

2 Formación de las imágenes

Los sistemas de Visión Artificial se encuentran con escenas visuales de información en el espacio continuo y sin embargo el desarrollo y diseño de los equipos se sustenta en tecnología digital. Luego se evidencia el paso de información continua en el tiempo a discreta. La captación de la información presente en el espacio continuo, así como la forma en que ésta va a ser alterada, como consecuencia de la transformación a datos discretos, requiere ser estudiada con cierta profundidad para comprender las limitaciones que impone el sistema de formación de imágenes en su conjunto.

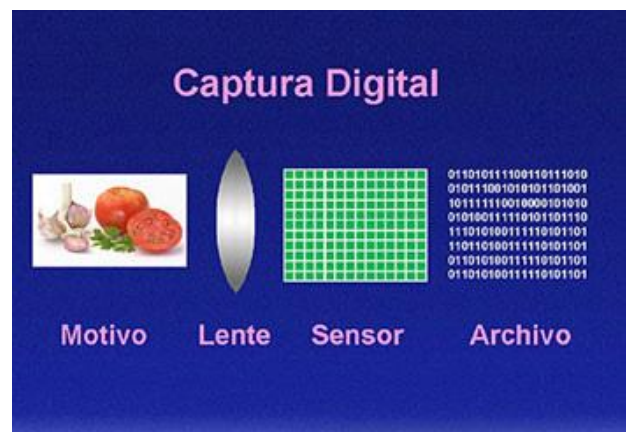


Figura 2.1 Captura digital: desde el espacio continuo al discreto

La figura 2.2 muestra un esquema de adquisición de escenas. El motivo es iluminado de manera óptima, tal que aumente el contraste entre el fondo y los objetos de interés; la luz reflejada o transmitida pasa a través de un sistema óptico, adecuando la escena al tamaño del elemento sensor. El transductor está constituido por elementos sensores que convierten los haces luminosos recibidos en una señal eléctrica. La cámara

de estado sólido da un formato a la imagen recibida en señal eléctrica pudiendo ser de tipo digital o analógico. Ésta es transmitida hasta el computador, ya sea para su monitorización o procesamiento. En el caso de utilizar señal de vídeo analógica se requiere de la conversión a formato digital. Sea como fuese, se habrá generado una correspondencia entre la realidad y lo almacenado en el computador.

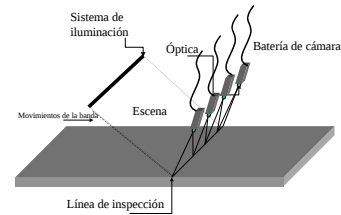


Figura 2.2. Elementos constituyentes de formación de imágenes

Se pretende abordar las capacidades de la adquisición y de formación de las escenas visuales. Como se ha podido observar, las posibilidades de estos equipos dependen de una multitud de factores. El diseño de una aplicación adecuada requiere del análisis de todos estos factores que influyen en las prestaciones, buscando, además, las fuentes de degradación en la adquisición de las imágenes. Estas últimas pueden ser eliminadas y/o compensadas mediante la selección apropiada de soporte físico, compromisos de medidas y técnicas de calibración. El exacto conocimiento de los factores que influyen en la calidad de la imagen permitirá con posterioridad la selección correcta de los componentes.

El desarrollo de este capítulo va a tratar de cómo afecta los diferentes elementos del sistema de formación de imágenes a la métrica virtual, centrándose especialmente en el subsistema de iluminación, en la óptica y en la selección de arquitecturas de cámaras de estado sólido. Los motivos son que éstos representan los elementos más críticos en la calidad de la imagen y por ende en la métrica. El capítulo finalizará con algunas técnicas de calibración de las cámaras.

2.1 Caracterización de imágenes digitales

La calidad de una imagen es una impresión subjetiva que va desde pobre hasta excelente. La habilidad de la percepción realizada por el cerebro viene afectada no sólo por el sistema de visión humano, sino también por otros factores como son: las emociones, el aprendizaje y la memoria. Las relaciones entre éstas son complejas, intrincadas y no comprendidas. Los juicios personales varían ante una misma imagen y las valoraciones pueden ser muy dispares.

Muchas expresiones han sido dadas para categorizar una imagen, cada una es apropiada según un particular conjunto de condiciones. Estas expresiones suelen ser obtenidas desde las observaciones empíricas, donde un grupo de imágenes son ordenadas y posteriormente se infiere alguna ecuación que justifique la clasificación.

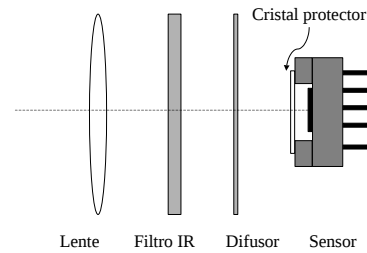
En un principio, la calidad de la imagen estaba relacionada con las lentes. Con el avance de la tecnología, la calidad se ha centrado en la minimización de la relación señal ruido, SNR, y en la arquitectura de las cámaras. Muchos tests han dado medidas sobre la calidad de la imagen. La mayoría de las medidas están relacionadas con la calidad de las ópticas, la resolución de la cámara o la relación señal ruido del elemento sensor. En general, las imágenes con altas prestaciones ópticas y bajo ruido son clasificadas como imágenes de gran calidad. Las prestaciones del sistema de visión

dependen de las cámaras, lentes, iluminación, filtros ópticos, etc. Sólo al final se puede plantear la calidad de la imagen.

Los más importantes factores que afectan a la calidad geométrica y radiométrica de una imagen adquirida por una cámara de estado sólidos son los siguientes:

- Iluminación
- Superficie a visualizar
- Óptica
- Cámaras
- Transmisión y digitalización de las imágenes
- Entorno

Figura 2.3. Partes constitutivas de una cámara



La selección de una apropiada iluminación es una pieza clave en las tareas de procesamiento de las imágenes. La uniformidad de luminosidad en la superficie, la estabilidad temporal y su respuesta espectral, así como las propiedades ópticas del entorno deben ser consideradas. La falta de homogeneidad de intensidad luminosa en la escena puede acarrear que cientos de píxeles aparezcan con bajo contraste. Por la extensión y la importancia en el buen funcionamiento de los equipos de visión, se ha dedicado un apartado para abordar el estudio de las fuentes luminosas.

El segundo aspecto se refiere a la naturaleza del producto a inspeccionar. El tipo, el color, el tamaño y la forma del detalle a analizar obviamente influyen en las exactitudes de las medidas. Éstas dependen, sobre todo, del tamaño del objeto de estudio y de su contraste.

Los aspectos referidos a las ópticas, sensores, cámaras y digitalizadoras van a ser abordados con mayor detalle y representan el núcleo del capítulo. Se plantearán los aspectos de focalización de la radiación luminosa que llegue al elemento sensor y cómo se modela, para pasar posteriormente al estudio del funcionamiento del sensor y de la composición de las cámaras de vídeo.

Por último, respecto al entorno, hay que citar la temperatura, la humedad, las vibraciones y las iluminaciones ambientales. Por ejemplo, el efecto de la temperatura puede producir profundos cambios en las cámaras. Hay multitud de cámaras que requieren más de dos horas antes de alcanzar su estado estable. Estos cambios provocan errores de gran magnitud. Sus causas se encuentran en las deformaciones mecánicas, en la sensibilidad de elemento sensor con la temperatura y en las variaciones en la frecuencia del reloj interno de la cámara.

2.2 Iluminación

La iluminación juega un papel primordial dentro del proceso de visión. Los sistemas de iluminación ayudan a simplificar considerablemente el análisis e interpretación de las escenas. Muchas veces, los algoritmos de procesamiento pueden ser reducidos drásticamente con una selección adecuada del sistema de formación de las imágenes. Por el contrario, la falta de estudio sobre la adquisición de las imágenes provocará el aumento de la complejidad de los algoritmos y las dificultades de generalización del equipo de Visión Artificial. Las imágenes recibidas por las cámaras no pueden estar, generalmente, en un ambiente de iluminación no estructurada, esto es, donde las condiciones de iluminación no hayan sido estudiadas. En estos casos, la propia sombra de otro objeto que interfiera hará peligrar el éxito del montaje.

Los sistemas de Visión Artificial trabajan, normalmente, con iluminaciones estructuradas. El espacio captado por las cámaras está aislado del resto. No hay contaminación lumínica, no hay sombras inesperadas, no existen objetos que interseccionen a los escenarios que se desean capturar. La iluminación se diseña para que sea capaz de resaltar los objetos de interés. La iluminación depende básicamente de la naturaleza de las superficies a inspeccionar y del tipo de fuente luminosa.

Normalmente, la primera reducción consiste en tratar los objetos como si fuesen opacos. La radiación transmitida no es considerada. En Computación Gráfica se estudia la interacción entre la luz y las superficies y cuánta radiación le llega al sensor. En una primera aproximación se estudia la interacción entre la fuente luminosa y los cuerpos. Sin embargo, estas iteraciones no son suficientes para crear imágenes sintéticas creíbles. Así, en el movimiento de los objetos sobre un escenario no se producirían cambios de iluminación o sombras. Estas consecuencias se derivan de no tener en cuenta las radiaciones de los cuerpos existentes en la escena. Una de las técnicas de renderización más empleadas se llama *Ray Tracing*. Consiste en lanzar rayos de luminosidad y ver cómo interactúan con las superficies de los objetos que hay en el escenario, tomando en cuenta la interacción entre todos ellos.

