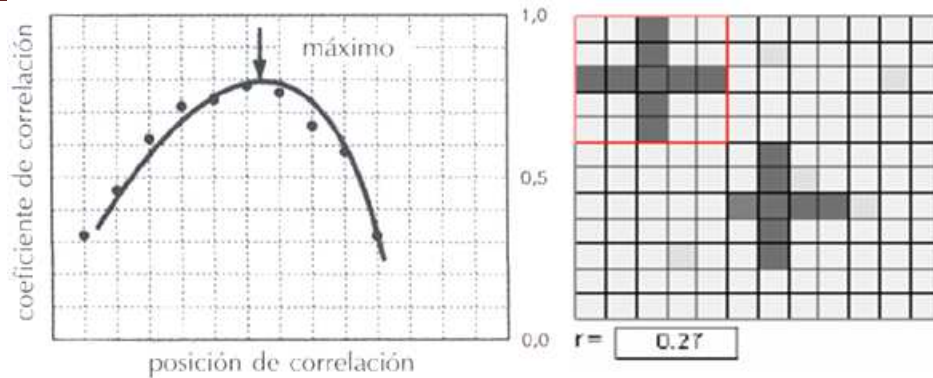


Ilustración 46

Procedimiento

El procedimiento descrito a continuación describe los pasos generales en una correspondencia por áreas con la correlación como índice de semejanza.

1. Seleccionar el centro del patrón de la imagen.
2. Determinar las ubicaciones aproximadas de las posiciones homólogas en la otra imagen.
3. Determinar tanto para el patrón como para la ventana de correlación los tamaños mínimos que cumple con el criterio de unicidad. Elegir el mayor de los dos valores como tamaño de la ventana para la posición actual de correspondencia.
4. Calcular los coeficientes de correlación $p_{f,c}$ con la ecuación anterior para todas las posiciones f,c de la ventana de correlación en la ventana de búsqueda.
5. Analizar los factores de correlación. Debe alcanzarse un valor umbral mínimo para que la correspondencia sea válida. Además del máximo debe determinarse éste de modo inequívoco como un índice de la calidad de medida.
6. Repetir los pasos 2-5 para una nueva posición del patrón hasta que se haya hecho la correspondencia en todas las posiciones.
7. Analizar los resultados de la correspondencia con una visión global para determinar la consistencia y la compatibilidad con el conocimiento a priori de la escena o imagen.

4.6.2 Correspondencia por mínimos cuadrados

Las primeras experiencias de correspondencia por mínimos cuadrados (LSM) las presentaron los investigadores a comienzos de los años ochenta, por ejemplo (Ackermann, 1984); (Grün, 1985); (Grün y Baltsavias, 1987); (Förstner, 1982, 1986); (Rosenholm, 1987); (Thurgood y Mikhail, 1982). La idea en la correspondencia por mínimos cuadrados es minimizar las diferencias en los niveles de gris entre el patrón y la ventana de correspondencia, motivo por el que la posición y la forma de la ventana de correspondencia son parámetros que han de determinarse en el proceso de ajuste. Es decir, la forma y posición de la ventana de correspondencia cambian hasta que las diferencias de nivel de gris entre la ventana deformada y el patrón (constante o fijo) alcancen un mínimo. Si bien el concepto de mover la ventana de correspondencia hasta encontrar la posición conjugada u homóloga es sencillo de entender, la modificación del aspecto puede no ser tan obvia a primera vista. Después de tratar sobre las distintas distorsiones geométricas producidas por el desconocimiento de los parámetros de

Análisis de imágenes basado en correlación de imágenes

orientación, la inclinación de la superficie, el relieve, etc., parece evidente que debe cambiar la forma de la ventana de correspondencia para que todos los píxeles de la misma sean conjugados de los que les corresponden en el patrón.

Supóngase que la ventana de correspondencia está centrada exactamente en la posición conjugada. Las diferencias de niveles de gris entre los píxeles correspondientes del patrón y la ventana de correspondencia se deben a distintos factores, entre los que se detallan:

- Las diferencias de reflectancia e iluminación entre las dos imágenes.
- La propia cámara.
- El proceso de revelado fotográfico y de barrido o escaneado en el caso de fotografías digitalizadas.
- La distorsión geométrica de la teselación regular de la imagen causada por la orientación, la inclinación de la superficie, el relieve.

El proceso habitual sigue los siguientes pasos:

Modelo aleatorio

Expresa que entre la función imagen (su máscara de niveles de gris) de referencia, sobre el píxel (x,y) , y la función imagen candidata (su máscara de niveles de gris) para establecer el punto homólogo sobre el píxel (x,y) , existe una discrepancia residual que sigue una distribución gaussiana y, por tanto, susceptible de minimización.

$$F(\bar{x}, \bar{y}) + v = G(x, y) \quad (\text{Ecuación 199})$$

$$p(v) \sim N(v; 0; \sigma_v) \quad (\text{Ecuación 200})$$

Modelo funcional radiométrico

Expresa que entre las dos funciones existe una variación sistemática reducible generalmente a un modelo lineal: variación de brillo (r_0) y de contraste (r_1).

$$F(\bar{x}, \bar{y}) + v = G(x, y)r_1 - r_0 \quad (\text{Ecuación 201})$$

Modelo funcional geométrico

Expresa que el entorno de los dos puntos de las imágenes (las máscaras) cuyos valores radiométricos se comparan pueden estar relacionados mediante una transformación lineal (afín) y, por tanto, que el valor de la imagen observado $G(x_0, y_0)$ se puede relacionar con el teórico ideal $G(x, y)$ a través de dicha transformación lineal.

$$F(\bar{x}, \bar{y}) + v = G(ax_0 + by_0 + \Delta x + cx_0 + dy_0 + \Delta y)r_1 + r_0 \quad (\text{Ecuación 202})$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_0 & y_0 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_0 & y_0 & 0 & 1 \end{bmatrix} \quad (\text{Ecuación 203})$$

Análisis de imágenes basado en correlación de imágenes

La tarea a realizar consiste en la estimación de los parámetros radiométricos r_0 , r_1 y geométricos (a , b , c , d , Δx , Δy) mediante la minimización de $\sum v^2$. Una vez linealizado el modelo mediante el desarrollo en serie de Taylor y las correspondientes aproximaciones iniciales llegamos a:

$$[F(\bar{x}, \bar{y}) - G(x, y)] + v = \begin{bmatrix} 1 & G_0(x, y) & G'_x x & G'_x y & G'_y x & G'_y y & G'_x & G'_y \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \\ da \\ db \\ dd \\ de \\ \Delta x \\ \Delta y \end{bmatrix}$$

(Ecuación 204)

donde los términos G'_x y G'_y representan las derivadas parciales y discretizadas de la imagen en la dirección X y en la dirección Y, respectivamente. Es decir, los términos G'_x se obtienen restando dos a dos los niveles de gris de dos columnas contiguas de la máscara y los términos G'_y se obtienen restando dos a dos los niveles de gris de dos filas adyacentes contiguas de la máscara. Una máscara de $n \times n$ píxeles permite formar un sistema de $(n-1) \times (n-1)$ ecuaciones.

El modelo puede y suele si suponemos que:

- Se ha llevado a cabo un procesamiento previo de las imágenes de manera que han eliminado las discrepancias radiométricas. Es decir, se han determinado los parámetros r_0 y r_1 .
- Las aproximaciones iniciales son suficientemente buenas o las imágenes suficientemente parecidas entre sí (poco relieve y caso normal) como para suponer que no hay giros ni cambios de escala a nivel local. Es decir supone conocidos los coeficientes de la transformación afín.

$$[F(\bar{x}, \bar{y}) - G(x, y)] + v = \begin{bmatrix} G'_x & G'_y \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad \text{(Ecuación 205)}$$

Como en el caso de la correspondencia por áreas esta estrategia es de carácter geométrico y local. Exige una suficientemente buena aproximación inicial. Da precisiones subpíxel. Es más flexible que la correspondencia por áreas.

Procedimiento de ajuste

La correspondencia por mínimos cuadrados no es un problema de ajuste lineal, y por tanto hay que buscar una solución iterativa. Sin embargo, hay algunas diferencias con el ciclo iterativo habitual. La primera iteración se inicia para una posición aproximada

Análisis de imágenes basado en correlación de imágenes

de la ventana de correspondencia. Con la matriz reseñada anteriormente se calculan los coeficientes de la matriz de diseño del sistema y los parámetros incógnita de la transformación.

Después del muestreo comienza la siguiente iteración y se repite el ciclo hasta que se cumpla el criterio de convergencia. Se utiliza un número máximo de iteraciones como tope en caso de que se produzca una falta de convergencia o, lo que es peor, exista divergencia. En circunstancias normales se satisface antes el criterio de variaciones poco significativas de los parámetros que el del máximo número de iteraciones. El punto homólogo o conjugado del centro del patrón es el centro de la ventana de correspondencia después de haber sido transformada a su forma y posición finales.

