

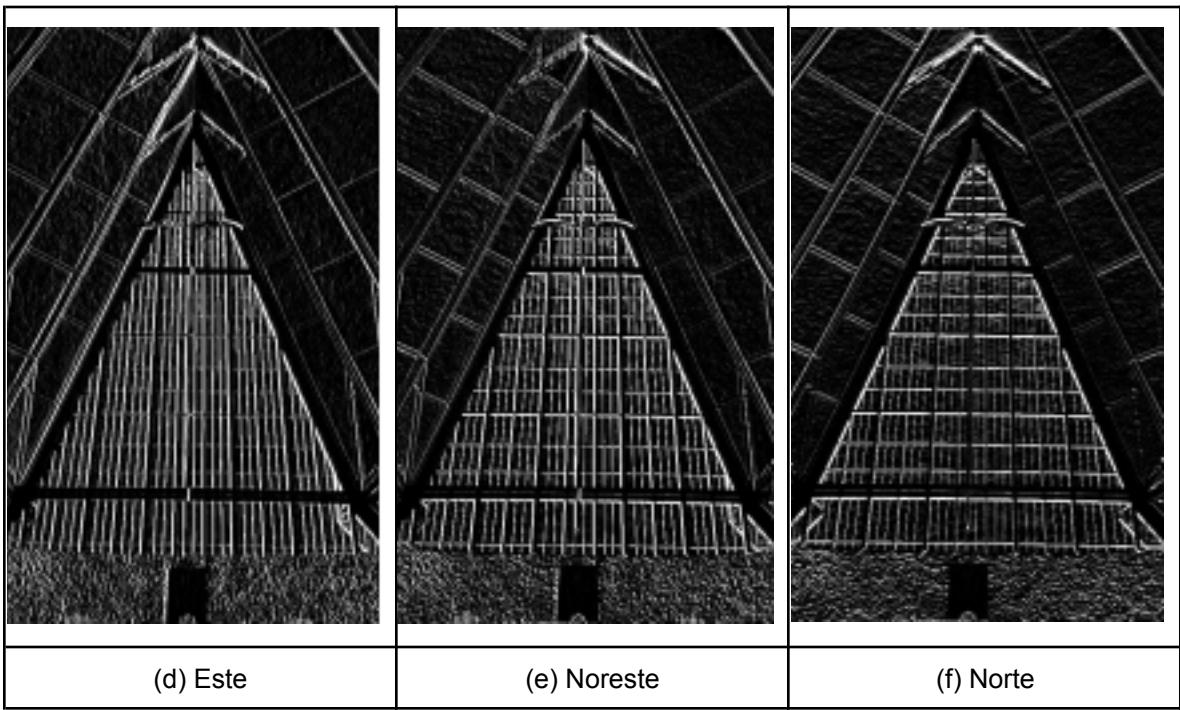
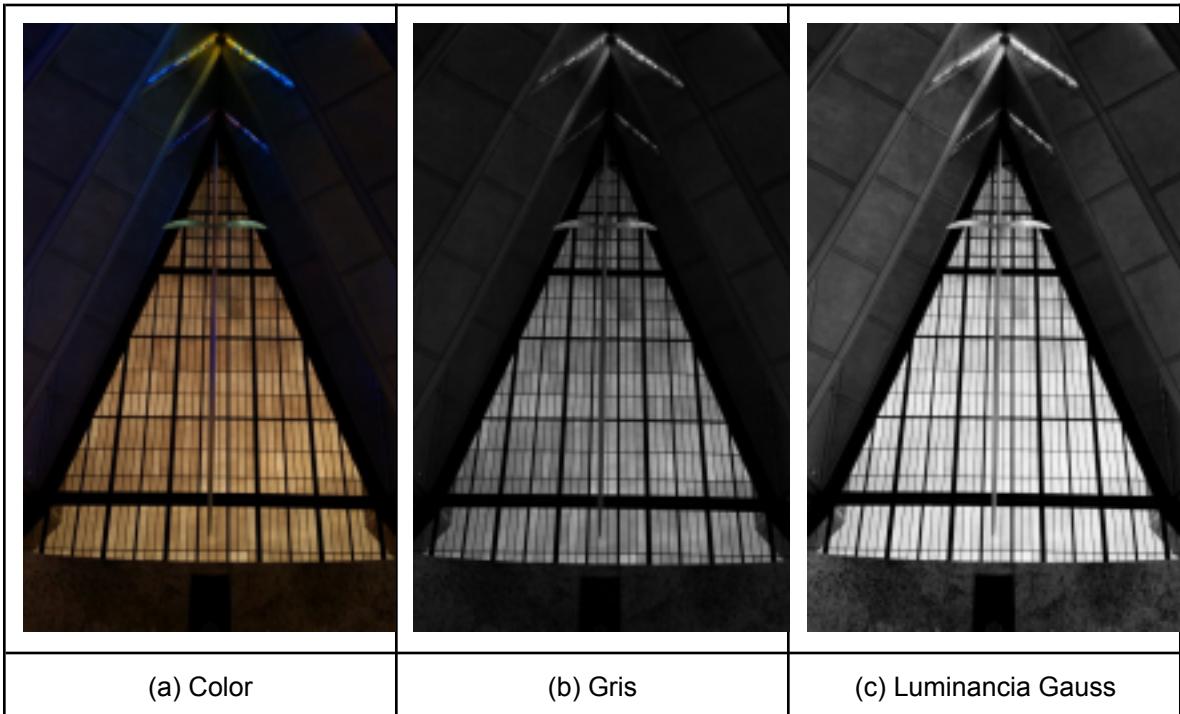
3.3. Extracción de esquinas