El problema en Visión Artificial es el inverso de la Computación Gráfica. No se trata de generar una imagen sintética, sino de ubicar correctamente la iluminación, los sensores y los cuerpos a examinar para resaltar las características visuales de los objetos.



Figura 2. 4 a) Dibujo sin interacción con la luz, b) Gráfico con interacción entre la luz y la materia

2.2.1 Modelos de superficies opacas

La manera en la que la luz es reflejada en una superficie depende de entre muchos factores, de las formas macroscópicas que constituyen las superficies. De esta manera, en superficies lisas al ser iluminadas con un haz incidente, la luz reflejada sobre ellas toma una sola dirección; en cambio, sobre superficies rugosas tienden a difuminarse en múltiples direcciones. Para establecer con exactitud el fenómeno de la reflexión es necesario tener un conocimiento a priori de las estructuras de las superficies, es decir, un modelo de éstas. Se pueden catalogar en dos clases de modelos de superficie: mapas de perfiles y modelos aleatorios. Un exacto perfil puede ser determinado con las medidas de alturas en cada punto de la superficie, por algún tipo de dispositivo existente en el mercado. No obstante, este método es caro e impracticable en muchos casos. Por lo tanto, suele modelarse mediante funciones de densidad estadísticas que tienden aproximarse al verdadero comportamiento de la superficie ante un haz de luz incidente.

Una superficie puede ser modelada a partir de una colección de microsuperficies planas, tal como queda reflejado en la figura 2.5. Cada microsuperficie tiene su propia orientación. El ángulo de inclinación, α , es el parámetro que determina el grado de oblicuidad respecto a la normal del plano medio. En el caso de superficies isotrópicas, la probabilidad de distribución de α sería uniforme y las microsuperficies estarían rotadas respecto a la normal, n , con igual probabilidad en todas las direcciones. Esto lleva a que puede modelarse la superficie mediante una función de distribución unidimensional. Por ejemplo, la superficie puede responder a una función de densidad gaussiana para el nivel de microfacetas con $N(\mu_\alpha, \sigma_\alpha^2)$. Grandes valores en la varianza, σ_α^2 , indicarán la presencia de superficies rugosas, y al contrario, valores pequeños señalarán superficies lisas. Los modelos basados en la función de distribución del ángulo de inclinación en las microsuperficies son muy populares, pues la difusión de la luz reflejada es explicada por los niveles de inclinación y no por la cartografía (altura) de la superficie. Por esta razón, los modelos de inclinación, aunque relativamente ambiguos, son ampliamente utilizados en los problemas de reflexión de la luz en superficies.

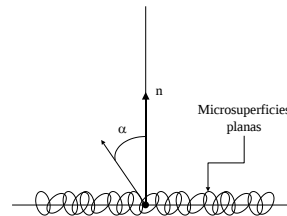


Figura 2. 5. Superficie formada por microfacetas

La impresión visual de los humanos es bastante vaga con el concepto de rugosidad. De manera que una superficie vista de cerca parece rugosa y algo más alejada, puede llevar a la conclusión de calificarla como lisa. De hecho, los humanos hacen cambios de vistas para ayudarse en el juicio de la rugosidad de la superficie.

A diferencia de las valoraciones humanas, la teoría de reflexión tiene bien definida el concepto de rugosidad; ésta debe relacionar las irregularidades de la superficie con las direcciones que tome la luz ante un haz de luz incidente. Para superficies cuyas irregularidades son menores que la longitud de onda de la luz, $\sigma_\alpha \ll \lambda$, la mayor parte del flujo incidente será reflejada especularmente en una sola dirección. En caso contrario, si las irregularidades superan la longitud de onda del haz, la

superficie difuminará la luz en múltiples direcciones, produciendo una radiación difusa. En la práctica, la radiación reflejada es una combinación de ambos casos.

El modelo más utilizado en Visión Artificial se basa en las conclusiones de Nayar. La reflexión de la luz sobre una superficie está formada por tres componentes: la primera corresponde con la reflexión difusa explicada por el modelo lambertiano. La segunda justifica el halo especular que se fundamenta en el modelo de Torrance-Sparrow y la última se emplea en superficies lisas y hace referencia al pico especular.

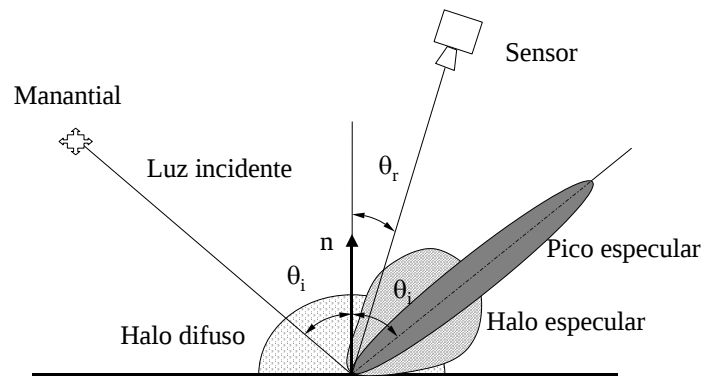


Figura 2. 6 Tipos de radiación al iluminar una superficie según el modelo de Nayar

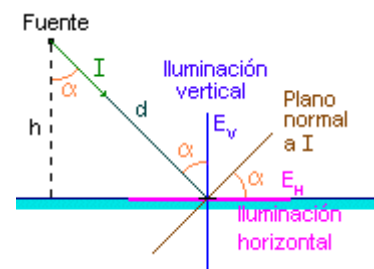
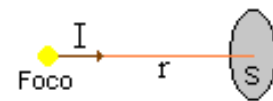
2.2.2 Cálculo de la iluminación y la luminancia

La iluminación sobre una superficie, tal como se vio en el capítulo anterior, viene determinada por la intensidad del foco luminoso y por las posiciones geométricas entre foco y superficie. Para el caso de intensidad isotrópica, puntual y de posición en el eje axial de la superficie respecto del foco, la iluminación es igual a:

$$E = \frac{d\phi}{dA} = \frac{I \cdot d\omega_r}{dA} \cong \frac{I}{r^2}$$

Cuando la iluminación no es paraxial hay que descomponerla en una componente horizontal y en otra vertical a la superficie:

$$\text{Iluminancia} = \frac{\text{Intensidad luminosa}}{\text{Cuadrado distancia}}$$



$$E_H = \frac{I \cdot \cos(\alpha)}{d^2}$$

$$E_V = E_H \tan(\alpha)$$

En general, si un punto está iluminado por más de un foco, su iluminancia total es la suma de las iluminancias recibidas:

$$E_H = \sum_i \frac{I_i \cdot \cos(\alpha_i)}{d_i^2}$$

$$E_V = \sum_i E_{H_i} \tan(\alpha_i)$$

Ejemplo 2.1

Una superficie está iluminada por una fuente luminosa puntual de 0.1 W/sr de intensidad constante en todas direcciones situada a 2 m de altura. Calcular la iluminancia de superficie y la iluminación vertical, si la inclinación entre la superficie y la fuente, α , es de 30° .

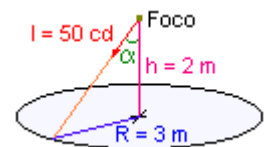
$$E_H = \frac{I \cos^3 \alpha}{h^2} = 16.2 \text{ mW} / \text{m}^2$$

$$E_V = E_H \tan \alpha = 9.4 \text{ mW} / \text{m}^2$$

$$|E| = 18.72 \text{ mW} / \text{m}^2$$

Ejemplo 2.2

Una superficie circular de 3 m de radio está iluminada por una bombilla de 50 cd de intensidad constante en todas direcciones situada a 2 m de altura sobre el centro de la plataforma. Calcular la iluminación máxima y mínima sobre la superficie.



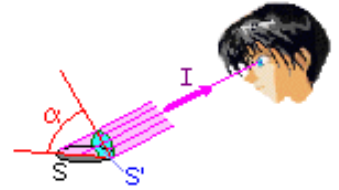
La iluminación máxima se dará en la vertical del foco:

$$E_{\max} = \frac{I}{h^2} = 12.5 \text{ lx}$$

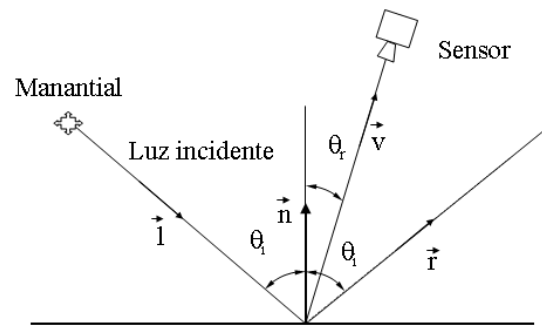
La mínima se dará en la frontera:

$$E_{\min} = \frac{I \cdot \cos(\alpha)}{(h / \cos(\alpha))^2} = 2.13 \text{ lx}$$

Hasta ahora se ha hablado sobre la iluminación radiada, pero para medir la recibida por un observador se emplea el concepto de luminancia o radiación. Ésta se definió como la relación entre la intensidad radiada por una fuente, en una dirección determinada, por la superficie aparente. Es ésta la que llega al ojo o a la cámara y no la iluminación.



