

TRANSFORMACIONES GEOMÉTRICAS

SOBRE IMÁGENES DIGITALES

Samuel Barreto Melo – sbarreto@udistrital.edu.co

Facultad de Ciencias – Carrera de Matemáticas

Universidad Distrital Francisco José de Caldas

RESUMEN

Las imágenes digitales (o imágenes discretas) provienen del mundo circundante, un mundo continuo. Es así como actualmente se suele representar una gran variedad de fenómenos por medio de fotografías digitales o secuencias de imágenes digitales. Éstas se obtienen a partir de imágenes análogas mediante muestreo y cuantificación, procesos que dependen de los dispositivos de captura de imagen de las cámaras digitales. El área del procesamiento de señales e imágenes se ha venido desarrollando a través del estudio teórico de las diferentes transformaciones, que se manifiesta en la creación de algoritmos que proyectan soluciones a problemas de la vida real, soportados por avanzados elementos de software y de hardware. Este artículo presenta algunas de las transformaciones más usadas sobre imágenes digitales tanto a nivel de píxeles, como de sus valores de intensidad. Transformaciones lineales y no lineales se implementan a través de una herramienta de software como Matlab. Igualmente se establecen aspectos teóricos del álgebra lineal: transformaciones lineales y transformaciones afines. Finalmente se estudian aspectos numéricos de interpolación sobre imágenes.

Palabras clave: transformación lineal, transformación afín, interpolación, matriz e imagen digital.

1. INTRODUCCIÓN

En el procesamiento de señales e imágenes se utilizan algunas técnicas sustentadas en el conocimiento de los operadores lineales y no lineales. El avance de las tecnologías de comunicación e información permiten hoy en día la aplicación de imágenes (matrices) y transformaciones del álgebra lineal a diversas áreas de las ciencias puras y aplicadas así como de la ingeniería.

Los conceptos de espacio vectorial y álgebra lineal resultan ser muy naturales en el procesamiento de señales. Y más aun en el procesamiento de imágenes digitales, ya que una imagen digital es una matriz. Las operaciones matriciales juegan un papel importante en el procesamiento de imágenes digitales. Una noción fundamental en el desarrollo de este estudio tiene que ver con que una imagen puede visualizarse de diferentes formas, y la manipulación de su *representación* es una de las más poderosas herramientas disponibles. Las manipulaciones útiles pueden ser lineales (ejemplo, las transformaciones) o no lineales (ejemplo, la cuantificación).

Transformaciones de diverso tipo son aplicadas a las imágenes con diferentes propósitos, produciendo así el desarrollo de nuevos y revolucionarios algoritmos a través de la matemática, que son soportados por poderosas aplicaciones de software, unificando las distintas áreas de las ciencias puras y aplicadas. Es importante resaltar, que lo significativo es el germen o la ciencia que hay detrás de los métodos y no los algoritmos en sí.

En la sección 2 se describen los conceptos básicos de la transformación afín 2D en coordenadas homogéneas, la forma matricial, y coordenadas espaciales. En la sección 3 se presenta las transformaciones espaciales. En la sección 4 se exhiben los métodos más comunes de interpolación y ejemplos. Y finalmente en la sección 5 se presentan algunas conclusiones, además de la bibliografía en la sección 6 .

2. TRANSFORMACIONES GEOMÉTRICAS

Las transformaciones Geométricas modifican la relación espacial entre píxeles. En términos del procesamiento de imágenes digitales una transformación geométrica consiste de dos operaciones básicas:

1. Una transformación espacial que define la reubicación de los píxeles en el plano imagen.
2. Interpolación de los niveles de grises, los cuales tienen que ver con la asignación de los valores de intensidad de los píxeles en la imagen transformada.

En términos Matemáticos las transformaciones afines son las más usadas en imágenes digitales 2D por su representación y manejo matricial.

Una Transformación afín es aquella (transformación) en la que las coordenadas (x', y') del punto imagen son expresadas linealmente en términos de las del punto original (x, y) . Es decir, la transformación viene dada por las ecuaciones:

$$\begin{aligned} x' &= ax + by + m \\ y' &= cx + dy + n \end{aligned} \quad (1)$$

Cuando $m = n = 0$ las ecuaciones anteriores se convierten en el prototipo de una transformación lineal multiplicación por una matriz a la izquierda, esto es,

$$\begin{cases} x' = ax + by \\ y' = cx + dy \end{cases} \longrightarrow \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (2)$$

El sistema tiene solución única sí y sólo si $\Delta = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$, es decir, la matriz es no singular, por lo tanto la transformación inversa existe y viene dada por:

$$\begin{cases} x = a'x' + b'y' \\ y = c'x' + d'y' \end{cases} \longrightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \quad (3)$$

La imagen de una recta es una recta:

$$\begin{aligned} Ax + By + D = 0 &\longrightarrow A(a'x' + b'y') + B(c'x' + d'y') + D = 0 \\ &\longrightarrow (Aa' + Bc')x' + (Ab' + Bd')y' + D = 0 \\ &\longrightarrow A'x' + B'y' + D = 0 \end{aligned}$$

Por lo anterior, no es difícil mostrar que las transformaciones afines transforman una grilla (o malla) en otra grilla, caracterizándolas como las más naturales a las imágenes.