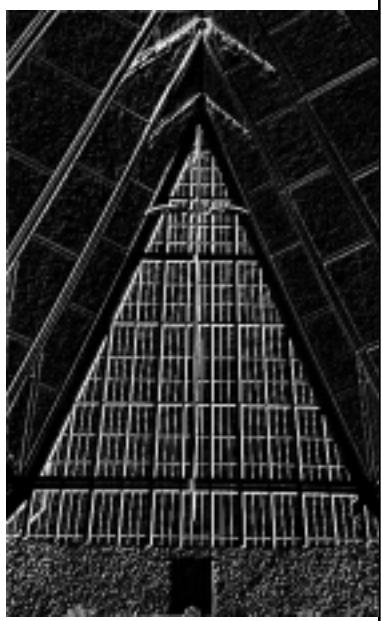
El estado del arte en el procesamiento de imágenes es tal que a medida que aumenta la complejidad de la tarea, disminuye el número de técnicas adecuadas para enfrentar esas tareas. Esto es particularmente cierto cuando se trata de descriptores de funciones aplicables a imágenes completas que son miembros de una gran familia de imágenes. Analicemos dos de los principales métodos de detección de características que se utilizan actualmente para este propósito.

Uno se basa en la detección (extracción) de esquinas y el otro funciona con regiones enteras en una imagen. Intuitivamente, pensamos en una esquina como un cambio rápido de dirección en una curva. Las esquinas son características altamente efectivas porque son distintivas y razonablemente invariantes para cualquier punto de vista. Debido a estas características, las esquinas se utilizan habitualmente para hacer coincidir las características de la imagen en aplicaciones como el seguimiento para la navegación autónoma, los algoritmos de visión artificial estéreo y las consultas de bases de datos de imágenes.

Una esquina (ángulo formado) se define como el espacio entre dos líneas. Un área también puede definirse como un área rodeada por dos vías principales y vías alternas. Los lugares de interés son lugares en la imagen que se conocen como lugares famosos que seguramente conoces, es decir, lugares que pueden ser rincones, pero también lugares alejados de la cima, al menos o del final. Una línea o punto en una curva que se ajusta perfectamente a una curva local.

La técnica de detección de esquinas se denomina punto de interés y, debido a la ubicación del punto de esquina, se debe investigar el punto de interés para determinar cuál es el punto de esquina. Hay detectores de contornos que pueden usarse para descubrir esquinas con post-procesamiento, ellos son el operador de mascara de Kirsch y el operador de mascara de Frei-Chen, en la figura 3.37 se muestra un ejemplo con todas las direcciones del operador de Kirsch y en la figura 3.38 el operador de Frei-Chen.