La reflexión de la luz sobre una superficie rugosa se aproxima a un comportamiento de reflexión en todas las direcciones. Una superficie que refleja la luz de esta forma se la llama lambertiana. Para explicar el modelo de interacción entre la luz y la materia se definen los vectores unitarios $\vec{h}$, $\vec{l}$, $\vec{v}$ y $\vec{r}$, según muestra la figura. La explicación del halo difuso se formaliza con:



$$I_d = \rho_d I_d^{in} (\vec{h} \cdot \vec{l})$$

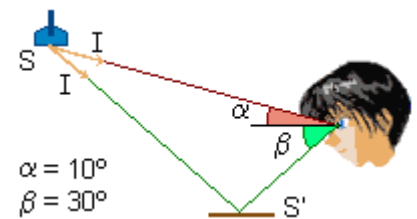
Siendo ρ_d el coeficiente de reflexión difuso adimensional e I_d^{in} es la intensidad incidente en la superficie. En cambio para superficies lisas, la reflexión es de carácter especular y su modelo es:

$$I_s = \rho_s I_s^{in} (\vec{r} \cdot \vec{v})^k$$

Donde ρ_s el coeficiente de reflexión especular adimensional, I_s^{in} es la intensidad incidente en la superficie y k es un parámetro del modelo para disminuir el efecto especular cuando se desvía del vector $\vec{r}$. Normalmente se considera $k=100$.

Ejemplo 2.3

Se tiene un proyector situado en el techo de 0.04 m^2 de superficie que ilumina con una intensidad de 150 mW/sr en cualquier dirección. Una mesa de 0.5 m^2 de superficie recibe dicha iluminación y tiene un comportamiento lambertiano de factor de reflexión de 0.8 . Calcular la luminancia de la fuente y la luminancia de la mesa para el observador de la figura.



$$L_p = \frac{I}{S_p (\vec{h}_p \cdot \vec{v}_p)} = \frac{I}{S_p \sin \alpha} = 21.59 \text{ W} / \text{m}^2 \cdot \text{sr}$$

$$L_m = \frac{\rho I}{S_m (\vec{h}_m \cdot \vec{v}_m)} = \frac{\rho I}{S_m \sin \beta} = 0.48 \text{ W} / \text{m}^2 \cdot \text{sr}$$

2.2.2.1 Modelo de Phong de Reflexión de la luz

Este modelo local es utilizado en Computación Gráfica para aplicaciones en tiempo real. Considera los dos mecanismos de reflexión de la luz en superficies: difusa y especular. En cuanto a la iluminación, las clasifica en difusa, especular y ambiental. Su forma de proceder es descomponer la luz en sus distintas formas de onda, normalmente utiliza las franjas RGB y calcula la interacción de cada parte del espectro con la materia. Con posterioridad, al aplicar superposición, se suma las interacciones de cada franja y se obtiene la intensidad en cada punto del escenario.

Para facilitar el cálculo computacional, no se considera la interacción propia entre cuerpos próximos entre sí al radiar unos sobre otros. Este término es aproximado por una iluminación ambiental, $\rho_a I_a^{in}$. La intensidad de emisión de un punto para una determinada longitud de onda es calculada como:

$$I = I_a + I_d + I_s + I_e = \rho_a I_a^{in} + \rho_d I_d^{in} \left(\frac{\mathbf{r} \cdot \mathbf{l}}{|\mathbf{r}| |\mathbf{l}|} \right) + \rho_s I_s^{in} \left(\frac{\mathbf{r} \cdot \mathbf{v}}{|\mathbf{r}| |\mathbf{v}|} \right)^k + I_e$$

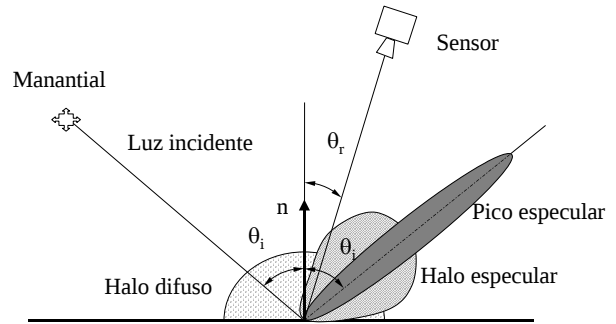
donde I_e es la componente de intensidad correspondiente a la emisión de luz o fuente de iluminación. Este cálculo es realizado sólo para cuerpos opacos. Es posible extender el modelo a cuerpos translúcidos, modelizando la refracción difusa y especular mediante la ley de Snell.

Ejemplo 2.4

Una fuente luminosa de 0.1 W/sr ilumina a una superficie. Los factores de reflexión difusa y especular sobre esta superficie son $\frac{1}{\sqrt{3}}, 0.5$, respectivamente.

Los ángulos de incidencia y de colocación de la cámara son $\theta_i = \frac{\pi}{4}, \theta_r = \frac{\pi}{6}$, respecto a la

normal de la superficie. Determinar la intensidad recibida en el sensor. Considérese que no hay interferencias de otras superficies y su emisión es nula. Utilizar el modelo de Phong. Dato: $k=100$.



$$I_d = \rho_d I_d^{in} \left(\frac{\mathbf{r} \cdot \mathbf{l}}{|\mathbf{r}| |\mathbf{l}|} \right) = \frac{1}{\sqrt{3}} 100 \cos\left(\frac{\pi}{4}\right) = 40.82 \frac{mW}{sr}$$

$$I_s = \rho_s I_s^{in} \left(\frac{\mathbf{r} \cdot \mathbf{v}}{|\mathbf{r}| |\mathbf{v}|} \right)^k = \frac{1}{2} 100 \cos\left(\frac{\pi}{4} - \frac{\pi}{6}\right)^{100} = 1.56 \frac{mW}{sr} \quad I = 42.38 \frac{mW}{sr}$$

2.2.3 Técnicas de iluminación

Existen muchas técnicas de iluminación y resulta difícil de generalizar sobre los distintos tipos de iluminación, ya que suele depender de cada caso concreto. Sin embargo, se pueden clasificar en cuatro grandes grupos: direccionales, difusas, a contraluz y estructuradas. Se presentan éstas y posteriormente se dará un repaso a combinaciones de las iluminaciones principales.

2.2.3.1 Iluminación direccional

Consiste en dar una iluminación orientada al objeto empleando un haz altamente direccional o colimado. Se utiliza en la localización y reconocimiento de piezas, inspección de superficies, seguimiento de puntos para soldadura o pegamento, etc. La forma del haz luminoso depende de la aplicación y del objeto a analizar. Cabe destacar los siguientes casos: a) Punto luminoso b) Línea de luz c) Plano de luz y d)

Corona circular
luminosa.

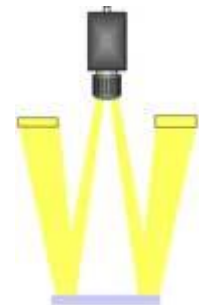
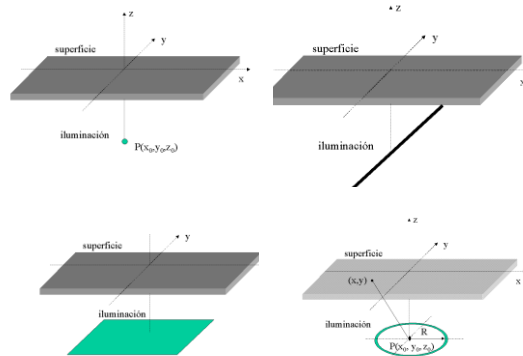


Figura 2.7
Técnica
direccional

2.2.3.2 Iluminación difusa.

En ella se intenta que los haces luminosos incidan sobre el objeto desde todas las direcciones y no sólo desde la fuente luminosa. Proporciona el mínimo contraste del objeto y elimina las sombras. Se utiliza cuando los objetos a analizar presentan superficies suaves.

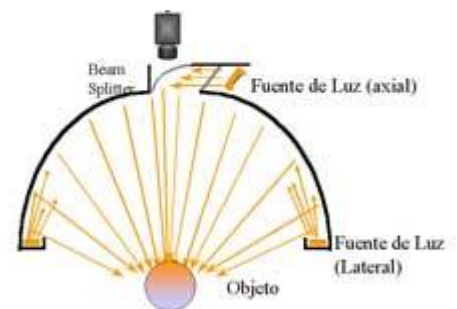


Figura 2. 8. Técnica difusa

Para crear una iluminación difusa se utilizan difusores de color blanco mate.