El procedimiento de correspondencia por mínimos cuadrados puede resumirse en los siguientes pasos:

1. Seleccionar el centro del patrón en una imagen.
2. Determinar las posiciones aproximadas para la ventana de correspondencia.
3. Determinar el tamaño mínimo del patrón y de la ventana de correspondencia que satisfagan los criterios de unicidad. Seleccionar el mayor de los dos valores como tamaño de la ventana.
4. Iniciar a primera iteración con la ventana de correspondencia en la posición aproximada.
5. Transformar la ventana de correspondencia y determinar los valores de gris para la teselación (remuestreo).
6. Repetir la secuencia de ajuste y remuestreo hasta que se alcance los criterios de convergencia.
7. Valorar la calidad del punto homólogo.
8. Repetir los pasos 1-7 para una nueva posición del patrón.

Como se mencionó anteriormente deben realizarse los ajustes radiométricos como, por ejemplo, una ecualización del histograma, antes del proceso de correspondencia de cara a evitar la presencia de más incógnitas en el modelo matemático.

En resumen, la correspondencia por mínimos cuadrados, es una técnica que pretende, por una parte, contrarrestar las limitaciones geométricas de la técnica de correspondencia por áreas (para lo que concede cierta flexibilidad a la máscara) y, al mismo tiempo, ajustar (mediante el procedimiento mínimo – cuadrático) la correspondencia por ‘debajo’ del píxel. La técnica se basa en la minimización de las discrepancias de los niveles de gris entre la máscara de referencia y la máscara de correlación.

4.6.3 Correspondencia basada en características

La correspondencia basada en características utiliza como entidades homólogas características o elementos obtenidos de los niveles de gris de la imagen original, entre lo que se encuentran puntos, bordes y regiones. Las características que se usan con más frecuencia son los bordes y los puntos (puntos de interés).

Análisis de imágenes basado en correlación de imágenes

La correspondencia basada en características se hizo popular al final de los setenta en la Visión Computacional, cuando se cayó en la cuenta de que la prodigiosa capacidad estereoscópica humana se basa más en buscar bordes o límites conjugados (Grimson, 1981; Horn, 1983).

Una vez estos elementos son extraídos, es necesario asignarles una serie de atributos de posición, de forma, orientación, gradiente, longitud, curvatura, de topología, etc. Finalmente debe compararse el grado de similitud existente entre los atributos de los elementos o de las relaciones mediante funciones de coste. Para ello debe acometerse una tarea de búsqueda a través de los registros estructurados de la imagen.

Estas herramientas presentan un carácter más abstracto (no son función inmediata de la geometría y la radiometría) y global (afectan a toda la imagen). En esta media son más invariantes al relieve y la orientación.

4.6.3.1 Extracción de características

La extracción de características es un proceso monocular, que se realiza de modo independiente en cada una de las imágenes. Aunque en el apartado Análisis de imágenes ya hemos hablado sobre ellas, a continuación realizaremos un breve repaso sobre las técnicas más empleadas en este contexto.

Puntos de interés

La extracción de distintos puntos en una imagen o en una tesela imagen tiene como idea básica la identificación de zonas de gran varianza, áreas que son potencialmente útiles para la correspondencia. El proceso de extracción se realiza mediante un operador de interés, y aquellos puntos que tienen características diferenciadas se denominan puntos de interés.

Moravec (1976) desarrollo un operador de interés que medía las diferencias de una tesela imagen con relación a sus alrededores. Actúa calculando las varianzas de pequeñas áreas y umbralizando las mismas local y globalmente.

El operador de interés de Förstner es invariante con la rotación y ofrece una precisión subpíxel, (Förstner, 1986); (Förstner y Gülch, 1987). Con él pueden identificarse diferentes características o elementos puntuales, tales como esquinas y elementos superpuestos a la imagen original.

Detección de bordes

Existen numerosos operadores de borde, cada uno con sus propias características distintivas, y muy a menudo la elección de uno frente a otro se toma por preferencia personal o por la propia experiencia. La detección de bordes implica la identificación de los píxeles de borde y el agrupamiento de los píxeles de borde origina los llamados contornos de borde.

Los bordes se corresponden con diferencias de luminosidad en las imágenes, que pueden ser bruscas (borde marcado o definido) o pueden aparecer sobre un área mayor (borde suave). En teoría un operador de borde debería ser capaz de identificar bordes suaves y definidos. Como habitualmente los bordes aparecen en cualquier

Análisis de imágenes basado en correlación de imágenes

orientación, se necesita un operador que sea independiente de la dirección. Pero la detección de las diferencias de luminosidad supone conocer o determinar las derivadas o, en el caso de funciones discretas como las imágenes digitales, las diferencias de orden n .

4.6.3.2 Correspondencia a través de puntos de interés

La correspondencia por áreas constituye la manera más fácil de hacer corresponder puntos de interés. La determinación de puntos de interés tuvo como motivación original la búsqueda de teselas de imagen con información suficiente para que la correlación tuviera posibilidades de éxito. Como se vio anteriormente, las teselas de imagen con intensidad uniforme producían una función de correlación plana sin un máximo claro, por lo que las áreas citadas deben evitarse.

El concepto de correspondencia de puntos de interés se muestra en la siguiente figura. En el par estereoscópico aparecen los puntos de interés extraídos mediante un operador de Förstner. La plantilla está centrada en uno de ellos, en este caso en la imagen derecha. Ahora, con el método descrito en la sección anterior se calcula la ventana de búsqueda en la otra imagen, en la que se encuentran varios puntos de interés. Se centra en todos estos puntos la ventana de correspondencia y se calculan los coeficientes de correlación. Se considera normalmente como punto conjugado del centro de la plantilla, a aquel punto al que corresponde mayor factor de correlación. Se repite el procedimiento hasta que se ha centrado la plantilla en todos los puntos de interés de la imagen derecha.

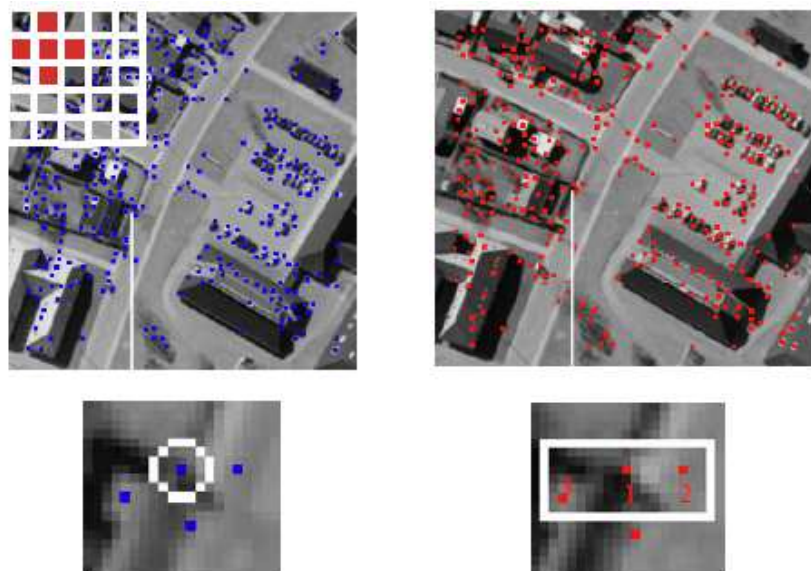


Ilustración 47

4.6.3.3 Correspondencia según puntos de interés

Esta forma de realizar la correspondencia de puntos de interés es una aproximación lógica a la correspondencia por áreas, en la que la plantilla se sitúa en posiciones con

Análisis de imágenes basado en correlación de imágenes

una función de imagen distintiva y no en posiciones arbitrarias. Persisten, no obstante, inconvenientes asociados a las líneas de ruptura, ocultamientos o sombras, acortamientos. Es más, no todos los puntos de interés tienen su punto correspondiente de interés en la otra imagen, como demuestra la misma figura. A pesar de todo, el coeficiente de correlación puede superar el umbral para dar por válida una correspondencia, lo que produciría una actuación incorrecta, y por consiguiente debe realizarse un control posterior, que permita identificar posibles correspondencias erróneas. Como una alternativa a la correlación cruzada, puede aplicarse mínimos cuadrados.

4.6.3.4 Correspondencia de los píxeles de borde

Supóngase que se quiere hacer corresponder el píxel de borde remarcado de la tesela de imagen izquierda de la figura. Lo primero que hay que hacer es determinar la ventana de búsqueda en la tesela de imagen derecha, su centro y dimensiones. Aunque se supone que las imágenes están normalizadas, la búsqueda no puede reducirse estrictamente a la línea bipolar, sino que debe incluir algunas filas por encima y por debajo de la misma.