En el estudio del significado de los coeficientes de la matriz de la transformación se puede comprobar que sí

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} SC_x & S_{Horiz.} \\ S_{Vert.} & SC_y \end{pmatrix} \quad (4)$$

donde, $a = SC_x$ corresponde al escalado en x , $d = SC_y$ escalado en y , $b = S_{horiz}$ presión o inclinación en dirección horizontal y $c = S_{vert}$ presión o inclinación en dirección vertical.

Ejemplo: Sea $T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ y \end{pmatrix}$. El efecto de los coeficientes de la matriz de la transformación puede observarse en la figura 1.

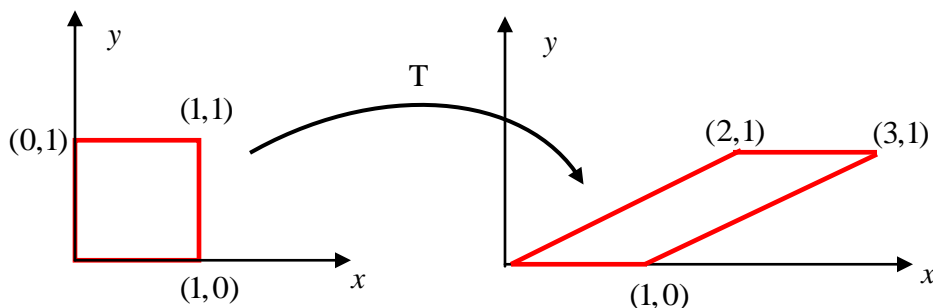


Figura 1. Transformación Presión o empuje en dirección horizontal.

Existen también otras acciones de las transformaciones afines tal como la de rotar una región geométrica un ángulo θ alrededor de un punto fijo y que se define como:

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (5)$$

Coordenadas homogéneas:

Un punto (x, y) en coordenadas homogéneas se representa por la terna (xw, yw, w) . Es así, como el punto $(3, 2)$ del sistema de coordenadas cartesianas se representa como $(6, 4, 2)$, o como $(15, 10, 5)$ entre otros. Y la representación normalizada de $(3, 2)$ en coordenadas homogéneas es $(3, 2, 1)$ que es la convención más usada.

La representación matricial en coordenadas homogéneas de la transformación afín 2D:

$$\begin{cases} x' = ax + by + m \\ y' = cx + dy + n \end{cases} \longleftrightarrow \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & m \\ c & d & n \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (6)$$

donde los escalares a, b, c, d tienen el mismo significado que los de la ecuación (4), m corresponde al desplazamiento horizontal mientras n al desplazamiento vertical. En la figura 2 se ilustran la aplicación de una transformación afín a una imagen cuya acción es la de una presión o empuje en dirección horizontal y luego una acción compuesta de rotación, escalado y traslación.

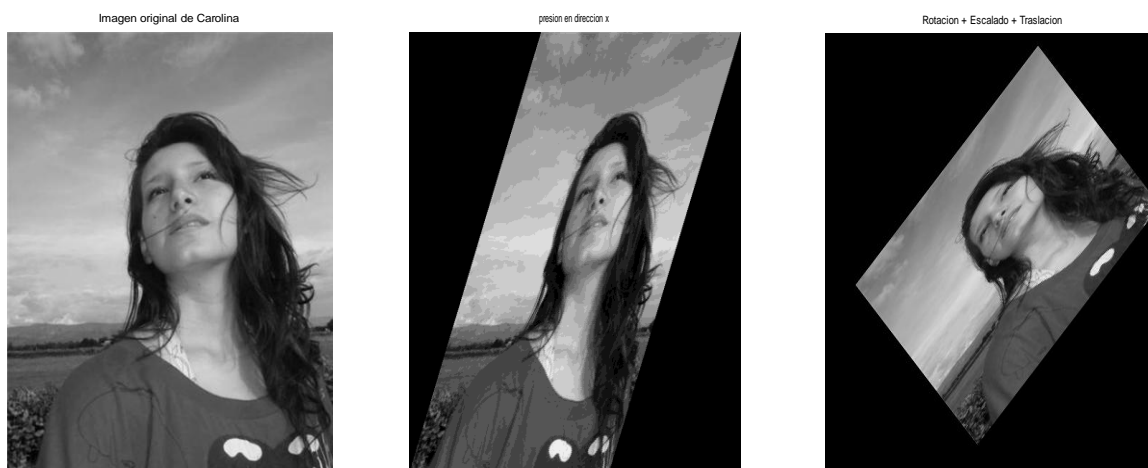


Figura 2. Transformación afín actuando sobre una imagen digital.

3. TRANSFORMACIONES ESPACIALES

La mayoría de los métodos para transformar imágenes se clasifican:

1. Métodos que usan el mapeo directos o hacia adelante
2. Métodos que usan el mapeo inverso.

Los mapeos hacia adelante son más complejos y costosos de implementar computacionalmente. El método del mapeo inverso es más fácil de implementar que el mapeo directo. El procedimiento de mapeo inverso valora cada píxel de salida y calcula las correspondientes posiciones en la imagen de entrada usando e interpolando entre los píxeles de la imagen de entrada más cercanos para determinar el valor del píxel de salida

En el caso de la distorsión geométrica si suponemos que esta se puede modelar mediante el par de ecuaciones bilineales

$$\begin{aligned}x' &= r(x, y) = c_1x + c_2y + c_3xy + c_4 \\y' &= s(x, y) = c_5x + c_6y + c_7xy + c_8\end{aligned}\quad (7)$$

Con 8 puntos de *unión* conocidos estas ecuaciones se pueden resolver para los c_k con $k=1,2,3,\dots,8$. Estos coeficientes modelarían la distorsión geométrica y por tanto podrían ser usados para transformar todos los píxeles dentro de un cuadrilátero definido por los puntos de atadura que se usaron para obtener los coeficientes.