(g) Noroeste



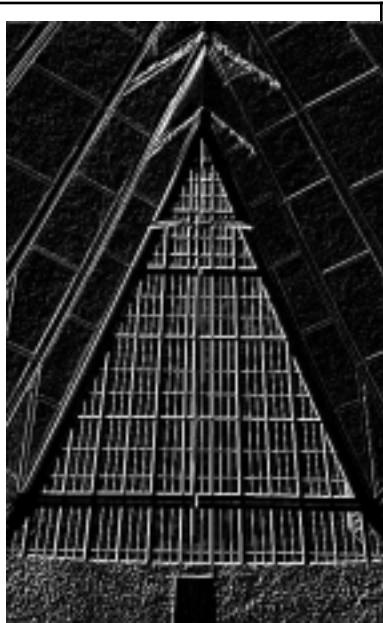
(h) Oeste



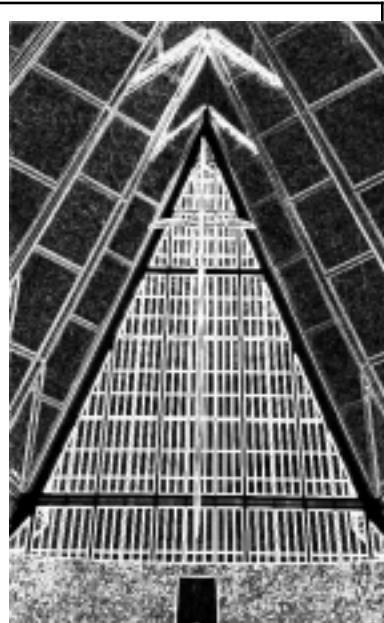
(i) Suroeste



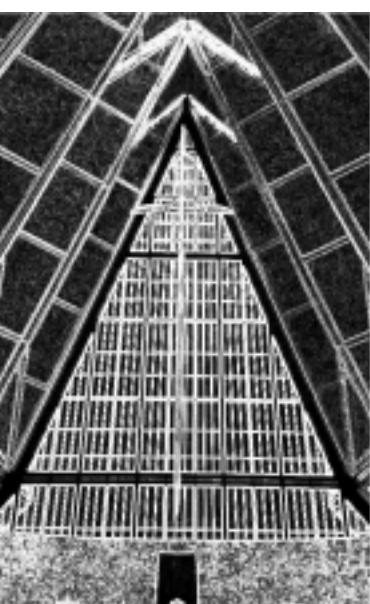
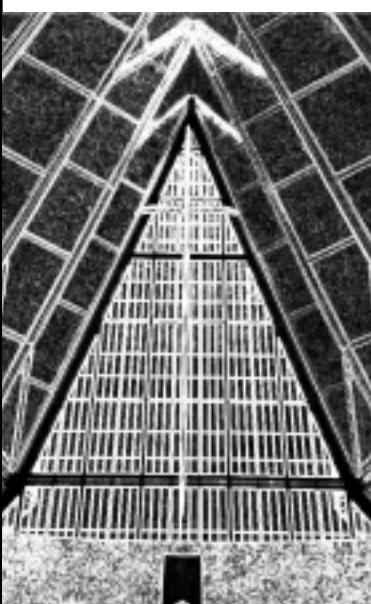
(j) Sur

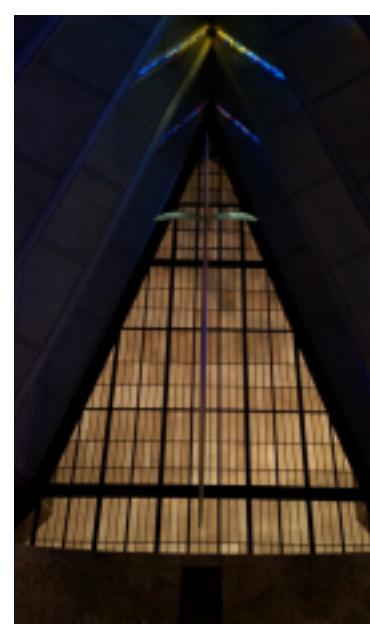
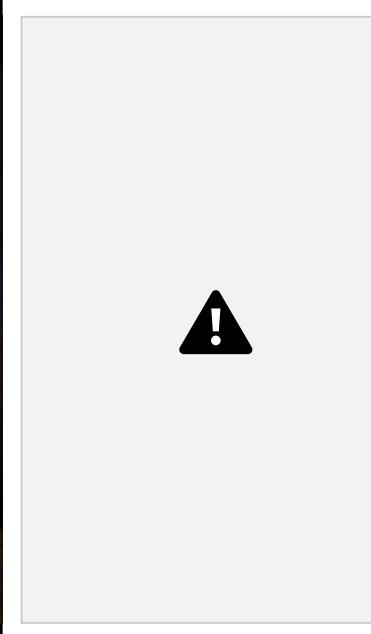
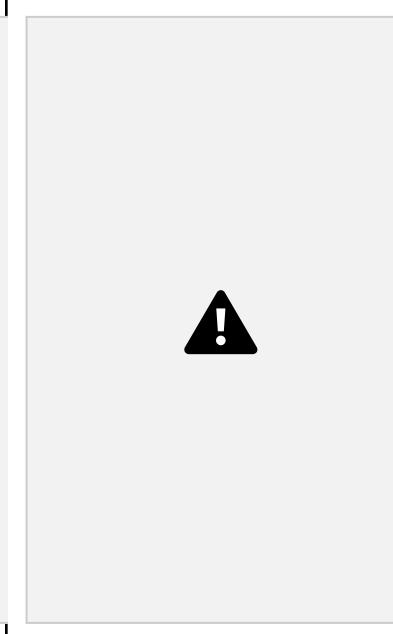