2.2.3.3 Iluminación a contraluz.

Consiste en colocar la fuente luminosa por detrás del objeto, de manera que la información capturada por el sensor proviene del flujo luminoso

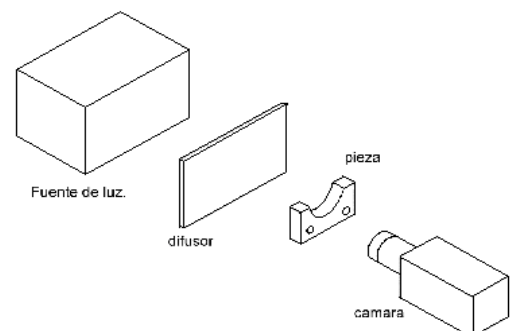


Figura 2. 9 Técnica a contraluz

transmitido por el cuerpo. Si el objeto es opaco, éste aparecerá como una sombra y las imágenes se constituyen básicamente en sólo dos niveles de grises. Se emplea preferentemente para la localización de piezas y análisis dimensional. El principal inconveniente es la pérdida de los detalles de los cuerpos. Una preocupación a tener en cuenta es el posible deslumbramiento de la cámara, producido por la cantidad de radiación que alcanza al sensor, lo que origina un recorte en la dimensión de los objetos.

En caso de cuerpos translucidos dejará información sobre las propiedades internas de la materia. Éste es el fundamento de la tomografía o de la radiografía. La imagen capturada dependerá de la absorción de la materia. En este ejemplo se observa que la fuente de iluminación no tiene necesariamente que pertenecer al espectro visible.

2.2.3.4 Iluminación estructurada.

Este tipo de iluminación se sirve de la proyección de puntos, franjas o rejillas sobre la superficie de trabajo. En función de cómo se deforme este patrón de luz sobre la superficie se puede detectar las singularidades de la pieza objeto de análisis. Este método posee dos ventajas con respecto a las anteriores.

1. Establece un patrón de luz conocido sobre la superficie de trabajo y las diferencias con este patrón indican la presencia de un objeto, simplificándose así el problema de detección de un objeto.
2. Analizando la forma en que el patrón de luz es distorsionado, es posible obtener información de las características tridimensionales del objeto.

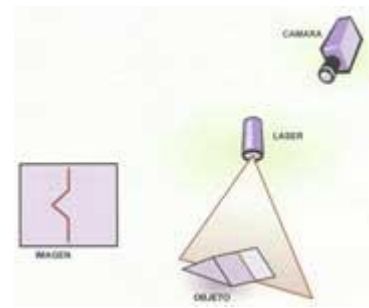


Figura 2. 10. Técnica de luz estructurada

2.2.4 Luz polarizada

Los haces de luz no polarizada consisten en radiaciones electromagnéticas con vibraciones transversales de igual magnitud en un número infinito de planos perpendiculares a la dirección de desplazamiento. Algunos materiales actúan selectivamente absorbiendo la energía en ciertos ángulos y transmitiéndola en otros. El resultado de un polarizador es un haz de luz que, en esencia, parece vibrar en una sola dirección (polarizador total), o vibra en todas direcciones pero con distinta magnitud (polarizador parcial). El estado de polarización depende de la relación existente entre las fases y amplitudes del campo eléctrico y magnético.

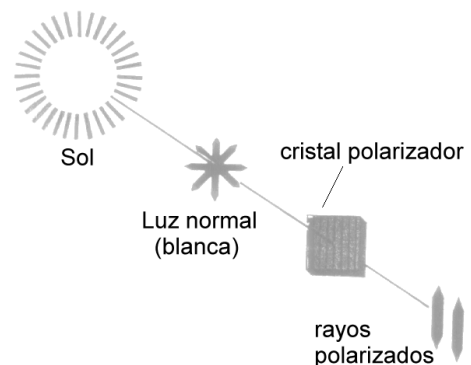


Figura 2. 11. Luz polarizada

La percepción de un objeto es el resultado de la reflexión difusa y especular de la luz incidente sobre éste. La reflexión especular es acromática y está parcialmente polarizada, mientras que la reflexión difusa no está polarizada y es la que al actuar selectivamente sobre las distintas longitudes de onda dan la sensación de color. Para obtener imágenes de una superficie coloreada habrá de emplear un filtro polarizador que impida la reflexión especular, polarizada y acromática, haciendo llegar al sensor sólo la reflexión difusa.

Se suele emplear dos polarizadores, uno en la iluminación y otro en la cámara. Generalmente se emplean $0^\circ/90^\circ$ y $45^\circ/135^\circ$.

2.2.5 Ejemplos de iluminaciones.

A continuación se enumeran algunas iluminaciones utilizadas en Visión Artificial, mezclas de las anteriores. Se hace una breve descripción así como sus ventajas e inconvenientes.

1. **Iluminación Frontal Direccional:** Es la técnica más empleada. La cámara es posicionada mirando al objeto en el mismo plano que la iluminación. La cámara recibe la luz reflejada del objeto. **Ventajas:** especialmente útil para observar superficies, fácil de implementar y con posibilidad de utilizar la fibra óptica. **Desventajas:** Puede crear reflejos no deseadas.

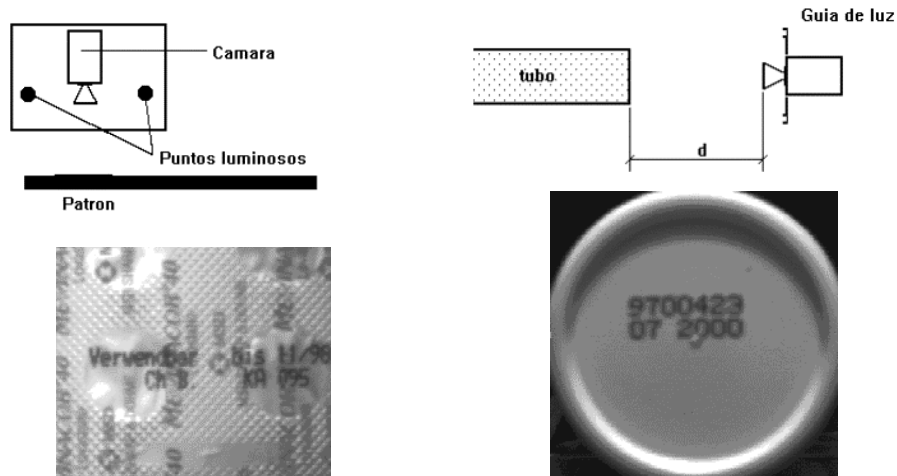


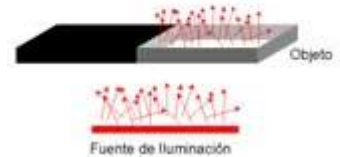
Figura 2. 12 Ejemplos de iluminación frontal empleado en la inspección de etiquetas

2. **Iluminación Frontal Difusa:** Usada para generar iluminación “superior” (la fuente está más elevada que la escena) . **Ventajas:** Es suave, regular no direccional, reduce el brillo en superficies metálicas, relativamente fácil de implementar. **Desventajas:** Los bordes pueden quedar difuminados y el contraste es bajo en zonas monocolor.

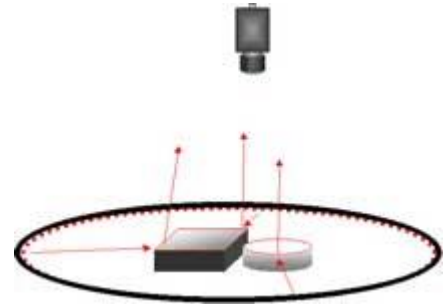
3. **Iluminación de Tienda (día nublado):** No direccional, totalmente difusa, iluminación superior que produce iluminación como la de un día nublado. Bueno para piezas metálicas (bujías, cojinete de bolas) y componentes electrónicos. **Ventajas:** Elimina brillos y sombras. **Desventajas:** Debe rodear la pieza, puede ser costoso, el tamaño puede ser un problema.



4. **Iluminación Colimada Posterior:** Iluminación posterior a través de una lente colimada, de modo que los rayos de luz son casi paralelos. **Ventajas:** Produce bordes muy contrastados para mediciones de precisión. **Desventajas:** Difícil de implementar si el material se mezcla en la dirección del haz y puede ser demasiado brillante para cámaras sin filtro.

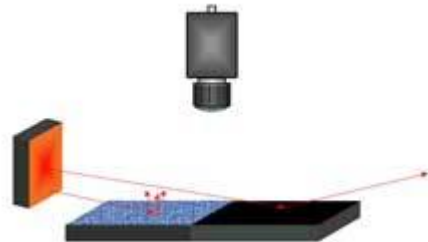


5. **Iluminación de Campo Oscuro:** La luz incidente se refleja lejos de la cámara y la iluminación es creada por reflexión especular. **Ventajas:** Ilumina defectos y produce una imagen de alto contraste en algunas aplicaciones. **Desventajas:** no ilumina superficies planas.



6. **Contraluz Difusa:** La luz está en el lado opuesto de la cámara y va a través de un material difusor, tal como papel de estraza o vidrio translúcido. **Ventajas:** Fácil de implementar, crea silueta del objeto, imagen de muy alto contraste y bajo coste. **Desventajas:** los bordes pueden quedar difuminados, puede ser difícil de implementar dependiendo de la posición del material.

7. **Iluminación de Bajo Ángulo:** La luz incide casi horizontal a la superficie del objeto. **Ventajas:** Muestra defectos topológicos, **Desventajas:** Una sola fuente producirá una iluminación desigual en la superficie.