Estos puntos de atadura actualmente se obtienen mediante técnicas muy distintas. Algunos sistemas de generación de imágenes poseen físicamente mallas de puntos metálicos dentro del sensor de imágenes produciendo un conjunto conocidos como mascarar obtenidos directamente sobre la imagen cuando ésta es capturada.

4. INTERPOLACION

Cuando se aplican transformaciones afines a las imágenes como por ejemplo una rotación se presentan cambios de 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. Existen tres métodos conocidos que son *el del vecino más cercano*, *de interpolación bilineal* y *de interpolación bicúbica*.

El método de interpolación de niveles de grises más simple es el fundamentado en el del vecino más cercano, que consiste en escoger como el valor del píxel de una imagen de salida el valor del píxel de entrada más cercano en la dirección inversa. Este método, aunque es simple de implementar, presenta frecuentemente artefactos indeseables en las imágenes, tales como distorsión de líneas en imágenes de alta resolución, motivo por el cual no se aconseja para el suavizado de imágenes. Este tipo de interpolación logra producir un error espacial relativo a lo más de $\frac{1}{\sqrt{2}}$ píxel por unidad. En la figura 3 se ilustra el método del vecino más cercano en la interpolación de niveles de grises de una imagen deformada $g(x, y)$.

El esquema consiste en tomar el píxel de la posición $(x, y) = (a, b)$ en la imagen original, y transformarlo mediante las ecuaciones de distorsión

$$\begin{aligned}x' &= r(x, y) = c_1x + c_2y + c_3xy + c_4 \\y' &= s(x, y) = c_5x + c_6y + c_7xy + c_8\end{aligned}$$

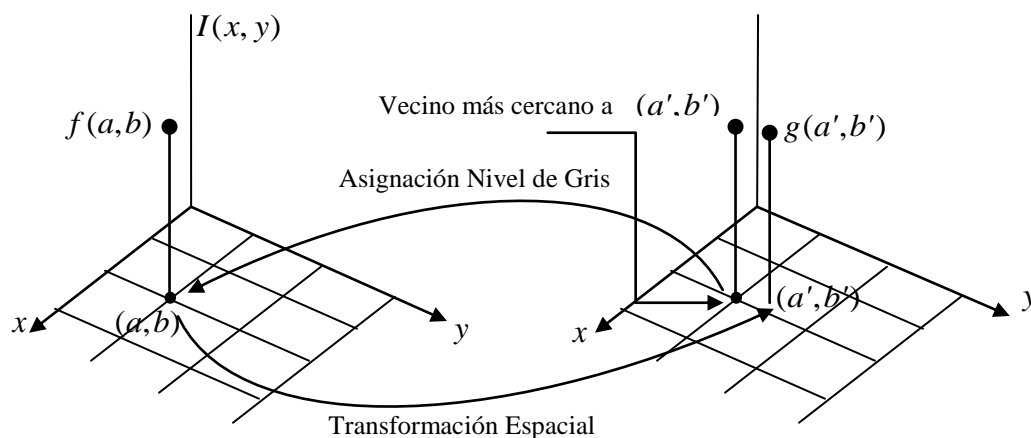


Figura 3. Interpolación de nivel de gris basado en el concepto del vecino más cercano.

para obtener las correspondientes coordenadas (a', b') que pueden ser enteras o racionales según los coeficientes, posteriormente se elige el vecino de coordenadas enteras más cercano a (a', b') y finalmente se asigna el nivel de gris de este vecino al píxel ubicado en (a, b) lográndose un valor aproximado al píxel original. El método se repite de la misma forma con cada uno de los píxeles restantes de la imagen hasta lograr una imagen restaurada $\hat{f}(x, y)$ que aproxima a $f(x, y)$. Nótese en el análisis la imagen como señal 2D.

El error de interpolación por muestra se puede reducir significativamente cuando se utilizan los cuatro vecinos más cercanos en la interpolación. En este método, llamado de interpolación bilineal, se interpola linealmente a lo largo de cada fila de una imagen, y posteriormente el resultado se interpola linealmente en la dirección por columnas. En la figura 4 se ilustra una vecindad alrededor de (p', q') , donde $a = p' - p$, $b = q' - q$