(k) Sureste



(l) Cuatro puntos cardinales

		<p>Figura 3.37(a) Imagen en color original de 512 x 805 pixeles, (b) Imagen de gris RGB a YIQ, (c) Realce de luminancia por Gauss de (b), de (d) hasta (k) se aplica la convolución de la imagen con las máscaras de la figura 3.24, (l) es el aplicado de este, norte, oeste y sur, (m) es el aplicado de noreste, noroeste, suroeste y sureste, (o) se aplican todas las máscaras. Para (l)y(m) se efectúan cuatro convoluciones por cada una y para (n) son ocho convoluciones con sus respectivas máscaras.</p>
(m) Cuatro puntos diagonales	(n) Todas	

		
(a) Original a color	(b) Gris Y	(c) Realce de Gauss

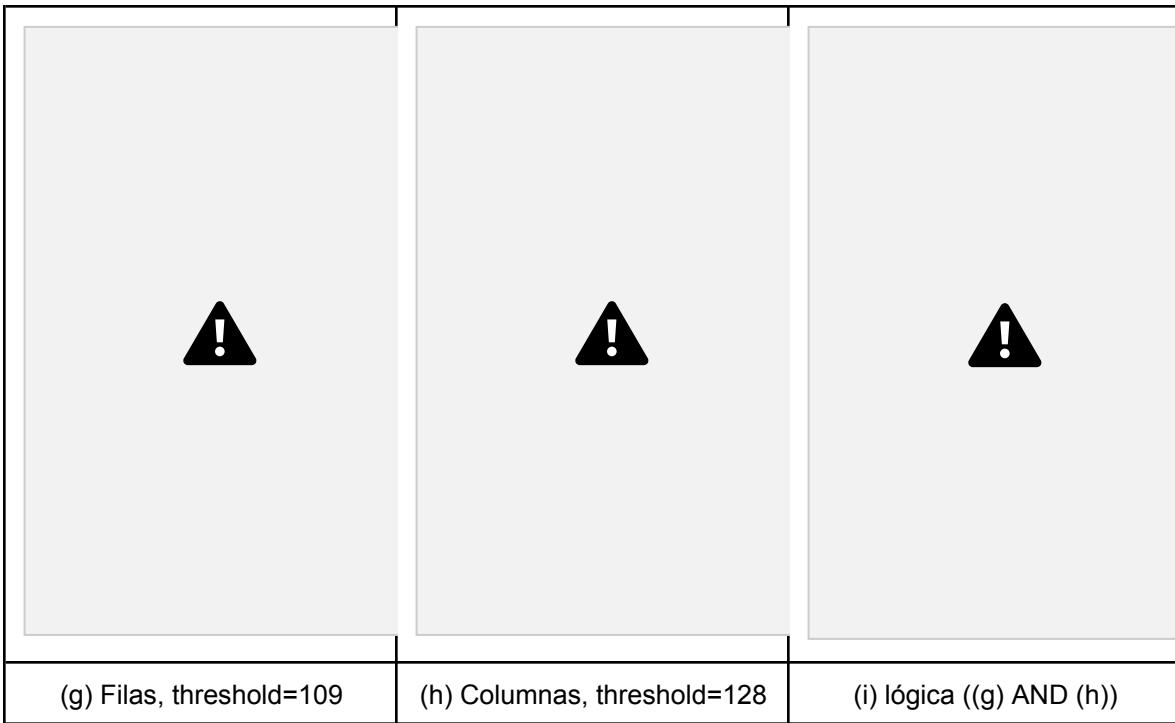


Figura 3.38 (a) Imagen en color original de 512 x 805 pixeles, (b) Imagen de gris RGB a YIQ, (c) Realce de luminancia por Gauss de (b), de (d) Filas de Frei-Chen (e) Columnas de Frei-Chen, (f) Operación lógica OR de (d) y (e), (g) Aplicado de umbral a (d), (h) Aplicado de

umbral a (e), (i) Operación lógica AND entre (g) y (h).

Esquinas, puntos de interés y rasgos, se usan de manera diferente en el texto. Algunas luces se utilizan para marcar áreas, llamadas balizas, pero a veces se confunden con luces de esquina. Los escáneres de esquina no son muy robustos y a menudo requieren una inspección especial o el uso de escáneres más grandes para evitar el impacto de errores individuales en el proceso de búsqueda.