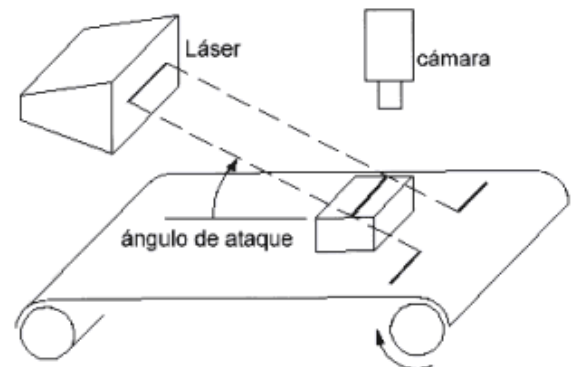
8. **Iluminación Frontal Polarizada:** Iluminación frontal con un polarizador en la luz y un polarizador cruzado en la lente. **Ventajas:** elimina brillos. **Desventajas:** reduce significativamente la cantidad de luz a través de la lente.

9. **Contraluz Polarizado:** El polarizador y el polarizador cruzado están en lados opuestos del objeto. **Ventajas:** alta iluminación en ciertos tipos de

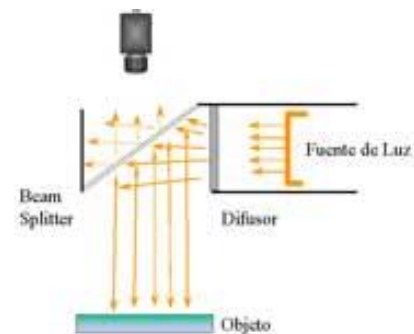
características o defectos en materiales translúcidos y relativamente fáciles de implementar. **Desventajas:** los bordes pueden quedar difuminados y difícil de implementar para algunas posiciones del material.

10. **Iluminación Estroboscópica:** Una iluminación periódica que dura microsegundos se usa para congelar la imagen de objetos en movimiento. **Ventajas:** toma imágenes sin movimiento, puede ser de área, fibra óptica o led. **Desventajas:** Más costoso que las fuentes usuales, requiere una precisa sincronización con la cámara y debe estar fuera del alcance visual de las personas.

11. **Iluminación Estructurada:** Un plano de luz generado por una luz estructurada mediante conversión de lente cilíndrica de punto a línea de láser. **Ventajas:** muestra información 3D, produce alto contraste en la mayoría de las zonas. **Desventajas:** el láser por encima de 5 mW es peligroso, la luz puede ser absorbida por algunas superficies.



12. **Iluminación Coaxial:** La iluminación va en la misma dirección que la luz que entra por la cámara reflejada por los objetos, esto es, la iluminación y el eje de la cámara son colineales. La fuente de luz está colimada (todos los haces luminosos van en la misma dirección). **Ventajas:** elimina sombras e iluminación uniforme. **Desventajas:** complicado de implementar y áspera iluminación para superficies lustrosas.



2.2.6 Fuentes luminosas.

Entre los elementos comerciales que se pueden utilizar para una correcta iluminación son: luces incandescentes, focos halógenos, fluorescentes, láseres y leds.

Las luces incandescentes presentan un coste bajo, son fáciles de usar y además pueden estar funcionando durante largos periodos de tiempo. Éstas suelen ser puntos luminosos, no direccionales y con baja eficiencia energética. Generalmente presentan la posibilidad de ajustar la intensidad lumínica, aspecto importante, ya que interesa que el área efectiva de la lente sea lo más pequeña posible, por lo que para estimular al sensor se requiere mayor potencia lumínica. El principal inconveniente es el elevado calor que desprenden, producto de su funcionamiento, y que puede limitar su aplicación.

Cuando las fuentes luminosas emplean corriente alterna, la radiación luminosa varía de intensidad y puede provocar la variación de una imagen a otra. Por eso, a veces, se emplean fuentes luminosas que puedan ser alimentadas con corriente continua, como así sucede con las lámparas halógenas. Estas bombillas tienen una temperatura menor que las incandescentes (350°C).



Figura 2. 13. Fluorescente tipo corona

Los fluorescentes son fuentes de luz que se caracterizan por no generar calor. Presentan formas y tamaños variables, así como de diversos colores de luz. Son comúnmente empleados en iluminación difusa. La alimentación debe ser en corriente alterna y para evitar fluctuaciones en las imágenes se emplean reactancias de altas frecuencias (de 100Hz a 25kHz).

El láser es empleado en iluminación estructurada. Es una fuente monocromática que puede presentar diferentes patrones de luz: punto, línea, rejillas, ... Su gran inconveniente es la necesidad de emplear sistemas de seguridad, para evitar daños en los operarios que trabajen cerca de estos sistemas de iluminación.



Figura 2. 14. Luz láser

El uso de fibra óptica, como guía de luz, está asociado a la transmisión de la iluminación pero no del calor, por lo que suele ir asociado a fuentes incandescentes. Su aplicación está principalmente en la iluminación de pequeñas áreas y concavidades, como por ejemplo en la endoscopia.

Las luces estroboscópicas son ideales para analizar objetos en movimiento o piezas en una cinta transportadora. Dan una iluminación intensa, por lo que la iluminación ambiental influye poco. Existen elementos con pulsos de 2 microsegundos. El mayor inconveniente es el proceso de sincronización entre la iluminación y la cámara.



Figura 2. 15. Guía de luz (fibra óptica)

2.2.7 Aspectos a considerar en las fuentes de iluminación

- Variación de la iluminación en pequeños periodos de tiempo: alimentación continua, alimentación alterna en alta frecuencia y sincronización.
- Variación de la iluminación en largos periodos de tiempo: arranque y autocalibración.
- Temperatura y disposición física.
- Fuentes luminosas de alta precisión.
- Suciedad.
- Carácter colimado del haz.
- Mantenimiento, coste y facilidad de recambio.

2.3 La óptica

La óptica es el elemento del sistema de visión que permite concentrar los rayos luminosos del entorno sobre el plano sensor. Una vez definido el sistema de iluminación, la elección de la óptica influirá en la calidad de la imagen recibida y del tamaño de los objetos. En este apartado se hace un repaso de los parámetros ópticos con un modelo de lente simple. Posteriormente, se explicarán varias cuestiones prácticas sobre las ópticas comerciales empleadas en las cámaras de vídeo.

2.3.1 Cámara oscura

Desde hace mucho tiempo es conocida la experiencia de la cámara oscura. Cardan, en 1550, empleó la cámara oscura con una lente. Se trata de realizar en un habitáculo cerrado, en una de sus paredes, un pequeño orificio. La luz, tras entrar en la cámara, se proyecta sobre la pared opuesta a la que tiene el orificio. Por cada rayo procedente del exterior se obtiene un punto luminoso en el plano proyectivo. Extendiendo el razonamiento a los infinitos rayos que atraviesan el agujero se obtiene una imagen invertida y proyectiva del mundo exterior al plano sensor.

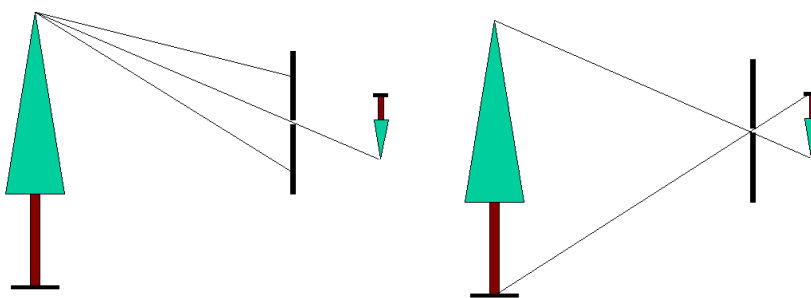


Figura 2. 17. Fundamento de la cámara oscura

También se observa que a medida de que la separación entre las paredes de la cámara aumente, la imagen proyectiva se incrementa de tamaño. A este modelo de cámara oscura se la denomina *pin-hole*.

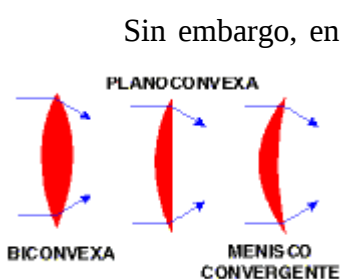


Figura 2. 18. Tipos de lentes convergentes o positivas

Sin embargo, en la cámara oscura, la energía incidente que atraviese no es lo suficientemente elevada como para excitar al plano sensor. Por lo que se resuelve aumentando el tamaño del agujero. El efecto provocado es una degradación de la imagen, ya que son varios los rayos los que inciden, pero no necesariamente del mismo objeto, produciéndose una imagen borrosa. Estos inconvenientes son subsanados con el uso de las lentes.