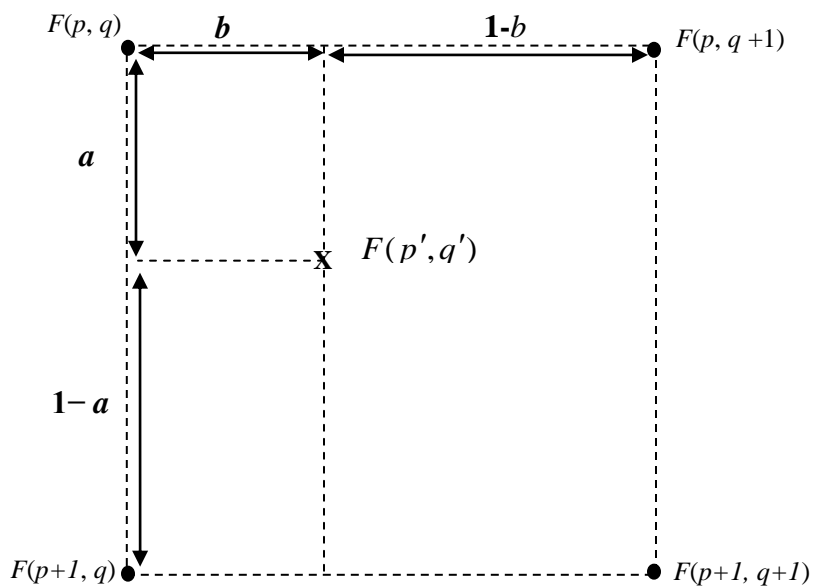


Figura 4. Vecindad en el proceso de Interpolación bilineal.

Realizando las dos interpolaciones lineales horizontales de $F(p, q)$ a $F(p, q+1)$ y de $F(p+1, q)$ a $F(p+1, q+1)$, se tiene como resultado $(1-b)F(p, q)+bF(p, q+1)$ y $(1-b)F(p+1, q)+bF(p+1, q+1)$, ahora bien, realizando nuevamente la interpolación lineal vertical sobre los dos valores hallados anteriormente se obtiene el valor de intensidad del pixel interpolado como se muestra en la figura 5, y cuyo valor es:

$$F(p', q') = (1-a)[(1-b)F(p, q)+bF(p, q+1)] + a[(1-b)F(p+1, q)+bF(p+1, q+1)] \quad (8)$$

Aunque las operaciones horizontal y vertical son lineales por separado, en la mayoría de los casos, su aplicación sucesiva da como resultado una superficie no lineal adaptada entre los cuatro pixeles vecinos.

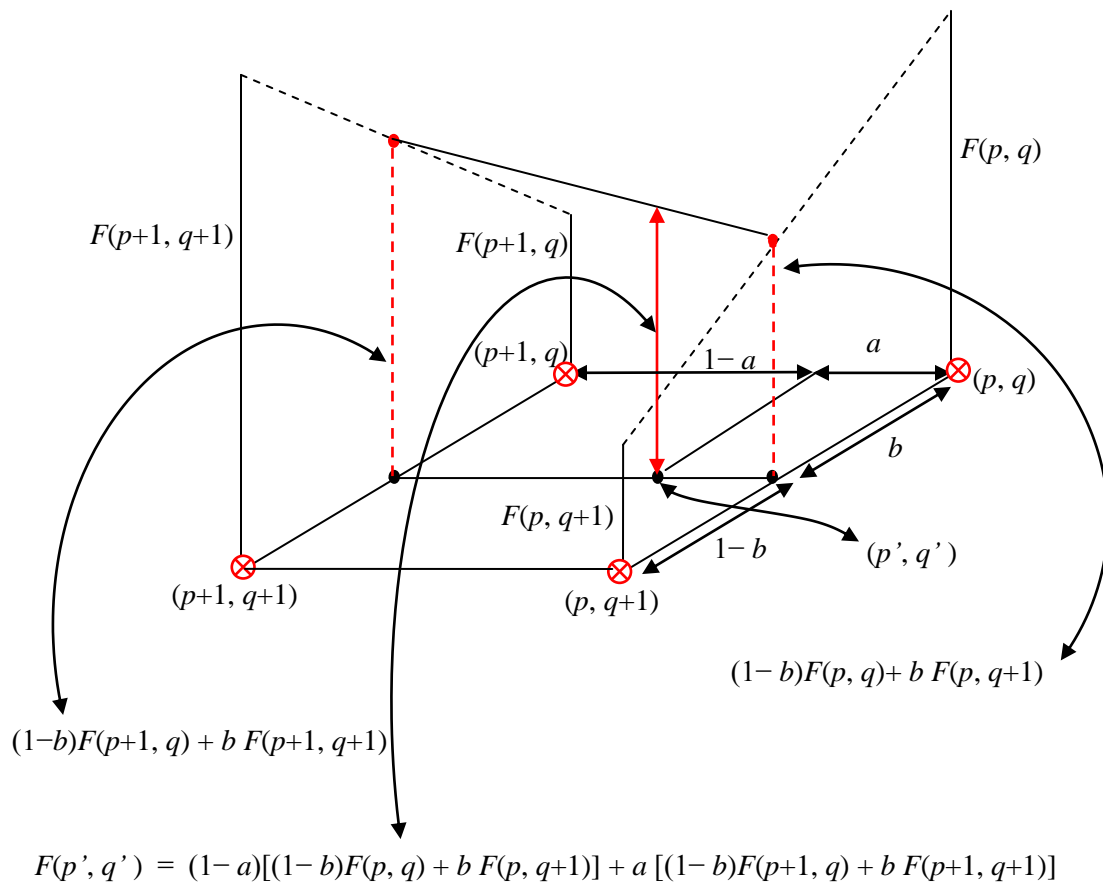


Figura 5. Ilustración geométrica de la interpolación bilineal interpretando la imagen como una señal 2D.