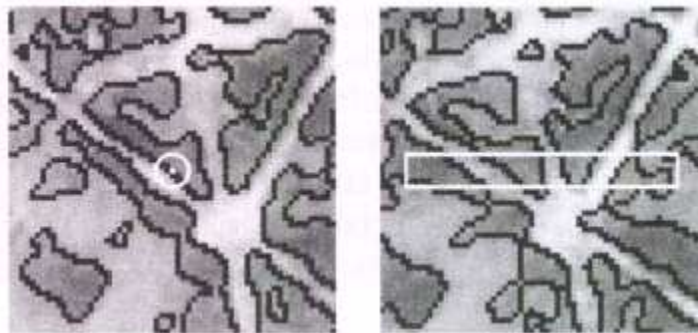


Ilustración 48: Correspondencia de borde. Se selecciona el píxel de borde en el centro del círculo de la imagen izquierda y se predice la ventana de búsqueda (el rectángulo de la imagen derecha). La altura de la ventana de búsqueda es de 3 píxeles por encima y por debajo de la línea epipolar para permitir los desplazamientos verticales. La longitud depende de la incertidumbre de la cota (profundidad) del píxel de borde homólogo.

A mayor número de correspondencias potenciales halladas, la dificultad para determinar la correcta aumenta. Por lo tanto la cuestión básica es cuántas posibilidades se tendrán de media y cómo puede modificarse dicho número para alcanzar un máximo. El número de pasos o cruces por cero depende del tamaño del operador LoG (Laplaciano + Gaussiana).

La selección del píxel homólogo correcto entre un conjunto de correspondencias posibles se rige por el principio de la continuidad a lo largo de las discontinuidades. El borde marca una discontinuidad en los niveles de gris en toda su dimensión. Los bordes pueden corresponderse con límites de objetos reales que normalmente son continuos, o al menos lo son a trozos. Por tanto la paralaje a lo largo de los bordes correspondientes cambia de forma gradual y en muy pocas ocasiones bruscamente. Una posible forma de trabajar es calcular todas las paralajes posibles entre el píxel que

Análisis de imágenes basado en correlación de imágenes

se quiere hacer corresponder y las correspondencias posibles. Estudiando el histograma de todas las paralajes posibles puede encontrarse fácilmente la paralaje predominante. La presencia de más de un máximo indicaría una discontinuidad a lo largo del borde. Si en este método se incluye un análisis de tendencias del histograma se logra una mejora sobre lo anterior que podrían reflejar cambios graduales de la paralaje.

Por último, otros atributos de borde adicionales, que no se han considerado en la primer correspondencia, pueden ayudar cuando se presentan asociaciones ambiguas que el análisis del histograma es incapaz de resolver. Como por ejemplo el valor del atributo fuerza del borde. Es improbable que un borde que aparece nítido o marcado en una imagen esté borroso en la otra.

Por el mismo motivo, es bastante improbable que un borde sea el conjugado de varios. Por lo tanto, la asociación de píxeles del mismo borde favorecería más la resolución de las ambigüedades.

4.6.3.5 Parametrización de bordes. Método Ψ -s

El método Ψ -s constituye un acercamiento a la correspondencia de bordes en base a una parametrización en el espacio de parámetros Ψ -s, siendo Ψ la orientación del borde a través de la derivada función gradiente y s su longitud. El resultado es invariante respecto a la posición del borde y conlleva una pequeña traslación con respecto a la rotación del borde.

En la ilustración 49 aparecen ejemplos de funciones en la representación cartesiana tradicional x,y así como en la representación Ψ -s. A las líneas rectas en el dominio espacial le corresponden líneas rectas horizontales en el dominio Ψ -s (la derivada es constante), El valor Ψ de las líneas horizontales es proporcional a la orientación de la línea recta, como se ve en (a). Los círculos se representan por líneas rectas de pendiente proporcional a la curvatura como ilustra en (c). Una rotación en el dominio espacial supone un desplazamiento vertical en la representación Ψ -s. (b) es una versión rotada de (a). Análogamente (d) es una versión rotada de (c).

Los ejemplos representados en la figura aclaran las ventajas de la representación Ψ -s en la correspondencia de contornos de borde. Por un lado la representación es invariante con respecto a la posición del borde en la imagen. Una rotación entre las dos imágenes supone una simple traslación o desplazamiento de la posición cero. Además se reduce en uno el grado de la línea original.

Los bordes se representan como una secuencia de píxeles de borde, por ejemplo unos códigos en cadena. El comienzo del borde viene dado por las coordenadas imagen, pero los siguientes píxeles están relacionados con el anterior por la dirección. El código en cadena es la primera diferencia de la representación lineal discreta (correspondiente a la primera derivada de la representación continua). Por lo tanto la representación discreta Ψ -s es en esencia una representación de un código en cadena.

Se usa la representación Ψ -s en la correspondencia de formas por su capacidad para la extracción de elementos o características formales distintas. Un ejemplo de

Análisis de imágenes basado en correlación de imágenes

característica de forma distinta es el cambio de curvatura. En términos de formas, las líneas rectas son francamente aburridas, mientras que las líneas con cambios de curvatura frecuentes ofrecen una gran riqueza de información sobre la forma. ¿Cómo puede explicitarse dicha información para explicitarla en la correspondencia de bordes?

Supóngase dividida o segmentada la curva Ψ -s en una sucesión de líneas rectas. Como los segmentos de líneas horizontales se corresponden con líneas rectas en el dominio x,y con una orientación igual a Ψ y las líneas rectas con pendiente son arcos circulares, la división o segmentación de las curvas Ψ -s supone una segmentación curvilínea en el dominio espacial. Del análisis de la curva Ψ -s segmentada pueden deducirse o extraerse características o elementos de forma. Los vértices indican los cambios de curvatura y los bordes con formas similares se caracterizan por vértices similares.

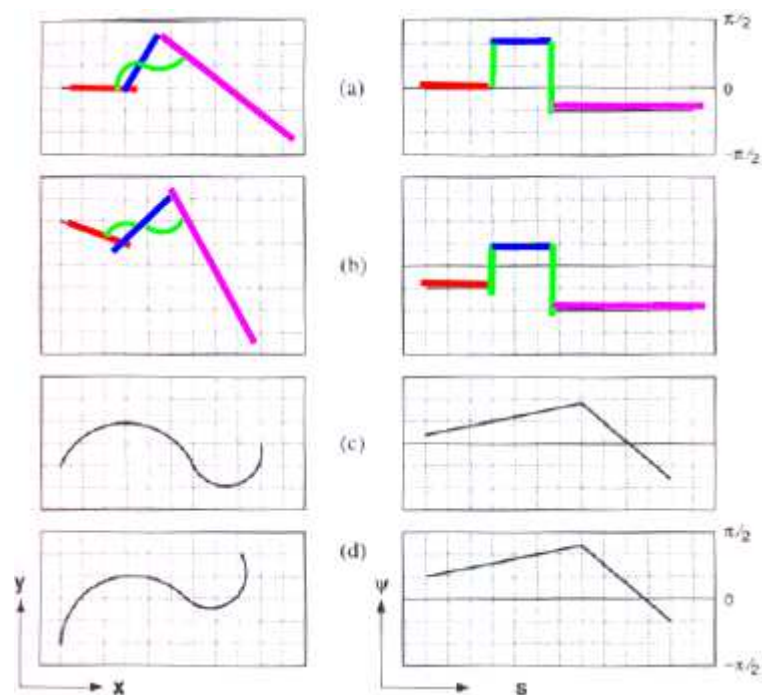


Ilustración 49. Ejemplo de líneas en el dominio x y junto con su representación en el dominio Ψ -s.

4.6.4 Correspondencia relacional

El método Ψ -s de correspondencia tratado con anterioridad tiene una característica común: entidades de correspondencia individuales se comparan con entidades en la otra imagen de una en una. Por ejemplo, se selecciona un contorno de borde en una imagen y se compara con los bordes existentes en la ventana de búsqueda de la otra imagen, buscando una semejanza en la forma. Se repite el proceso de moda más o menos independiente para todos los bordes. Una vez hecha de esta forma la correspondencia de todas las entidades, se selecciona la correspondencia más probable de entre todas las posibles mediante un procedimiento separado que asegura la consistencia local y global partiendo de supuestos a priori sobre la superficie en el espacio objeto.

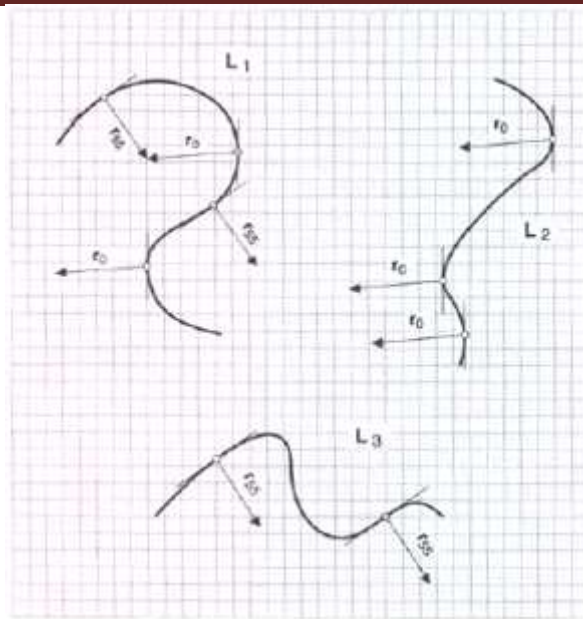


Ilustración 50

La incorporación de las relaciones entre entidades mejora el procedimiento de correspondencia, con lo que podemos hablar de correspondencia de características relacionales, correspondencia por áreas relacional o bien simplemente correspondencia relacional.