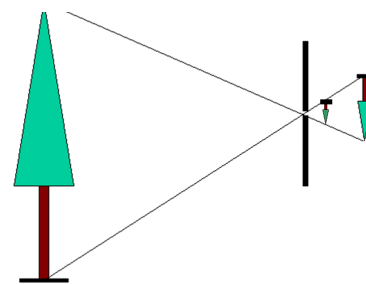


Figura 2. 16 Tamaño del objeto proyectado

2.3.2 Lente simple

Con el objeto de aumentar la luminosidad en el interior de la cámara se utiliza una lente, sin que la imagen se haga borrosa al aumentar el tamaño del agujero. Una lente es una sustancia transparente y refringente, limitada por dos caras, una de las cuales es curva y la otra puede ser plana o curva, pero con la condición de que sus centros de curvatura están en el mismo eje. El efecto deseado es la convergencia o divergencia del frente de ondas.

La figura representa una lente convexa¹ de radio R . El centro de curvatura de la lente se define por C . Sea un punto P exterior, situado a una distancia Z . Los infinitos rayos procedentes del emisor que chocan sobre la lente producirán una dirección especular, respecto a la normal de la superficie de la lente. La refracción del frente de ondas, gracias a la curvatura constante, se concentra en el punto P' , separado a una distancia z , respecto a la superficie. La relación entre estas distancias con el radio de curvatura viene definida por la ecuación de Gauss:

$$\frac{1}{Z} + \frac{1}{z} = \frac{2}{R}$$

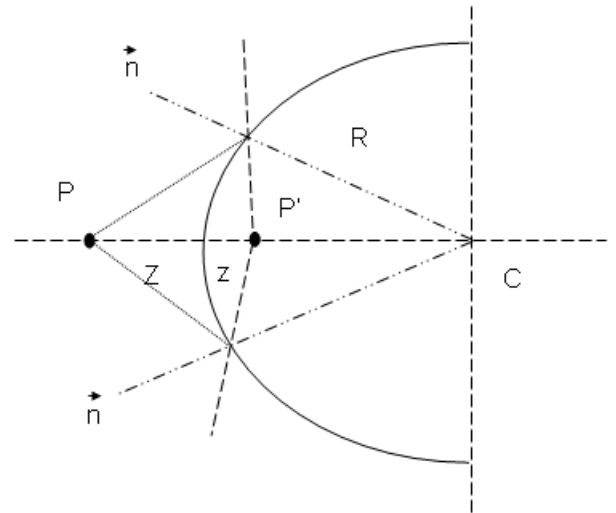


Figura 2. 19 Modelo de lente delgada

En la formación de imágenes se emplean tanto lentes convergentes o positivas como divergentes o negativas. Las lentes positivas tienen la propiedad de hacer converger en un punto los distintos rayos divergentes que proceden de un mismo lugar del objeto exterior. Las lentes divergentes hacen el trabajo al contrario, no focalizan la luz, sino que las hacen divergir.

Al emplear lentes convergentes, todos los puntos del plano imagen tienen, por tanto, la característica de que todos los rayos incidentes en cada uno de dichos puntos, provienen de un mismo punto del objeto, obteniéndose una imagen nítida con una mayor energía incidente.

¹ Una superficie es convexa cuando el centro de curvatura pertenece a la figura geométrica.

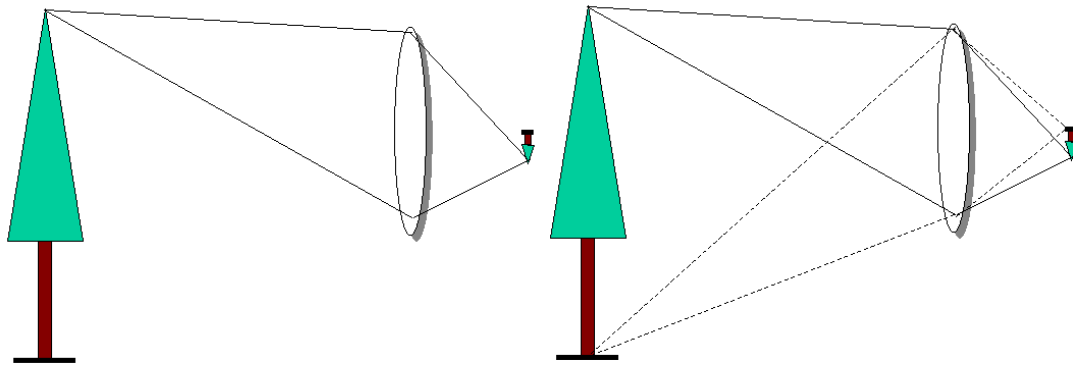


Figura 2. 20. Modelo de lente delgada

La zona sensible de la cámara deberá estar situada en el plano de convergencia de los rayos, en caso contrario, los rayos incidentes sobre el sensor provendrán de distintos objetos, obteniéndose una imagen borrosa o desenfocada.

2.3.3 Distancia focal

Se observa que, para un mismo objeto exterior, a medida de que éste se encuentre más alejado respecto a la cámara, el punto de convergencia de los rayos, al otro lado de la óptica, disminuye, ya que el grado de convergencia de la lente permanece constante y la radiación incidente tiene menor divergencia (ver figura 2.22). Al procedimiento de ajustar la distancia entre la lente y el plano sensor, con el objeto de tener una imagen nítida, se le denomina enfoque.

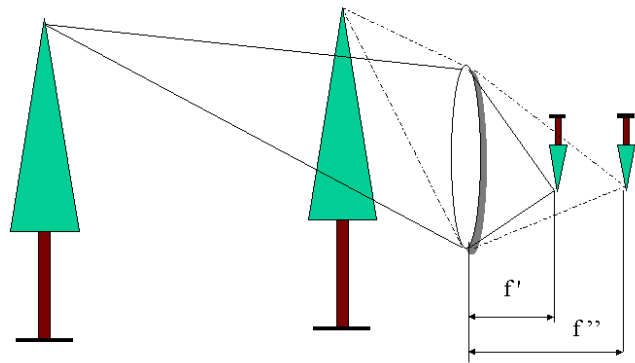


Figura 2. 22. Distancia de enfoque dependiendo de la situación de los objetos

Dependiendo de la curvatura de las lentes, éstas poseen distintos grados de convergencia de los rayos provenientes del objeto y por tanto será distinta la distancia entre el plano de la imagen y la lente. Con el fin de medir el grado de convergencia de una lente se utiliza la distancia lente-plano de imagen. Pero debido a que esta distancia es variable según la distancia a la que está enfocado el objeto, su medición se realiza cuando el objeto se encuentra enfocado a una distancia infinita delante de la cámara; en cuyo caso, los rayos

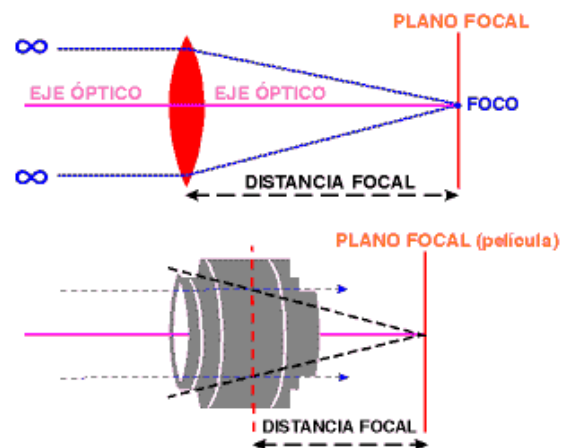


Figura 2. 21. Distancia focal

procedentes del mismo punto del objeto serán paralelos. La distancia focal, f , se define como la distancia entre la lente y el plano sensor o focal cuando se encuentra enfocado hacia el infinito. Esta definición equivale a decir que la distancia focal es la distancia entre la lente y el punto de convergencia de todos los rayos paralelos que inciden perpendicularmente a ella. A este punto de convergencia se le denomina foco de la lente. Si el objeto se encuentra a una distancia finita, la separación entre la lente y el plano focal será mayor que la distancia focal.

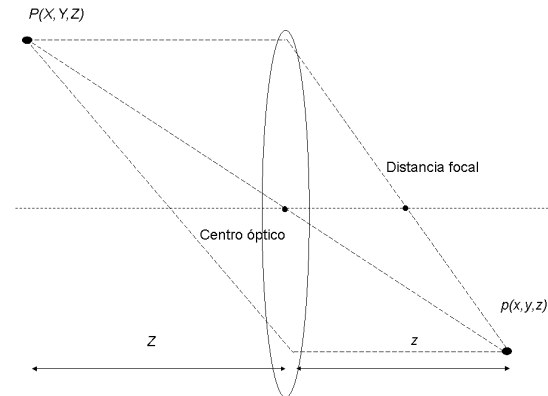


Figura 2. 23. Modelo de lente delgada

Para explicar la formación de la imagen se emplea el modelo de lente delgada. De forma que todos los rayos que pasen por el centro óptico continúan en línea recta y los rayos que atraviesen la lente paralelamente al eje axial de la óptica convergen en el foco. La proyección de un punto P del exterior, de coordenadas (X,Y,Z) , sobre el plano sensor, p de posición (x,y,z) vendrá definido, según la ecuación de Gauss, por:

$$\frac{1}{Z} + \frac{1}{z} = \frac{1}{f} \quad (2.1)$$

La lente delgada puede ser simplificada si se emplea el modelo *pin-hole*. Este modelo, como se ha comentado, reduce la óptica a un punto de convergencia situado a la distancia focal, f . Con esta suposición, se considera despreciable el movimiento de convergencia de los rayos a través del proceso de enfoque.