La expresión de interpolación bilineal anterior se puede generalizar para cualquier función de interpolación $H(x)$ que valga cero por fuera del rango de más o menos un espacio por muestra. Con la función $H(x)$ la interpolación bilineal puede expresarse como la suma de cuatro funciones *peso de interpolación* así:

$$F(p', q') = H[-a]H[b] \cdot F(p, q) + H[-a]H[-(1-b)] \cdot F(p, q+1) + \\ + H[1-a]H[b] \cdot F(p+1, q) + H[-(1-a)]H[-(1-b)] \cdot F(p+1, q+1) \quad (9)$$

Sí definimos $h(x) = \begin{cases} 1-|x| & \text{si } 0 < |x| < 1 \\ 0 & \text{de otro modo} \end{cases}$ como función núcleo de interpolación

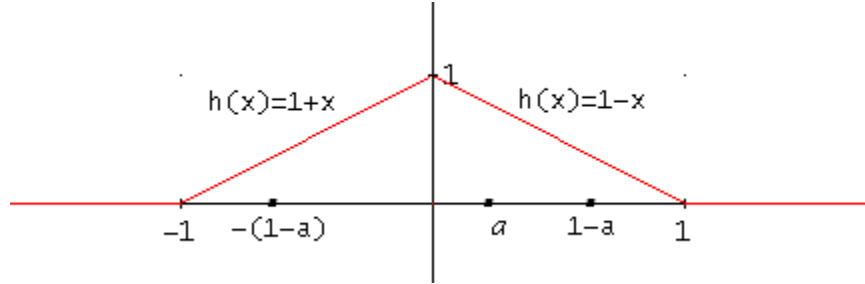


Figura 6. Grafica del núcleo de interpolación $h(x)$.

y tomamos como caso particular la interpolación lineal $H(x) = h(x)$ en la ecuación (9), se obtiene la expresión de la ecuación (8) correspondiente a la interpolación bilineal. Nótese que:

$$H[-a] = h(-a) = 1-a, \quad H[b] = h(b) = 1-b, \quad H[-(1-b)] = h(-(1-b)) = b, \\ H[1-a] = h(1-a) = a, \quad H[-(1-a)] = h(-(1-a)) = a$$

Usualmente, la interpolación se reduce a cuatro píxeles vecinos debido a la complejidad computacional. No obstante, en la figura 7 se exhibe un modelo de vecindad generalizada de interpolación bicúbica en la que de los 16 vecinos que rodean a $F(p', q')$, el píxel $F(p, q)$ se muestra como el vecino más cercano al píxel cuyo valor se va a interpolar. El píxel interpolado se puede expresar en forma compacta como:

$$F(p', q') = \sum_{m=-1}^2 \sum_{n=-1}^2 F(p+m, q+n) \cdot H_c[m-a] \cdot H_c[n-b] \quad (10)$$

$$F(p', q') = \sum_{m=-1}^2 \{ F(p+m, q-1)H_c[m-a]H_c[-1-b] + F(p+m, q)H_c[m-a]H_c[-b] \\ + F(p+m, q+1)H_c[m-a]H_c[1-b] + F(p+m, q+2)H_c[m-a]H_c[2-b] \}$$

Los 16 términos expresados por filas son:

$$\text{Fila (p-1): } \{ F(p-1, q-1)H_c[-1-a]H_c[-1-b] + F(p-1, q)H_c[-1-a]H_c[-b] + \\ + F(p-1, q+1)H_c[-1-a]H_c[1-b] + F(p-1, q+2)H_c[-1-a]H_c[2-b] \}$$

$$\text{Fila (p): } \{ F(p, q-1)H_c[-a]H_c[-1-b] + F(p, q)H_c[-a]H_c[-b] + \\ + F(p, q+1)H_c[-a]H_c[1-b] + F(p, q+2)H_c[-a]H_c[2-b] \}$$

$$\text{Fila (p+1): } \{F(p+1, q-1)H_c[1-a]H_c[-1-b] + F(p+1, q)H_c[1-a]H_c[-b] + \\ + F(p+1, q+1)H_c[1-a]H_c[1-b] + F(p+1, q+2)H_c[1-a]H_c[2-b]\}$$

$$\text{Fila (p+2): } \{F(p+2, q-1)H_c[2-a]H_c[-1-b] + F(p+2, q)H_c[2-a]H_c[-b] + \\ + F(p+2, q+1)H_c[2-a]H_c[1-b] + F(p+2, q+2)H_c[2-a]H_c[2-b]\}$$

El total de la suma de los cuatro números por fila es igual a $F(p', q')$. La función de interpolación bicúbica $H_c(x)$ puede tomarse ya sea como un spline B-cubico o una función de interpolación cubica.