La correspondencia relacional proporciona un mecanismo para considerar las relaciones entre las entidades de correspondencia. Puede definirse como la búsqueda de la mejor representación entre dos descripciones relacionales (Vosselman, 1982). Este esquema de correspondencia más general ha sido desarrollado por investigadores en Visión Computacional, por ejemplo (Shapiro y Haralick, 1987) y (Boyer y Kalk, 1998) presentándose en forma de etiquetado o clasificación consistente, relajación y correspondencia estructural. En aplicaciones fotogramétricas se ha usado con éxito en los últimos tiempos la correspondencia relacional, (Vosselman, 1994), (Zilberstein, 1992), (Cho, 1996) y (Wang, 1996).

En la correspondencia relacional deben considerarse tres aspectos igualmente importantes: la descripción, la función de evaluación y el árbol de búsqueda.

4.6.4.1 Descripción de las primitivas y relaciones

Los elementos básicos de la descripción relacional se denominan primitivas y se caracterizan por sus atributos. Prácticamente cualquier puede ser una primitiva. En la aplicación de correspondencia de imágenes, las primitivas son las entidades de correspondencia. Tómese el ejemplo de los bordes. En la siguiente figura hay dos bordes dibujados. Supóngase que se dividen los bordes en partes curvilíneas, por ejemplo con el método Ψ -s descrito anteriormente. La longitud de n segmento y su curvatura son atributos que describen adecuadamente la primitiva. Se tiene:

$$\rho_i = \{centroide(x_i, y_i), longitud(l_i), curvatura(c_i), arco(\alpha_i)\} \quad (\text{Ecuación 206})$$

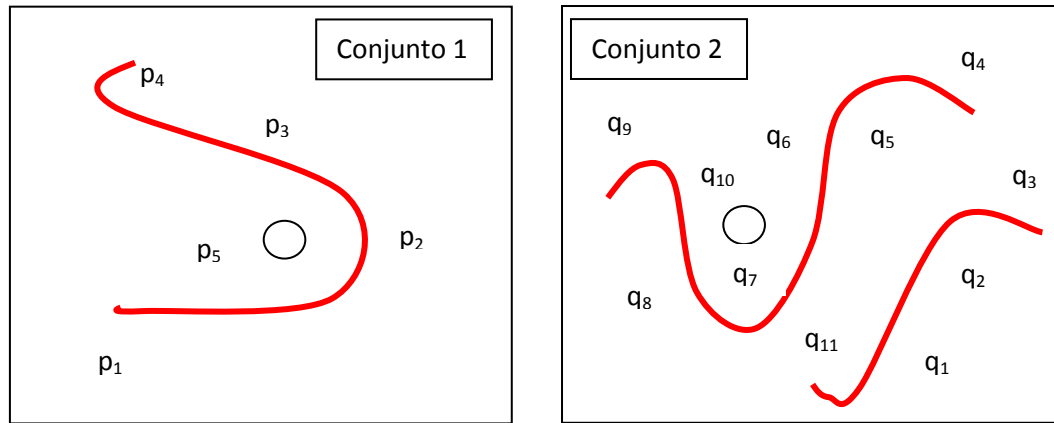


Ilustración 51: Ejemplo de correspondencia relacional. La figura de la izquierda tiene dos contornos de borde segmentados (conjunto 1). La figura de la derecha tiene tres contornos (conjunto 2). La mejor representación entre los dos conjuntos se encuentra mediante un árbol de búsqueda.

Además de la longitud, l , la posición del centroide (x_i, y_i) y el ángulo α_i , correspondiente a la longitud del arco, puede añadirse información adicional, como por ejemplo la fuerza o firmeza del borde, el signo del mismo, incluso quizá información espectral, añadiendo un conjunto de atributos más enriquecido, en función del esquema utilizado en la detección del borde. Sea un borde caracterizado como fuerte, un círculo cerrado de radio conocido, rojo y positivo. Pueden encontrarse cientos de bordes para hacerlos corresponder, de los cuales sólo algunos son cerrados, unos pocos tienen un radio pequeño y muchos son rojos. A mayor riqueza de la descripción, mayor facilidad en la correspondencia. Debería señalarse que la correspondencia puede realizarse solamente con la descripción de las primitivas. Los atributos deberían guardarse en un vector, denominado vector de características, que es precisamente lo descrito en la sección anterior bajo el nombre de correspondencia basada en características.

Retómese de nuevo el problema de la introducción de relaciones entre las primitivas. En el sencillo ejemplo dibujado en la figura anterior la conexión es una relación útil, que especifica simplemente qué segmentos están unidos o conectados. Por ejemplo p_2 está unido a p_1 ; p_5 no lo está, etc. Un ejemplo de interrelación entre bordes podría referirse a la posición de p_5 con respecto a las primitivas que forman el otro borde, que puede expresarse de distintas formas. Al tratar un borde hay un sentido que es la izquierda o la derecha del borde. Si se empieza con p_1 y se sigue hasta p_4 (basándose en la conectividad), la primitiva p_5 siempre estaría a la izquierda. Podría introducirse también una distancia, por ejemplo entre centroides. Para especificar un poco más, la relación debe restringirse sólo a los vecinos próximos. Se termina así con la potente relación de vecino.

Las relaciones se describen con el nombre de la relación un juego de tuplas de relación:

$$(conexión)(\{p_2, p_1\}, \{p_3, p_2\}, \{p_4, p_3\}) \quad (\text{Ecuación 207})$$

$$(conexión)(\{p_5, p_1\}, \{p_5, p_2\}, \{p_5, p_3\}) \quad (\text{Ecuación 208})$$

La relación de vecino puede tener atributos. La distancia y la posición (izquierda, derecha) son atributos muy útiles, como se comentó antes. Pueden combinarse añadiendo un signo a la distancia. Una distancia negativa podría significar, por ejemplo, que la primitiva está a la izquierda de la primitiva (con la que está conectada).

$$(vecino)\{p_5, p_1\}, (distancia - 18) \quad (\text{Ecuación 209})$$

$$\{p_5, p_2\}, (distancia - 12) \quad (\text{Ecuación 210})$$

4.6.4.2 Función de evaluación

Se entiende por correspondencia de imágenes la comparación entre dos imágenes y la determinación de una serie de características o detalles correspondientes, denominados entidades. En la correspondencia relacional las entidades son las descripciones de las primitivas y las relaciones entre ellas. La comparación entre este tipo de relaciones se lleva a cabo mediante una función de evaluación. Por lo general se supone que los atributos y las tuplas relacionales son independientes unas de otras. En ese caso la medida de la semejanza entre dos descripciones relacionales se determina sumando un número (costo), que expresa el grado de falta de semejanza entre las descripciones.

La función de costo es cero si las dos descripciones relacionales son idénticas. Como esta situación raramente se presenta, hay que trabajar con la medida de la falta de semejanza. La función de costo se estima bien por una medida de distancia o bien por un método probabilístico. Como ejemplo de medida de distancia puede tomarse la diferencia absoluta entre los atributos, tanto para las primitivas como para las tuplas relacionales. Sin embargo este método sencillo tiene problemas. En primer lugar, algunos atributos como el color son simbólicos. ¿Cómo se determina la función de costo entre el rojo y el púrpura? En segundo término los atributos tienen diferentes unidades, por ejemplo ángulos, píxeles o milímetros cuadrados. El último problema puede resolverse normalizando los valores de los atributos de modo que todos varíen en el mismo rango de [0, 1].

Como recomiendan (Boyer y Kak, 1988), es preferible el método probabilístico cuando se trabaja con atributos de naturaleza simbólica, como el color, la forma (recta, curva) y la topología (arriba, izquierda). En este caso, la función de probabilidad condicional mide la correspondencia entre las primitivas y relaciones de dos descripciones relacionales. Cuánto más parecidas sean las dos descripciones, más próxima a la unidad estará la función de probabilidad. Realmente lo que se computa como medida del costo es el logaritmo negativo de la función de probabilidad condicional. Vosselman amplía más este método al proponer una función de valor que estima el apoyo que proporcionan a la representación los atributos de las correspondiente primitivas y tuplas relacionales (Vosselman, 1982).

4.6.4.3 Árbol de búsqueda

Después del establecimiento de la descripción relacional de las entidades de correspondencia y la forma de medir el grado de semejanza, es necesario ejecutar un esquema de correspondencia. La forma más común de encontrar la solución es mediante un árbol de búsqueda. En la Visión Computacional e Inteligencia Artificial, se utilizan con éxito tres métodos de búsqueda (Bender, 1996).