Un algoritmo para la detección de esquinas fue formulado por Harris y Stephens en 1988. La idea básica del detector de esquina Harris-Stephens (HS) se ilustra en la figura 3.37. El enfoque es el siguiente, las esquinas son detectadas al ejecutar una pequeña ventana sobre una imagen, como se hace para el filtrado espacial. La ventana del detector está diseñada para calcular los cambios de intensidad en la imagen. Hay interés en tres escenarios:

- (1) Áreas de cero (o pequeños) cambios de intensidad en todas las direcciones, lo que sucede cuando la ventana está ubicada en una región constante (o casi constante), como en la ubicación A en la figura 3.39.
- (2) Áreas de cambios en una dirección, pero ninguno (o pequeño) cambios en la dirección ortogonal, lo que sucede cuando la ventana abarca un límite entre dos regiones, como en la ubicación B.
- (3) Áreas de cambios significativos en todas las direcciones, una condición que sucede cuando la ventana contiene una esquina (o puntos aislados), como en la ubicación C.

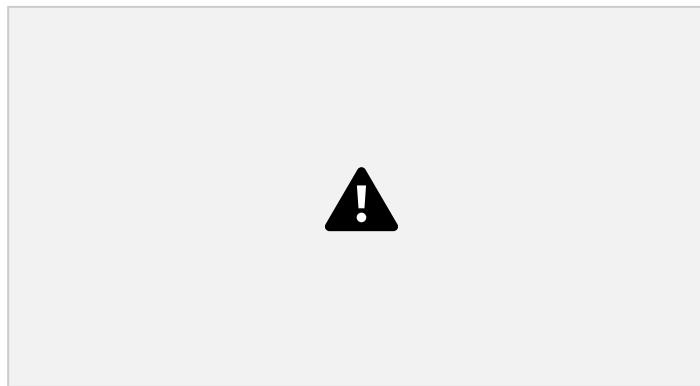


Figura 3.39 Detector de esquinas de Harris-Stephen (demonstración) A) sin esquina, plano, B) contorno, indica posible esquina si se sigue, C) esquina encontrada.

El detector de esquina de HS es una formulación matemática que intenta diferenciar entre estas tres condiciones. Sea f una imagen y sea $f(s, t)$ una parte de la imagen definida por los valores de (s, t) . Un parche del mismo tamaño, pero desplazado por (x, y) , está dado por $f(s + x, t + y)$. Entonces, la suma cuadrática media ponderada entre las dos gráficas está dada por

$$\text{P}(s, t) = \sum \sum P(s, t)[P(s + x, t + y) - P(s, t)]^2 \quad (3.20)$$

Donde $P(s, t)$ es una función de ponderación. El parche desplazado se puede aproximar mediante los términos lineales de una expansión de Taylor

$$\Phi(\Phi + \Psi, \Phi + \Psi) \approx \Phi(\Phi, \Phi) + \Phi\Psi\$ (\Phi, \Phi) + \Phi\Psi\% (\Phi, \Phi) \quad (3.21)$$

Donde $\Phi\$ (\Phi, \Phi) = \Phi\Phi\Phi / \Phi\Phi\Phi$ y $\Phi\% (\Phi, \Phi) = \Phi\Phi\Phi / \Phi\Phi\Phi$ ambas evaluadas en (Φ, Φ) . Se puede escribir la ecuación 3.20 como

$$\begin{aligned} \Phi(\Phi, \Phi) &= \sum \sum \Phi(\Phi, \Phi) 4\Phi\Phi\Phi\$ (\Phi, \Phi) + \\ &\Phi\Phi\Phi\% (\Phi, \Phi) 5! \end{aligned} \quad \# \quad (3.22)$$

De forma matricial:

$$\Phi(\Phi, \Phi) = [\Phi \Phi] \Phi 8^{\Phi} \Phi 9 \quad (3.23)$$

Donde

$$\Phi = \sum_{\#} \sum_{\#} \Phi(\Phi, \Phi) \Phi \quad (3.24)$$

y

$$\begin{aligned} \Phi &= ; \Phi\$! \Phi\$ \Phi\% \\ &\Phi\$ \Phi\% \Phi\$! < \end{aligned} \quad (3.25)$$

La matriz M se denomina matriz de Harris. Se entiende que sus términos son evaluados en (s, t) . Si $w(s, t)$ es isótropo, entonces M es simétrico porque A lo es. La función de ponderación $w(s, t)$ utilizada en el detector de HS generalmente tiene una de dos formas:

- (1) Es 1 (true) dentro del parche y 0 (false) en cualquier otro lugar (es decir, tiene la forma de un núcleo de filtro de paso bajo de caja), o
- (2) Es una función exponencial de la forma

$$\Phi(\Phi, \Phi) = \Phi^{& \# !} (\# !) / \# ! \quad (3.26)$$

La caja se utiliza cuando la velocidad computacional es primordial y el nivel de ruido es bajo. La forma exponencial se usa cuando el suavizado de datos es importante.