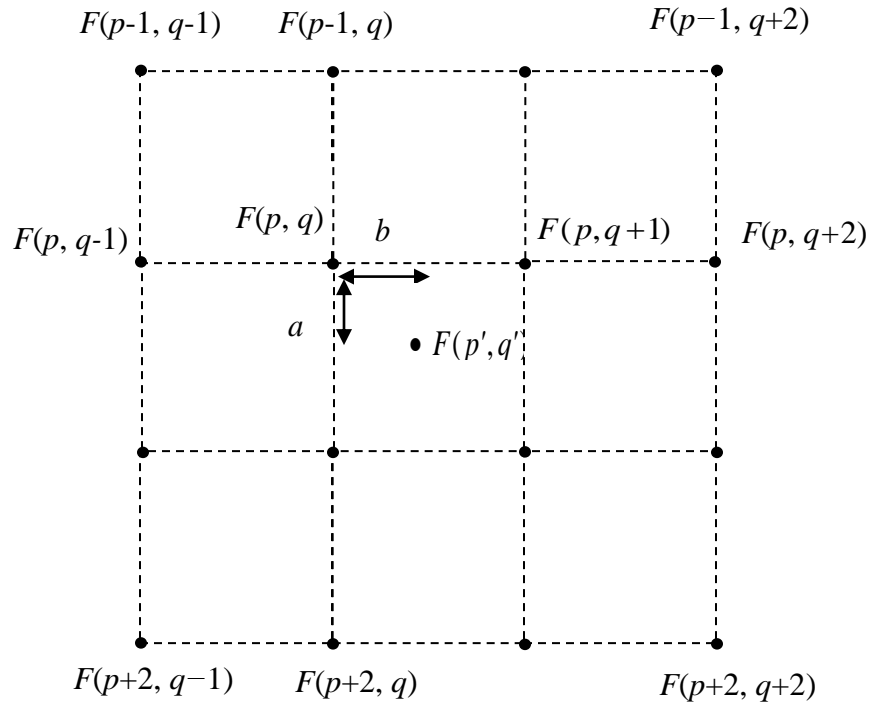


Figura 7. Vecindad en el proceso de Interpolación bicúbica

Sí tomamos como caso particular la interpolación cúbica $H_c(x) = h(x)$ en la ecuación (10), obtenemos la expresión correspondiente a la interpolación bicúbica. El núcleo más común de $h(x)$ es el siguiente y su grafica puede verse en la figura 8:

$$h(x) = \begin{cases} 1 - 2|x|^2 + |x|^3 & \text{si } 0 < |x| < 1 \\ 4 - 8|x| + 5|x|^2 - |x|^3 & \text{si } 1 < |x| < 2 \\ 0 & \text{otra parte} \end{cases}$$

Aunque el algoritmo del método de interpolación bicúbico es uno de los más utilizados en general se busca un óptimo entre tiempo de procesamiento y calidad de la imagen de salida.

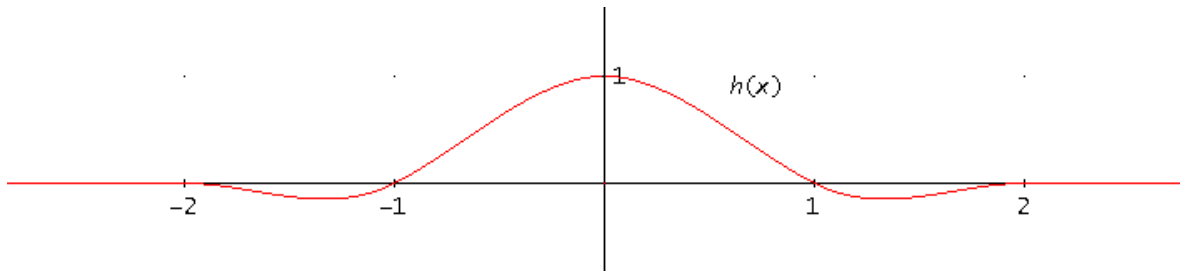
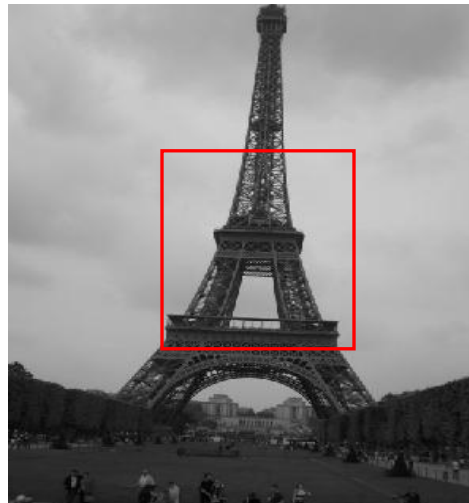


Figura 8. Gráfica de la función núcleo de interpolación bicúbica $h(x)$.

En la figura 9 se muestra la acción de una rotación de 45° bajo los tres métodos de interpolación a la imagen original y un zoom posterior para su comparación. Nótese los artefactos o distorsión que se produce en la imagen bajo la interpolación del vecino más cercano. En los otros dos métodos las distorsiones son menos pronunciadas pero el costo computacional es mayor.