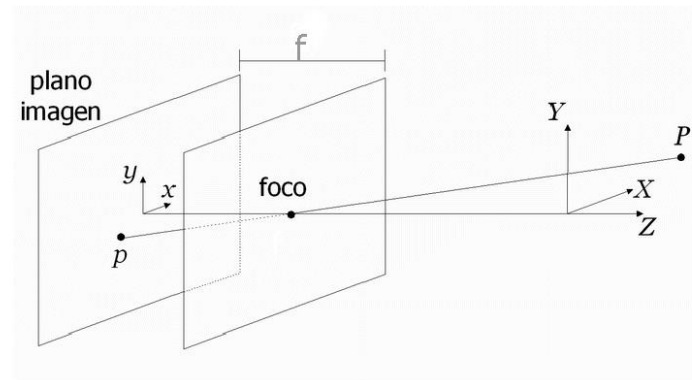


Figura 2. 24 Modelo pin-hole

Por tanto, es como si la óptica estuviera siempre enfocada hacia el infinito. La ventaja es la facilidad de modelado del sistema óptico. Por triangulación, la relación entre la geometría exterior y la proyectada mostrada en la figura es:

$$x = \frac{f}{Z} X \quad y = \frac{f}{Z} Y \quad (2.2)$$

Se denomina factor de magnificación a la relación entre el segmento proyectado sobre el plano imagen y el tamaño exterior del segmento físico:

$$M = \frac{x}{X} = \frac{y}{Y} = \frac{f}{Z} \quad (2.3)$$

En el modelo *pin-hole*, la distancia focal es proporcional al factor de magnificación, mientras en el modelo de lente fina es:

$$f = \frac{M}{M+1} Z \quad (2.4)$$

En cualquier caso, la distancia focal da una medida del grado de magnificación entre la imagen real y la proyectada. El tamaño de la imagen proyectiva depende de la distancia focal. A mayor distancia focal, menor ángulo visual. Una elección adecuada de la distancia focal depende de la resolución deseada del objeto. El grado de resolución espacial del sistema de adquisición de las imágenes vendrá determinado por el tamaño del objeto más pequeño que pueda distinguirse, sin llegar a confundirlo. Se analizará más adelante al introducir el teorema de muestreo.

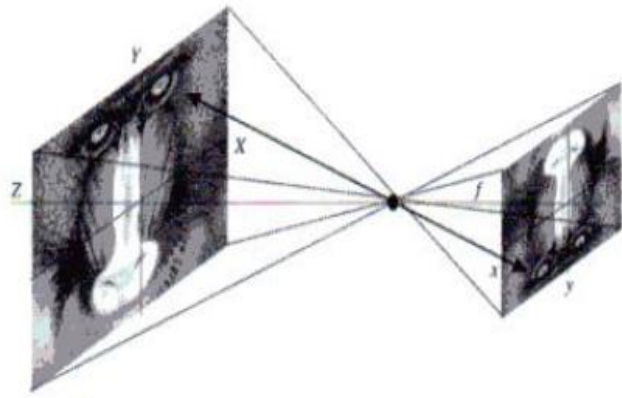


Figura 2. 25 Magnificación de la óptica

Ejemplo 2.5

Para la práctica de calibración de las cámaras se ha empleado una cuadrícula tipo de ajedrez. Los lados son de 27 mm y se ha puesto la rejilla a 1 metro de distancia respecto a la cámara. Se ha empleado una cámara de píxel cuadrado de 5.6μm. Las aristas de las caras se ven en 15 píxeles. ¿Cuál debe ser la distancia focal de la óptica?. ¿Qué área se visualiza, si la cámara está constituida por 357x293 píxeles?. Utilícese el modelo *pin-hole*.



$$f = \frac{15 \cdot 5.6 \cdot 10^{-6}}{27 \cdot 10^{-3}} \cdot 1 = 3.1 \text{ mm}$$

$$S = X \cdot Y = \frac{x \cdot y}{M^2} = \frac{357 \cdot 5.6 \cdot 10^{-6} \times 293 \cdot 5.6 \cdot 10^{-6}}{(3.1 \cdot 10^{-3})^2} = 0.341 \text{ m}^2$$

Ejemplo 2.6

¿Qué tipo de espejo se requiere para iluminar una pared situada a 3 m del foco luminoso, sabiendo que la separación entre la fuente y el espejo es de 10 cm? ¿Cuál es la proyección del foco de 5 mm sobre la pared?

El tipo de espejo debe de ser divergente, esto es, su curvatura es cóncava y tendrá un radio de curvatura de:

$$\frac{1}{3} + \frac{1}{0.1} = \frac{2}{R} \quad R = 19.4cm$$

El tamaño del foco proyectado será:

$$M \cong \frac{0.1}{3} \quad X = \frac{x}{M} = 5 \cdot 10^{-3} \cdot 30 = 150mm$$

Ejemplo 2.7

Un pequeño objeto se encuentra a 4 cm de un espejo cóncavo de radio 12cm. Calcular la posición y aumento de la imagen nítida.

$$\frac{1}{0.04} + \frac{1}{z} = \frac{1}{0.06} \quad z = 0.12m \quad M = \frac{0.12}{0.04} = 3$$

Ejemplo 2.8

Un objeto es proyectado por una esfera de vidrio hueca de radio R, 10cm, a una distancia respecto de la superficie de 3R, ¿A que distancia respecto del centro de la esfera se obtendrá la imagen nítida y cuanto ha sido ampliada?

La imagen invertida dentro de la esfera, al pasar por la zona convexa valdrá:

$$\frac{1}{3R} + \frac{1}{z} = \frac{2}{R} \quad z = \frac{3R}{5} = 6cm$$

Esta imagen al ser pasada por la zona cóncava será ampliada y la imagen nítida respecto del centro de la esfera será:

$$\frac{1}{Z'} + \frac{1}{7R/5} = \frac{2}{R} \quad Z' = \frac{7R}{9} = 7.78cm \rightarrow d = 17.78cm$$

El factor de magnificación será atenuante:

$$M \cong \frac{3R/5}{3R} \cdot \frac{7R/5}{7R/9} = \frac{9}{25}$$

2.3.4 Ángulo visual

El tamaño máximo de los objetos que se proyectan dentro del área sensible del sensor viene determinado por el área del sensor, por las distancias de los objetos a la cámara y por la distancia focal, f . Con el objeto de poder comparar los ángulos visuales de las distintas cámaras, se efectúa la medida con el enfoque hacia el infinito.

Con este supuesto, se calcula el ángulo visual cómo el ángulo que forman dos rayos que inciden sobre los bordes externos de la zona sensible de la imagen, cuando la imagen se encuentra enfocada a una distancia infinita. Puesto que la zona sensible es normalmente rectangular, su tamaño no es lo mismo en todas las direcciones, pudiéndose hablar del ángulo visual horizontal, vertical, así como del ángulo visual según la diagonal. Normalmente, los sensores de vídeo son caracterizados por su diagonal, D , por tanto, el ángulo visual se expresa en función de este valor.

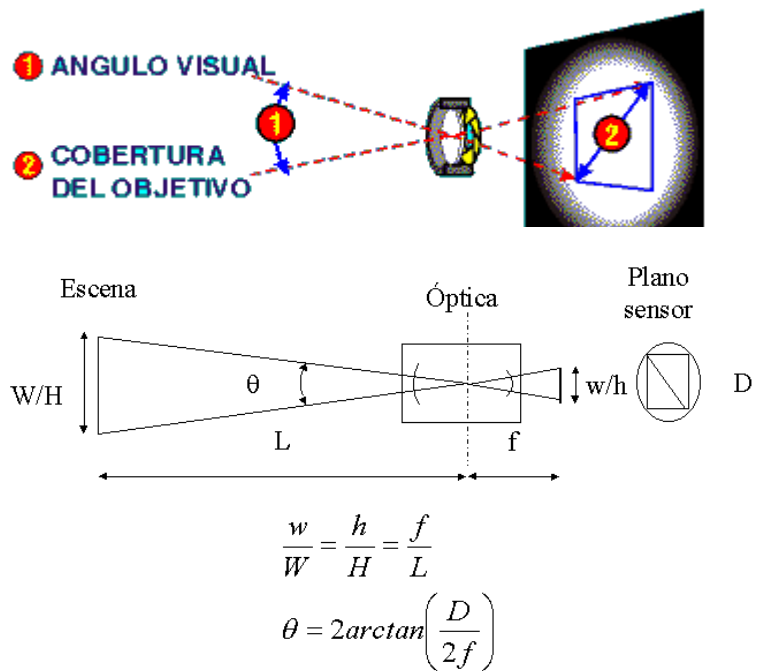


Figura 2. 26. Ángulo visual de la óptica

Se observa de la expresión que el ángulo visual disminuye con el aumento de la distancia focal de la lente.