Los árboles están formados por nodos y arcos, que comienzan en un nodo raíz, y descienden a través de los nodos antecesores, hasta los nodos hoja. La conexión entre dos nodos se realiza a través de arcos. Las primitivas $\{p_1, p_2, \dots, p_n\}$ de una descripción relacional se denominan unidades y las primitivas de la descripción que se quiere hacer corresponder, es decir $\{q_1, q_2, \dots, q_n\}$ se llaman etiquetas. La dimensión o profundidad del árbol se define por el máximo camino con el menor curso.

El principio se ilustra en la siguiente figura. Las unidades son las primitivas de los dos contornos de borde de la figura anterior y las etiquetas son las primitivas del segundo conjunto. La dimensión del árbol es cinco. Se comienza expandiendo el árbol en el primer nivel asociando unidades con etiquetas compatibles. Es decir p_1 se compara con $\{q_1, q_2, \dots, q_n\}$. Resulta que q_3, q_4, q_8 tienen atributos similares y por lo tanto estas tres etiquetas se convierten en los nodos del primer nivel. Los tres nodos se expanden en el segundo nivel mediante la unidad de correspondencia p_2 . Actuando de esta manera a través de todas las unidades se obtiene un árbol con 120 nodos hoja que representa el espacio de búsqueda de este problema de correspondencia, $P \geq Q$.

Cada camino desde la raíz a una hoja constituye una solución. Obviamente la mayoría de las 120 soluciones son incorrectas geométricamente. ¿Cómo se determina la solución más probable? Supóngase que se determina el costo para cada correspondencia unidad/etiqueta basándose en las diferencias entre atributos. La suma de los costes a lo largo de un camino da el coste total. El camino con coste mínimo representa la solución óptima.

Hasta aquí sólo se ha utilizado la descripción de las primitivas y la correspondencia, ya descrita, es simplemente correspondencia basada en características, realizada mediante el uso de un árbol de búsqueda (en oposición por ejemplo al método Ψ -s). El siguiente paso considera las relaciones entre las primitivas. Volviendo al primer nivel debería considerarse cualquier relación binaria (p_1, p_2) . Existe la relación conexión de p_2 con p_1 y se debería permitir añadir al árbol las etiquetas que tuvieran la misma relación. En otras palabras, de las cinco etiquetas originales que se corresponden con p_2 , sólo una satisface la relación de conexión y como resultado sólo se expande un nodo.

Considerando la relación conexión en todos los nodos, se obtienen tres soluciones como muestra la figura anterior. El costo de la solución $q_4, q_5, q_6, q_7, q_{10}$ será el mayor porque difieren más los atributos entre unidades y etiquetas que en las otras

Análisis de imágenes basado en correlación de imágenes

posibilidades, lo que resulta evidente comparando los contornos de la figura, donde la diferencia de forma entre los dos contornos es obvia.

¿Cuál de las dos soluciones restantes es más probable? Como los costos son similares, la respuesta concluyente se obtiene usando más información adicional, utilizando la relación vecino que aún no se ha empleado. La etiqueta q_{10} , emparejada con la unidad p_5 , sólo puede asociarse al camino $q_8, q_7, q_6, q_5, q_{10}$ debido a la relación vecino de p_5 con p_1, p_2, p_3 . Este ejemplo tan simple demuestra claramente que puede lograrse una correspondencia única $P \geq Q$ introduciendo las relaciones adecuadas.

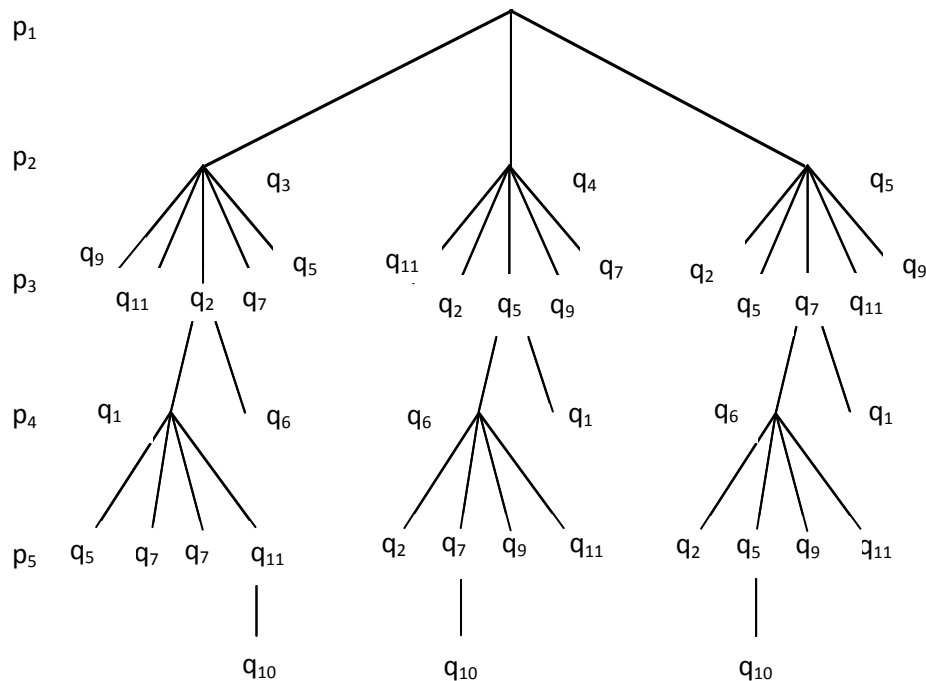


Ilustración 52

Ilustración 52: El árbol corresponde al ejemplo dibujado en la figura anterior. El conjunto de primitivas p_1, \dots, p_5 (unidades) se compara con el conjunto q_1, \dots, q_{11} (etiquetas). Se aceptan las correspondencias en función de las similitudes entre atributos. Únicamente satisfacen las relaciones entre las unidades aquellos nodos que se expanden posteriormente. Por ejemplo p_1 y p_2 tienen una relación de conexión. De todas las correspondencias posibles para p_2 , sólo q_2 en la rama izquierda, satisface la relación con q_3 . Igualmente en el camino central, sólo puede expandirse q_5 .

5. Procedimiento Experimental

Una vez conocida la teoría necesaria para la realización de nuestro proyecto, vamos a explicar brevemente el objetivo y desarrollo de este. El objetivo principal es la elaboración de un programa que nos permita de forma sencilla evaluar y encontrar diferencias existentes entre dos imágenes utilizando el método de correlación de imágenes. El programa utilizado para la realización de nuestro proyecto es el MATLAB.

MATLAB es un lenguaje de alto funcionamiento que integra tanto computación como visualización y programación. Debido a su entorno fácil de usar, los problemas y soluciones se expresan con la notación matemática más sencilla posible. Entre los usos de este programa se encuentran los siguientes:

- Matemática y computación.
- Desarrollo de algoritmos.
- Análisis de datos, exploración y visualización.
- Desarrollo de aplicaciones.

Como elemento básico, MATLAB utiliza la matriz, ya que tiene una característica fundamental, que es que no necesita dimensionamiento. A causa de esto, resuelve problemas en una fracción de tiempo similar al que se gastaría escribiendo un programa en cualquier tipo de lenguaje de programación.

Su nombre proviene de Matriz Laboratorio o Laboratorio de Matrices. Se ha desarrollado durante muchos años debido a las aportaciones realizadas por muchos usuarios en los entornos universitarios, ya que MATLAB es la herramienta estándar para cursos avanzados e introductorios en matemáticas, ingeniería y ciencia. En cuanto al ámbito profesional, MATLAB es utilizado en la industria como herramienta de alta productividad, desarrollo y análisis.

El sistema MATLAB consiste en cinco partes principales:

- Entorno de desarrollo

Se trata del conjunto de herramientas que permiten usar las funciones y archivos del programa. Algunas de estas herramientas son interfaces gráficas de usuario que incluyen desde el escritorio de MATLAB o la ventana de comandos a algunas librerías, que van desde funciones elementales como la suma, hasta funciones más complejas como la inversión de matrices.

- El lenguaje MATLAB

Es un lenguaje de alto nivel que permite la programación a pequeña escala para la creación rápida de programas, hasta programación a gran escala para la realización de aplicaciones complejas. Se trata de un lenguaje intérprete de comandos, es decir, es capaz de procesar de un modo secuencial una serie de comandos previamente definidos, obteniendo el resultado inmediatamente.

- Gráficas.

Permite la visualización de gráficas y vectores así como la posibilidad de incluir leyendas en ellas, explicando y comentando lo que ocurre en estas. Su visualización es tanto en 2 como en 3 dimensiones.

- Interfaces Externas.

Se tratan de un conjunto de librerías que permiten que un programa en el lenguaje MATLAB, interactúe con programas escritos en C o FORTRAN.

- Ayuda para MATLAB

El programa dispone de una gran documentación tanto impresa como en línea que permite conocer todas las características del programa. La ayuda online se obtiene en la barra de herramientas, pulsando en la ventana "Ayuda de MATLAB", o a través del comando help, introduciéndolo en la ventana de comandos.