Como se ilustra en la figura 3.39, una esquina se caracteriza por valores grandes en la región C,

en ambas direcciones espaciales. Sin embargo, cuando el parche se extiende por un límite, también habrá una respuesta en una dirección. La pregunta es: ¿Cómo podemos saber la diferencia? Los vectores propios de un real, matriz simétrica (como M arriba) apuntan en la dirección de máxima dispersión de datos, y los valores propios correspondientes son proporcionales a la cantidad de dispersión de datos en la dirección de los vectores propios. De hecho, el vector propio es el eje mayor de la elipse que ajusta los datos, y el tamaño del valor propio es la distancia desde el centro de la elipse hasta la intersección del eje mayor.

La figura 3.40 ilustra cómo podemos usar estas propiedades para diferenciar entre los tres casos en el que estamos interesados.

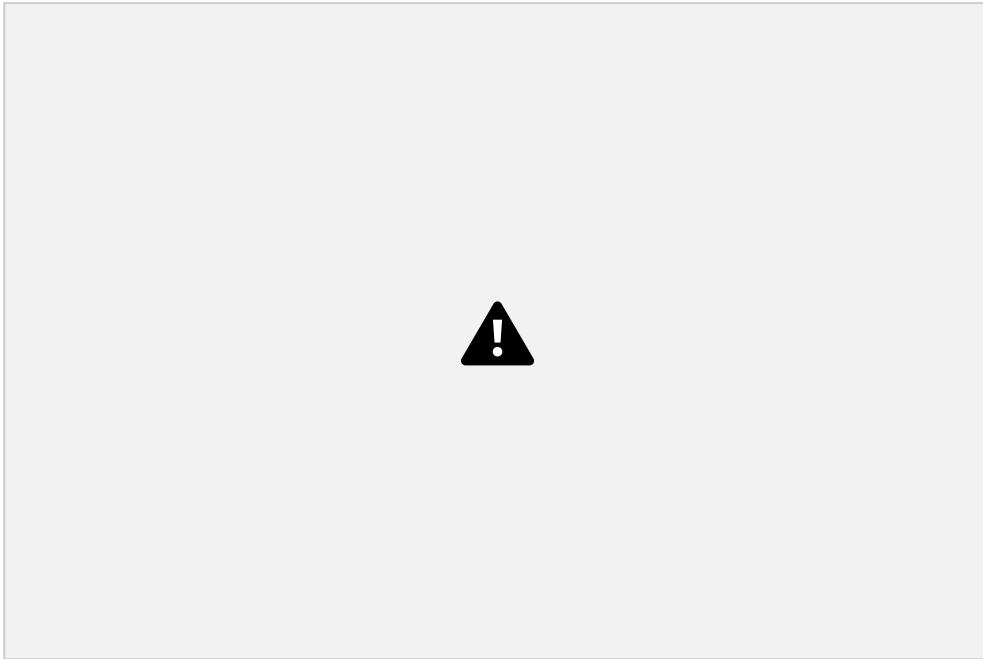


Figura 3.40

Imágenes con ruido procesadas con parches pequeños donde se comparan los pares $(\Phi_\nabla^1, \Phi_\nabla^2)$ mostrando las características de los eigenvalores de M empleados para encontrar esquinas. Los pequeños parches amarillos de la imagen en la figura 3.40 son representativas de las regiones A, B y C de la figura 3.39. En la abajo a la izquierda, se muestran valores de $(\Phi_\nabla^1, \Phi_\nabla^2)$ calculados usando los núcleos derivados $\Phi_\nabla^1 = [-1 0 1]$ y $\Phi_\nabla^2 = [0 1 0]$ (recuerde, usamos el sistema de coordenadas definido en la figura 3.40). Como calculamos las derivadas en cada punto del parche, las variaciones causadas por el ruido dan como resultado valores dispersos, estando la dispersión de la dispersión directamente relacionada con el nivel de ruido y sus propiedades. Como era de esperar, las derivadas de la región plana forman un grupo casi circular, cuyos valores propios son casi idénticos, produciendo un ajuste casi circular a los puntos (etiquetamos estos valores propios como "pequeño" en relación con las otras dos imágenes). La figura abajo en medio muestra las derivadas del parche que contiene el contorno. Aquí, la dispersión es mayor a lo largo del eje x, y aproximadamente casi lo mismo que la figura izquierda en el eje y. Por lo tanto, el valor propio Φ_∇^1 es "grande" mientras que Φ_∇^2 es "pequeño". En consecuencia, la elipse que se ajusta a los datos se alarga en la dirección x. Finalmente, la figura abajo a la derecha muestra los derivados del parche que contiene la esquina. Aquí la información se comparte entre ambas partes, lo que da como resultado dos valores cada vez más grandes, una curva circular. De esto concluimos que:

- (1) Dos valores pequeños propios indican una intensidad casi constante; (2) Un valor propio pequeño y uno grande implican la presencia de un límite vertical u horizontal;
- (3) Dos valores propios grandes implica la presencia de una esquina o (lamentablemente) puntos brillantes aislados.

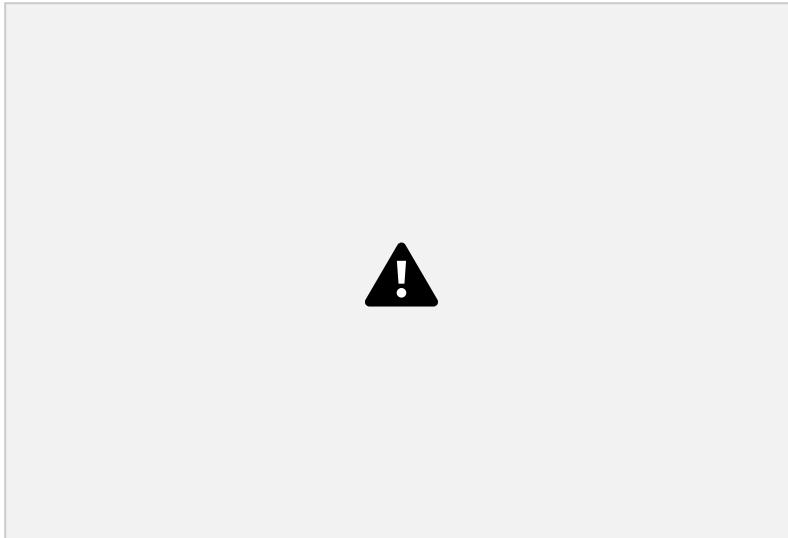


Figura 3.40 Coordenadas utilizadas para representar imágenes digitales. Porque los valores de las coordenadas son enteros, no es correspondencia de uno a uno entre x e y y las filas (r) y columnas (c) de una matriz.

Por tanto, vemos que los valores propios de la matriz formada a partir de las derivadas en el parche de la imagen se pueden utilizar para diferenciar entre los tres escenarios de interés. Sin embargo, en lugar de usar los valores propios (que son costosos de calcular), el detector HS emplea las mediciones de ángulos se utilizan asumiendo que el orden de una matriz cuadrada es igual a la suma de sus valores propios y que su exponente es igual al producto de los valores propios. La medida se define como

$$R = \frac{\lambda_1^2 + \lambda_2^2 - 2\lambda_1\lambda_2\cos(\theta) + (\lambda_1\lambda_2)^2}{\lambda_1\lambda_2} = \frac{2\lambda_1^2 + 2\lambda_2^2 - 2\lambda_1\lambda_2\cos(\theta)}{\lambda_1\lambda_2} = \frac{2\lambda_1^2 + 2\lambda_2^2}{\lambda_1\lambda_2} - \frac{2\lambda_1\lambda_2\cos(\theta)}{\lambda_1\lambda_2} = \frac{2}{k} - \frac{\cos(\theta)}{k} \quad (3.27)$$

donde k es una constante ($0 < k \leq 0.25$). La medida R tiene valores positivos grandes cuando ambos autovalores (eigen valores) son grandes, indicando la presencia de una esquina; tiene grandes valores negativos cuando un valor propio es grande y el otro pequeño, lo que indica un contorno; y su valor absoluto es pequeño cuando ambos valores propios son pequeños, lo que indica que el parche de imagen en consideración es plano. Puede interpretar a la constante k como un "factor de sensibilidad"; cuanto más pequeño es, más probable es que el detector encuentre esquinas. Por lo general, R se usa con un umbral, T ($0 < T \leq 0.3$). Decimos que se ha detectado una esquina en una ubicación de imagen solo si $R > T$ para un parche en esa ubicación.