TORRE EIFEEL - bilineal

TORRE EIFEEL - bicubica

TORRE EIFEEL - vecino mas cercano

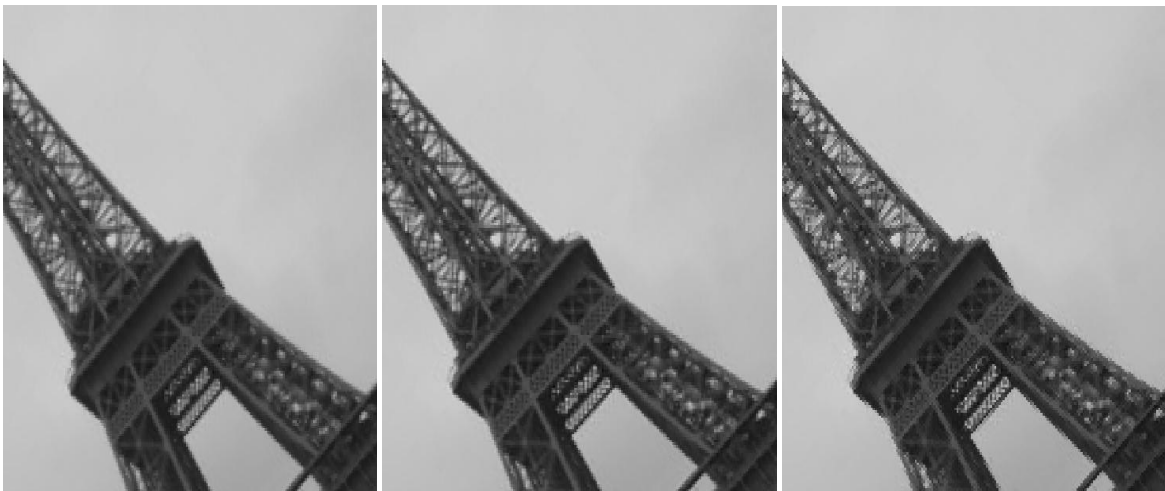


Figura 9. Imagen de la torre Eiffel. Nótese el perfil de la torre para la interpolación bilineal, bicúbica y la del vecino más cercano.

Ya que una imagen se puede visualizar como una función 2D cuya grafica es una superficie, los métodos de interpolación también se aplican allí con mucha frecuencia, para obtener una mayor información sobre datos desconocidos a partir de los ya conocidos con el objeto de predecir el comportamiento “continuo” de los datos. Es así como en el estudio de suavizado de superficies los métodos de interpolación expuestos anteriormente permiten un análisis comparativo. En la figura 10 se muestra la acción de la interpolación sobre una superficie con buenos resultados de suavizado para los métodos bilineal y bicúbico. Y en la figura 11 se exhibe de forma complementaria las curvas de nivel de las superficies interpoladas obteniéndose información relativa sobre la forma de la superficie.