Una vez conocidos los puntos principales del lenguaje y del programa a utilizar, vamos a explicar a continuación el funcionamiento de nuestro programa:

En primer lugar, introducimos las dos imágenes que vamos a comparar. Una de ellas la utilizaremos como Imagen Patrón, mientras que la otra será la Imagen de Búsqueda. Como hemos comentado anteriormente la estrategia de cálculo de correspondencias que utiliza nuestra aplicación es la correspondencia por áreas. A continuación comentaremos las transformaciones y cálculos que realiza nuestra aplicación.

Una vez introducidas las dos imágenes, el programa divide la Imagen Patrón en facetas. Debido a que las imágenes introducidas pueden tener distintas dimensiones, tanto pares como impares, utilizamos facetas de distintas dimensiones, con el objetivo de perder el mínimo de información posible. Por este motivo, disponemos de 5 facetas, (2,2), (3,3), (4,4), (5,5) y (7,7). Para que el usuario elija la faceta que más le convengan y en la cual pierda menos información, una vez elegidas las dos imágenes, se muestra en pantalla unos mensajes diciéndonos cuantas columnas y filas no evalúa nuestra aplicación utilizando una faceta u otra.

```
Disponemos de las matrices de (2,2), de (3,3), de (4,4), de (5,5) y de (7,7)
Si utilizamos la matriz(2,2) perdemos 0.000000 columnas y 0.000000 filas
Si utilizamos la matriz(3,3), perdemos 0.000000 columnas y 0.000000 filas
Si utilizamos la matriz(4,4), perdemos 2.000000 columnas y 2.000000 filas
Si utilizamos la matriz(5,5) perdemos 1.000000 columnas y 1.000000 filas
Si utilizamos la matriz(7,7), perdemos 6.000000 columnas y 6.000000 filas
```

Ilustración 53. Mensajes de aviso de pérdida de información

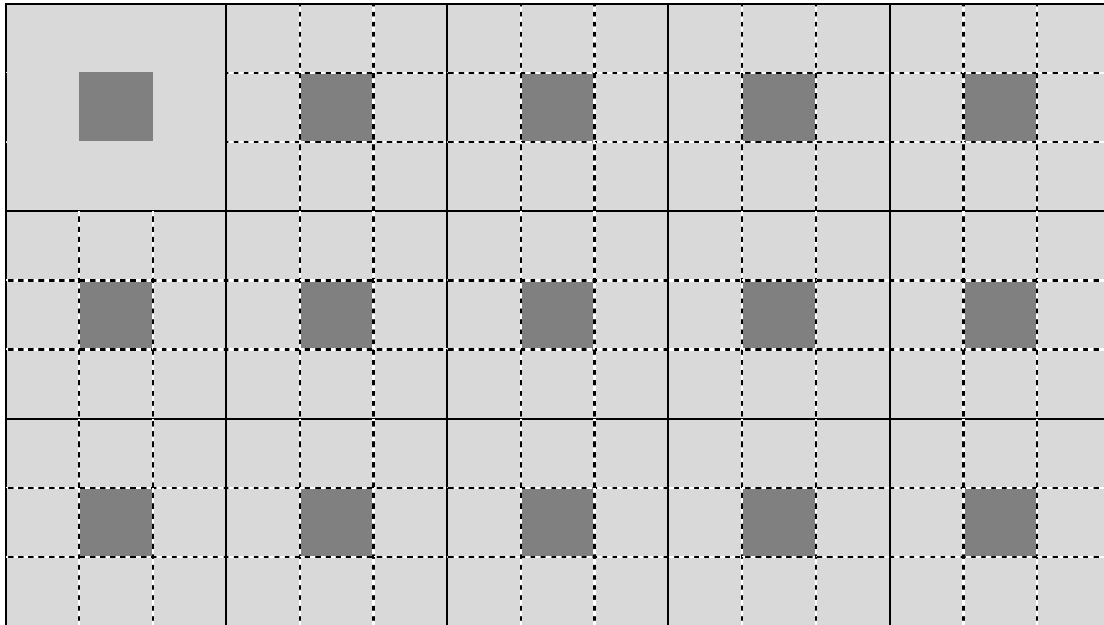


Ilustración 54. Imagen dividida en facetas (Seleccionada primera faceta, ej. Faceta(3,3))

Una vez obtenida la primera faceta patrón, la aplicación almacena de ella, la media aritmética de los píxeles que forman la faceta y la posición exacta del píxel central (o ancla), que será diferente dependiendo de si utilizamos una faceta par o una faceta impar. En las pares, el píxel central se sitúa en la parte inferior derecha de la faceta, mientras que en las impares, el píxel central se sitúa en el medio de la faceta.

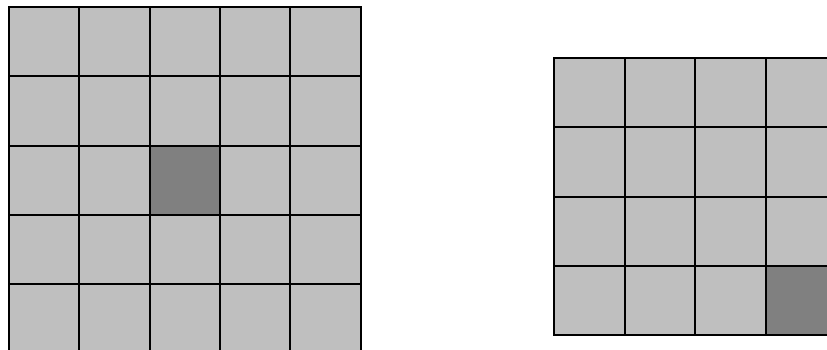


Ilustración 55. Facetas (5,5) y (4,4) con su respectivos anclas

Una vez obtenida la faceta patrón, tenemos que buscar en la imagen de búsqueda la faceta correspondiente. La búsqueda de esta faceta no se hace en la imagen completa, sino que se evalúa en un área cercana a la posición de la faceta patrón. Concretamente para obtener la faceta de búsqueda, cogemos las coordenadas de la faceta patrón y determinamos un área que corresponde a una faceta por encima de la faceta patrón, una faceta por debajo y una faceta tanto por la parte izquierda como por la parte derecha de la faceta patrón.

Análisis de imágenes basado en correlación de imágenes

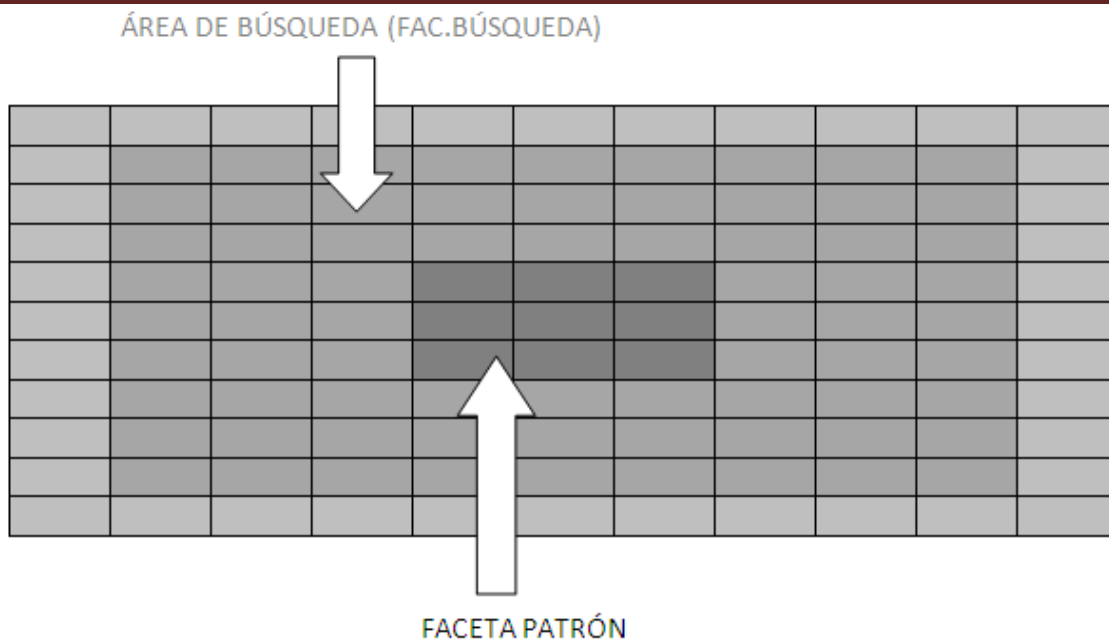


Ilustración 56. Ejemplo de la obtención del área de búsqueda (Faceta 3x3).