La figura 3.41 muestra un ejemplo de uso del detector de esquinas

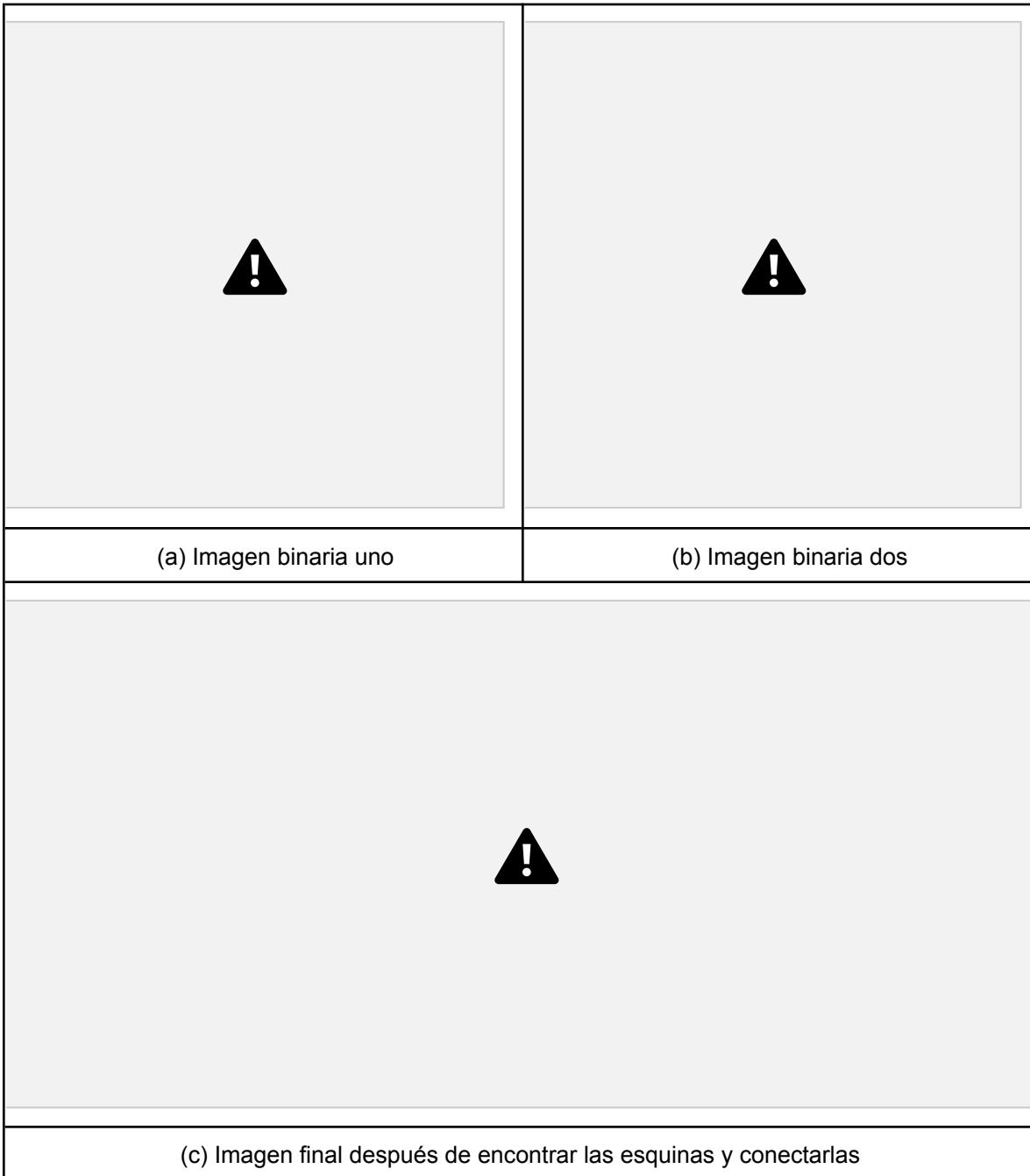


Figura 3.41 Algoritmo de Harris para encontrar esquinas y conectarlas.

O si lo que quiere es solo saber dónde están los puntos de una esquina esto se ve en las imágenes de la figura 3.42.



(a) Imagen RGB

(b) Imagen Y



(c) Realce reflectancia de Gauss

(d) Harris 1795 Puntos de esquina.

(e) Harris 1692 puntos de esquina	(f) Igualdad de (d) == (e)
	Figura 3.42 (a) imagen original de 512 x 512, (b) Transformación RGB => YIQ, (c) Imagen realizada por reflectancia de Gauss, (d) Puntos de esquina de la imagen b, (e) Puntos de esquina de la imagen c, (f) Comparación de igualdad de d y e, (g) Distinto entre d y e.
(g) Diferente de (d)!=(e)	