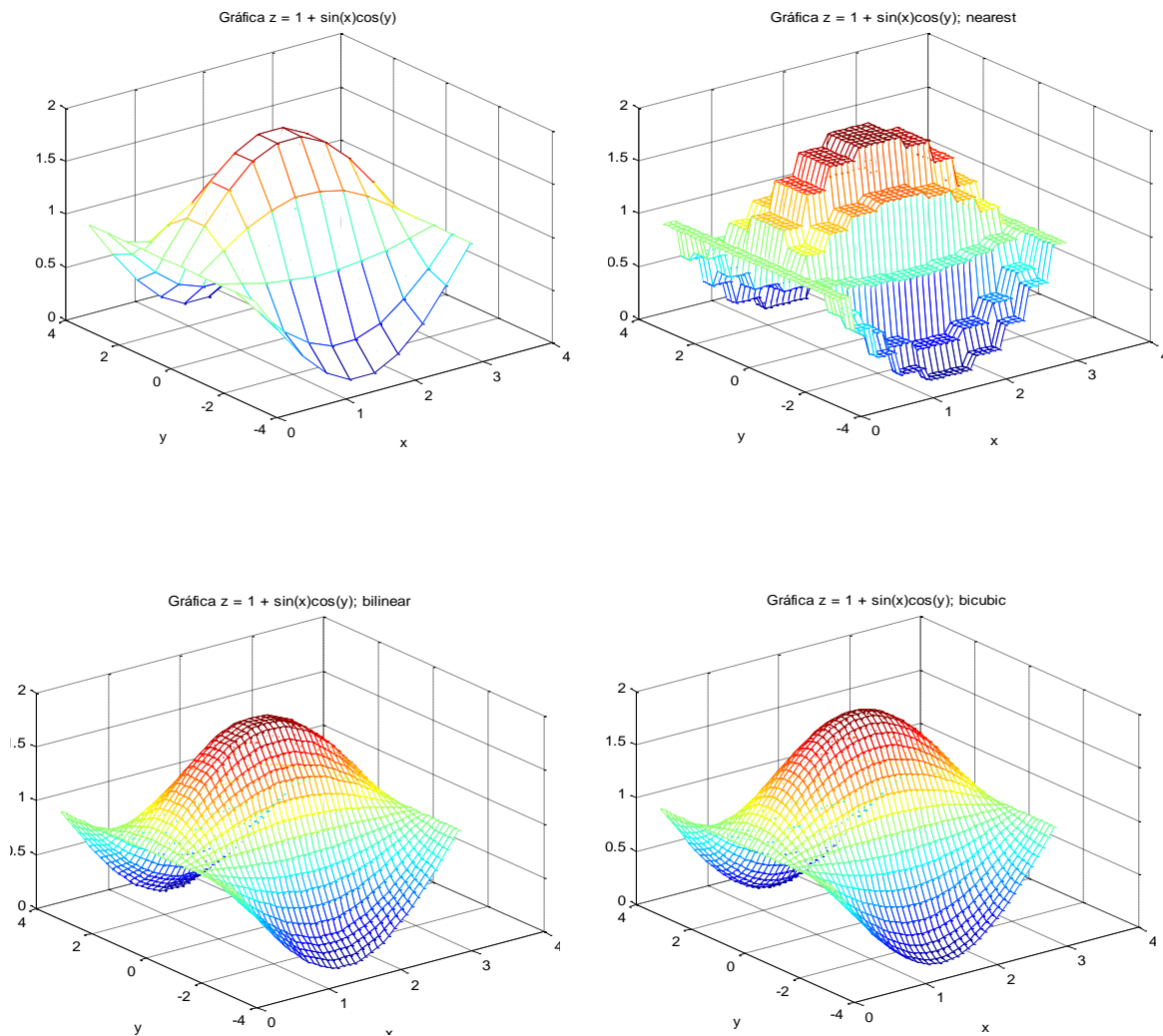


Figura 10. Superficie $z = 1 + \sin(x)\cos(y)$ y sus tres tipos de interpolación

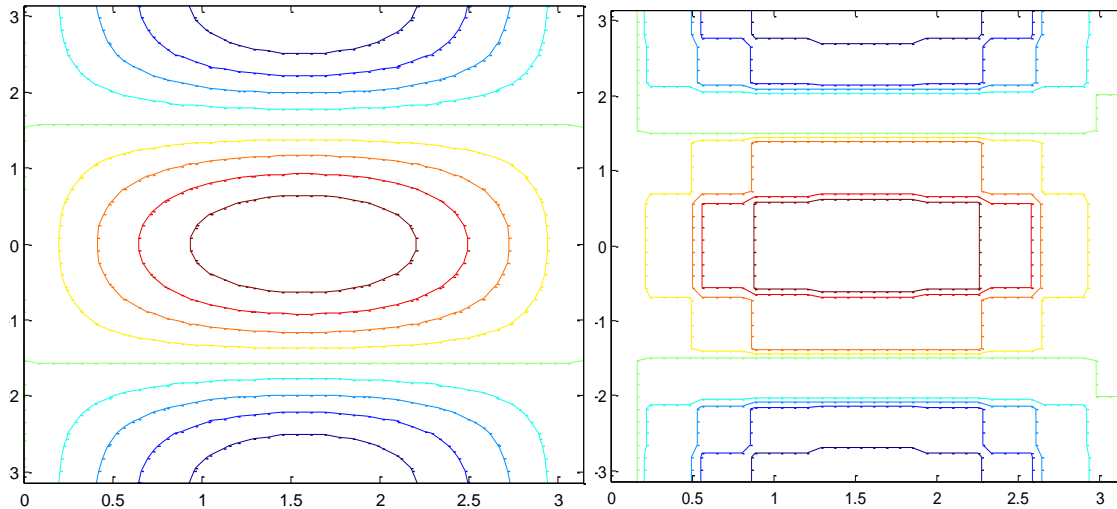


Figura 11. Las curvas de nivel de la superficie interpolada mediante el método bicúbico y el vecino más cercano.

5. CONCLUSIONES

1. Aunque la eficiencia de los métodos de interpolación depende de la imagen que se pretenda interpolar y de la máquina utilizada, los métodos de interpolación bilineal y bicúbica son muy usados debido a su buena aproximación en cuanto a la calidad de la imagen sin llegar a conseguir la calidad de la imagen original.
2. Los métodos de buena aproximación tienen el inconveniente de generar un alto costo computacional. Pero a su vez son muy requeridos en aquellas aplicaciones donde se necesita una alta fidelidad de la imagen transformada.
3. En caso de una imagen con alto contenido frecuencial los resultados del método bilineal y bicúbico son muy similares salvo el costo computacional.
4. Las transformaciones geométricas aplicadas a las imágenes exigen la solución de problemas de aproximación que varían según el tipo de aplicación.
5. Se destacan las propiedades de Matlab como software de análisis en el área de señales e imágenes, y como la representación computacional de estas (arreglos y matrices), se acomoda naturalmente al esquema de procesamiento del software.
6. Los métodos de interpolación se utilizan en aquellas situaciones donde se quiera recuperar información perdida o información desconocida. `
7. Los aspectos matemáticos de la interpretación de la imagen como superficie muestran que el ajuste de la superficie mediante el método bicúbico es mejor que el presentado por el bilineal y el del mejor vecino respecto a características como la de la suavidad

6. BIBLIOGRAFIA.

- [1] Gonzalez, R. & Woods R. Digital Image Processing. Second Edition. Prentice Hall. New Jersey. 2002.
- [2] Gonzalez, R. & Woods R. Digital Image Processing using MATLAB. 1^a Edition. Prentice Hall. New Jersey. 2004.
- [3] Strang, Gilbert. Introduction to Linear Algebra. Wellesley-Cambridge Press. 3^a Edition 2003.
- [4] Pajares Gonzalo, Jesus M. de la cruz, y otros. Imágenes Digitales Procesamiento práctico con Java. Edición Original .Alfaomega y Ra-Ma. 2004.
- [5] Pratt K. William., Digital Image Processing, 4^a Edition. Wiley-Interscience. Feb 2007.
- [6] Vetterli, M., Wavelets, Aproximation, and Compression, IEEE Signal Processing Magazine. Sept 2001.