Una vez determinado el área, se recorre este píxel por píxel y por cada faceta de búsqueda se calcula el factor de correlación correspondiente con su faceta patrón. El cálculo de este factor se calcula utilizando la siguiente fórmula:

$$\rho = \frac{\sigma_{LR}}{\sigma_L \sigma_R} \quad (\text{Ecuación 211})$$

El área de búsqueda se recorre entero, y solo se guarda en la matriz el valor de la faceta que tenga mayor correspondencia con la faceta patrón, es decir, que se guarda el valor máximo.

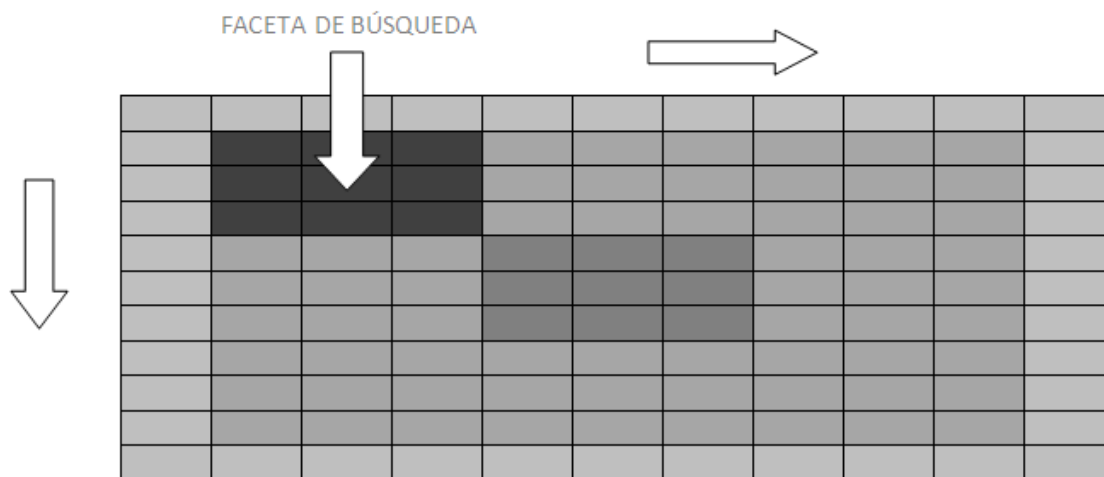


Ilustración 57. Faceta de búsqueda recorriendo el área de búsqueda (Faceta 3x3).

Una vez recorrida toda el área de búsqueda, se evalúa la siguiente faceta patrón y se vuelve a repetir el mismo proceso. El programa termina cuando la Imagen Patrón se recorre por completo y obtenemos una matriz con los valores máximos de correlación concretos de cada faceta. Esta matriz no tiene el mismo tamaño de las imágenes sino que depende del tamaño

Análisis de imágenes basado en correlación de imágenes

de las facetas elegidas y de las dimensiones de la Imagen Patrón. Por ejemplo, si disponemos de una Imagen Patrón de 80x80 píxeles y elegimos una faceta de (4,4), la matriz resultante tendrá unas dimensiones de 20x20. En la posición (1,1), estará el valor de la primera faceta patrón, en la posición (1,2) el de la segunda y así sucesivamente.

Con la información guardada, es decir, con la matriz de correlación obtenida y con la posición relativa de los píxeles centrales en cada faceta, nuestra aplicación procederá a la presentación de resultados de dos formas diferentes.

En la primera representación obtenemos una imagen con un tamaño reducido, dependiendo de la faceta utilizada, donde se muestran como los píxeles se han desplazado mediante líneas que unen las coordenadas del píxel central de la faceta patrón con las coordenadas del píxel central de su correspondencia.

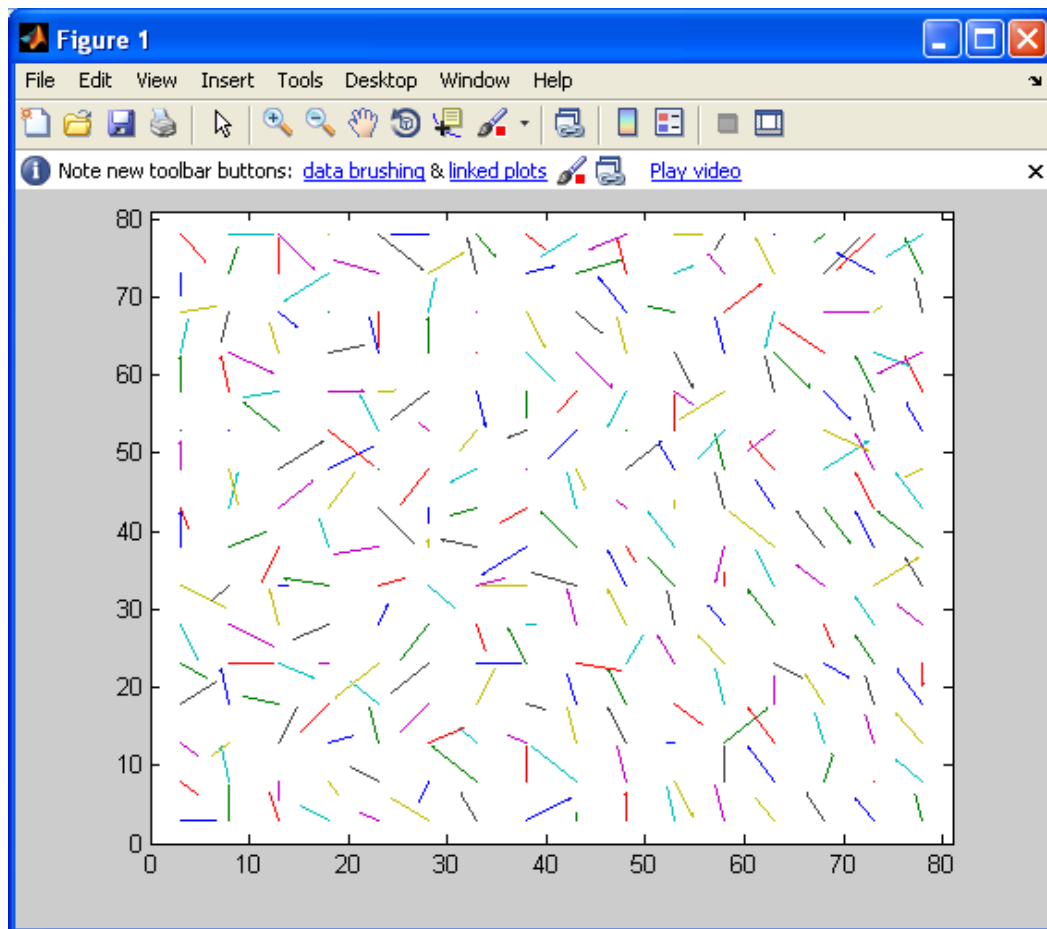


Ilustración 58. Vectores de movimiento de una imagen de 80x80.

La segunda representación, al igual que en la primera el tamaño es reducido con respecto a las imágenes introducidas al principio de la aplicación. Esta representación muestra una imagen con diferentes colores, asignando a cada faceta un color según un porcentaje de desplazamiento con respecto a un desplazamiento máximo, es decir, muestra diferentes colores por faceta patrón dependiendo si el valor del factor de correlación es grande o pequeño. Este grafico tiene 4 variantes según el factor obtenido. En la siguiente tabla aparecen los correspondientes colores según su factor de correlación.

FC (Factor de Correlación)	$-1 < FC < -0.8$	$-0.8 < FC < 0$	$0 < FC < 0.8$	$0.8 < FC < 1$
Color	Azul Oscuro	Azul Cyan	Amarillo	Rojo

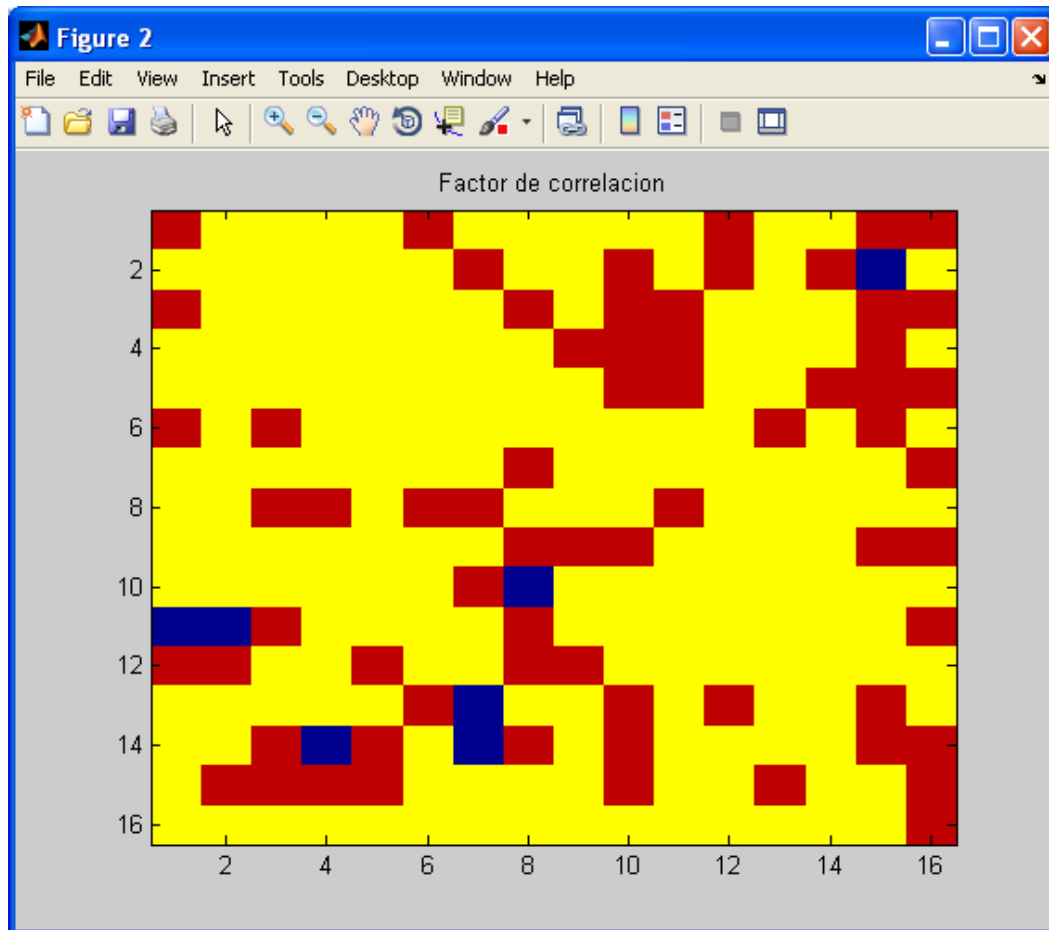


Ilustración 58. Facetas con su color correspondiente debido a su FC (Imagen 80x80).

6. Conclusiones

Para concluir quiero exponer las principales conclusiones que he obtenido durante la realización del proyecto.

Como primer punto importante me gustaría destacar la constante evolución que ha tenido el tratamiento digital de imágenes, debido en gran parte a la mejora de los dispositivos de captación y la velocidad de los procesadores.

El estudio realizado sobre el análisis de imágenes me ha aportado gran cantidad de información que desconocía, ya que el estudio por mi parte de esta materia había sido muy superficial y sin adentrarme en profundidad en el tema. Además, he podido observar la cantidad de campos en los que es útil el tratamiento de imágenes como es la industria, la biomedicina o la seguridad.

En cuanto al desarrollo de la aplicación, me ha servido de gran ayuda las prácticas en asignaturas como TDI (Tratamiento Digital de la Imagen) o sistemas de Televisión, ya que en estas ya me había familiarizado con el uso del programa MATLAB. Aún así he de afirmar que he tenido numerosos problemas ya que aunque había trabajado con el programa el desconocimiento de muchos comandos y la sintaxis tan peculiar de MATLAB, me han ocasionado algún problema.

Sin embargo, aunque la misión del proyecto se ha conseguido satisfactoriamente, me gustaría dar a conocer una serie de problemas que me han surgido y que podrían mejorarse en un futuro. El principal problema que veo a esta aplicación, es el rendimiento que tiene. Esto se debe en gran parte a que MATLAB es un programa interpretado, es decir que la ejecución del programa se traduce o interpreta comando por comando en tiempo de ejecución a un lenguaje intermedio o lenguaje máquina o a través de una máquina virtual, por lo que su funcionamiento es bastante tedioso y hablando de rapidez, es lento en comparación con un programa compilado. A continuación muestro una tabla resumen, en la que se ve la duración que tiene el programa para imágenes con diferentes dimensiones:

Tamaño de la Imagen	40 x 40 píxeles	80 x 80 píxeles	100 x 100 píxeles	128 x 128 píxeles
Tiempo cálculo	81 segundos	1365 segundos	3379 segundos	9508 segundos
Factor de Correlación	1,35 minutos	22,75 minutos	56,31 minutos	158 minutos

Tabla 1. Tiempo cálculo Factor de Correlación (Faceta 4x4)

Por este motivo, la principal mejora que encuentro para este proyecto es la realización del mismo utilizando un programa compilado. Estos programas trabajan de la siguiente forma, una vez que el programa está escrito, este se traduce a partir de su código fuente y por medio de un compilador, a un archivo ejecutable que pueda ser utilizado por cualquier plataforma, es decir, un lenguaje compilado es aquel traducido por el propio ordenador y en el cual las instrucciones generadas son interpretadas directamente por la máquina.

Análisis de imágenes basado en correlación de imágenes

Otro aspecto a mejorar y que también ocasionaría cambios en el rendimiento del programa es la optimización de código. Como ya he dicho anteriormente, mi familiaridad con el programa MATLAB es conocida pero breve, por este motivo considero que con una optimización del código utilizando los comandos de los que dispone el programa con mayor utilidad y sin utilizar tantos bucles, se podría conseguir una mejora en el rendimiento de la aplicación.

También se podría mejorar la presentación de resultados. Aunque en las figuras que obtenemos podemos ver claramente los aspectos destacados de la aplicación y la presentación de resultados cumplen los aspectos pedidos en el proyecto, si quiero destacar que se podía mejorar haciendo la parte gráfica más vistosa y cómoda para el usuario.

Sin embargo, un aspecto muy importante a mejorar y que tiene gran repercusión en los resultados obtenidos, son los valores que pueden tomar las facetas, tanto patrón como la de búsqueda. Por ejemplo, puede pasar que una faceta sea uniforme completamente, lo que produce que al calcular el factor de correlación en MATLAB, obtenemos NaN (Not a number) como resultado. Ejemplo:

Si disponemos la siguiente faceta patrón:

$$\begin{bmatrix} 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 & 4 \end{bmatrix}$$

Si sustituimos los valores en las

ecuaciones 196, 197, 198, obtenemos lo siguiente:

$$\rho = \frac{\sigma_{LR}}{\sigma_L \sigma_R} = \frac{0}{0} \text{ NaN}$$

Por esta razón, para tener menor probabilidad de que aparezcan estos errores, lo mejor es utilizar facetas más grandes. Para imágenes de 256 x 256 píxeles, podemos utilizar fácilmente facetas de 10x10 ó 15x15.

En conclusión, la mejora del proyecto es fundamentalmente utilizar de forma más eficaz el lenguaje MATLAB e ir actualizando el proyecto en cuanto surjan nuevas técnicas de procesado, ya que el análisis de imágenes es un campo en continua evolución.

7. Objetivos conseguidos

Al inicio del proyecto como objetivos principales se nos marcaba realizar una estudio teórico de las diferentes técnicas de análisis de imágenes, centrándonos con mayor profundidad y mayor detenimiento en la correlación digital de imágenes, más concretamente en la correspondencia por áreas. Una vez realizado este estudio, se nos pedía realizar un programa en lenguaje MATLAB, usando estas técnicas.

Tras realizar el proyecto, podemos finalizar que se han conseguido los objetivos marcados en un principio, ya que durante todo este documento hemos realizado un exhaustivo análisis sobre el tratamiento digital de la imagen, centrándonos en el análisis que era nuestro principal objetivo. Los conocimientos sobre estas técnicas, junto con el conocimiento del lenguaje MATLAB, me han permitido realizar la última parte de los objetivos, es decir, la realización de una pequeña aplicación sencilla en la cual pudiéramos evaluar y detectar las diferencias y variaciones entre dos imágenes, mostrándose el resultado de forma gráfica.

8. Bibliografía

- Apuntes de la asignatura Tratamiento Digital de la Imagen impartida por Alfonso Martín Marcos y José Manuel Díaz López.
- Proyecto fin de carrera: "Control de movimiento basado en Imágenes". Camilo Vázquez Herrero. Febrero 2012.
- Proyecto fin de carera: "Análisis de Imágenes". Carlos Martín Vallinas. 2001.
- Análisis de deformaciones en probetas planas mediante correlación digital de imágenes. E. López-Alba, F.A. Díaz, R. Dorado, R. López-García.

<http://www.uclm.es/actividades/2010/CongresoIM/pdf/cdarticulos/109.pdf>

- Lenguajes compilados vs Lenguajes interpretados.

<http://www.buenastareas.com/ensayos/Lenguajes-Interpretados-Vs-Lenguajes-Compilados-Desidia/3427582.html>

- Apuntes de Fotogrametría II. Universidad de Extremadura. Apartado 8.2: Principio de la correlación de imágenes.

<http://es.scribd.com/doc/36263964/98/Principio-de-la-correlacion-de-imagenes>

- Tutorial de Matlab.

http://www.usc.es/gir/docencia_files/tdd/tutorial_matlab.pdf

- Tutorial de Matlab. Antonio Souto Iglesias.

http://canal.etsin.upm.es/web_cnum/main_matlab.pdf

- Teoría de la Imagen Digital. Javier Monedero. 2008-2009.

<http://es.scribd.com/doc/67565862/15/Perturbaciones-caracteristicas-y-metodos-de-correccion>