2.3.5 Diafragma

El área efectiva de la lente va a influir en las características de la imagen proyectada en la zona sensor. Para poder modificar el área efectiva de la lente, sin tener que cambiarla, se emplea el diafragma. El diafragma es una superficie opaca, colocada en la óptica, que permite el paso de los rayos luminosos en un círculo central de la lente e impide que los rayos que inciden sobre la corona pasen. Los diafragmas se construyen de manera que la abertura central de la lente

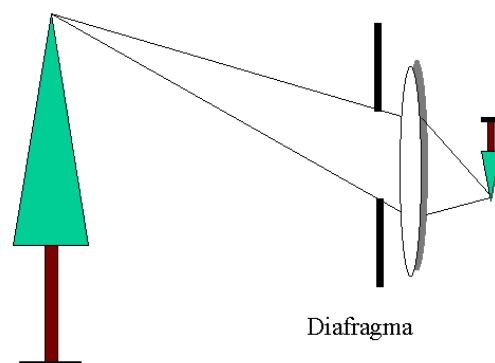


Figura 2. 27. Colocación del diafragma

sea variable, posibilitando de esta manera una graduación del tamaño efectivo de la lente.

Para una lente con una distancia focal fija, el efecto más inmediato de la variación de la abertura del diafragma se refleja en la cantidad de luz que incide sobre el plano sensor. Esta cantidad de luz es proporcional al área efectiva de la lente. Si se denomina con D el diámetro de la abertura efectiva, la energía incidente de los rayos será proporcional a D^2 . Pero la

luminosidad en el plano imagen no depende sólo de la cantidad de luz que atraviesa la lente, sino también de la distancia entre la lente y el plano sensor, la cual es inversamente proporcional al cuadrado de su distancia focal². Según estos razonamientos, la luminosidad de una lente, que expresa el flujo de luz en la superficie del área sensible, es proporcional a D^2/f^2 . Para cuantificar la luminosidad de una imagen se emplea el número F de una lente que se define como $F=f/D$. Por tanto, una disminución del número F en un valor de $\sqrt{2}$ equivale a un aumento del doble de la potencia luminosa.

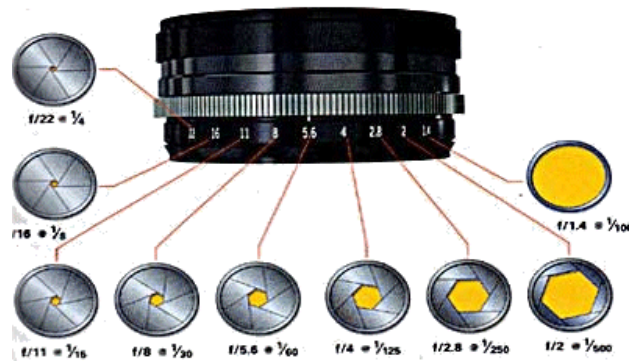


Figura 2. 28. Abertura de la lente dependiendo del número F

2.3.6 Profundidad de campo

Para que un objeto situado a una distancia L produzca una imagen nítida es necesario que la lente se encuentre situada a una determinada distancia del plano sensor. Esta distancia depende de la distancia L del objeto y de la distancia focal de la lente, f .

Cualquier objeto, situado delante o detrás del enfocado, dará lugar a que los rayos provenientes de un punto de éste no focalicen en el plano de la zona sensible. En el sensor quedará proyectado como una mancha, denominado círculo de confusión.

Si el diámetro del círculo de confusión es menor que la mitad de la resolución del sistema de adquisición de la imagen, dicho sistema ve también como un solo punto de imagen. Por el contrario, si el círculo de confusión generado por un punto fuera del foco es mayor que el valor mencionado, dicho punto aparecerá como borroso o desenfocado. A este valor crítico

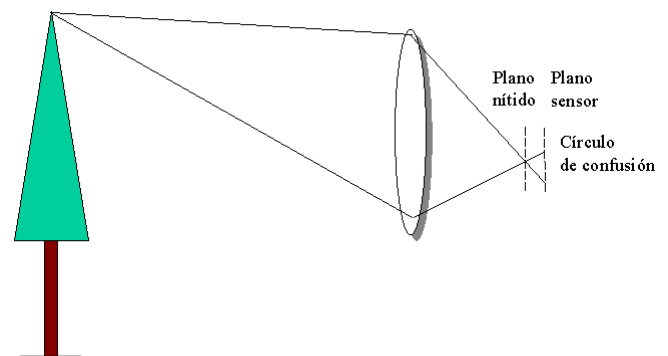


Figura 2. 29. Círculo de confusión

$$^2 \phi = I \cdot d\omega_r \cong I \cdot \frac{\pi/4 D^2}{f^2} \propto \frac{I}{F^2}$$

se le denomina círculo permisible de confusión, δ .

Cuando un punto del objeto se encuentra más cerca o más lejos de la distancia L , a la que se encuentra enfocado el sistema, pero a una distancia tal que el círculo de confusión generado sea menor del permisible, dicho punto también se encontrará enfocado.

Se denomina profundidad de campo a la zona del espacio dentro de la cual el objeto forma una imagen proyectada nítida o enfocada. Este valor viene determinado por el radio del círculo permisible de confusión, δ , la distancia focal, f , y la abertura del diafragma, F :

$$d = d_1 + d_2 = \frac{\delta \cdot F \cdot L^2}{f^2 - \delta \cdot F \cdot L} + \frac{\delta \cdot F \cdot L^2}{f^2 + \delta \cdot F \cdot L} \quad (2.5)$$

Siendo d_1 y d_2 las distancias límites de los puntos enfocados por detrás y por delante respectivamente respecto al punto de foco. La profundidad de campo, d , posee las siguientes características:

1. A mayor número F da mayor profundidad de campo.
2. A menor distancia focal da mayor profundidad de campo
3. A mayor distancia al objeto supone mayor profundidad de campo.
4. Cuanto mayor sea el radio del círculo de confusión permisible, menor resolución del sistema y tanto mayor será la profundidad de campo.
5. La profundidad de campo es mayor por detrás del objeto que por delante, $d_1 > d_2$.

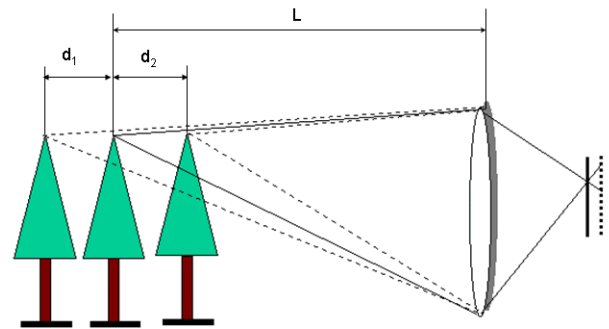


Figura 2. 30. Profundidad de campo

La primera característica supone que cuanto menor sea la abertura de la lente mayor será la profundidad de campo. La abertura de la lente hace algo más que controlar la cantidad de luz que penetra en la cámara. Tiene una segunda función muy importante, la de controlar la profundidad de campo.

2.3.7 Aberraciones y distorsión de la imagen

El término *aberraciones* designa las imperfecciones introducidas en las imágenes por los sistemas ópticos. Hasta ahora se ha supuesto que la lente hace converger los rayos divergentes de un punto del espacio en un único punto del plano sensor. Sin embargo, este hecho sólo sucede cuando la luz es monocromática y situada en las proximidades del eje óptico. Hay dos clases de distorsiones: cromáticas y geométricas. Las primeras están relacionadas con el espectro frecuencial de los rayos y las segundas aparecen si los rayos que inciden en la óptica están alejándose del eje axial de la lente.

2.3.7.1 Aberración cromática

Se debe al diferente índice de refracción de un rayo luminoso según su longitud de onda. Existen dos tipos de aberración cromática: longitudinal y lateral.

Aberración cromática longitudinal. Este tipo de aberración provoca que los rayos de distintas longitudes de onda procedentes de un punto sobre el eje óptico de la lente se enfoquen sobre diferentes planos de imagen.

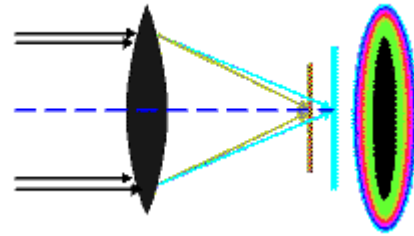


Figura 2. 31 Aberración cromática

Aberración cromática lateral. Ocurre debido a que la magnificación de la lente cambia según la longitud de onda del rayo, produciendo un desplazamiento lateral de los puntos de convergencia de los rayos de diferente longitud de onda.

2.3.7.2 Aberración geométrica

Las cinco aberraciones geométricas básicas o aberraciones de Seidel, en honor al investigador que las clasificó, son: esférica, coma, astigmatismo, curvatura de campo y distorsión. Las cuatro primeras afectan a la nitidez de la imagen y la quinta a la forma de los objetos proyectados.

Aberración esférica. Del haz de rayos que parten de un punto del objeto situado sobre el eje de la lente, aquellos que entran en la lente a una mayor altura sobre dicho eje son refractados a diferentes puntos sobre el eje óptico, produciéndose un desplazamiento del foco. La aberración esférica puede ser subsanada mediante el diafragma; cuanto mayor sea el valor de F , menor será el radio de la lente, reduciéndose

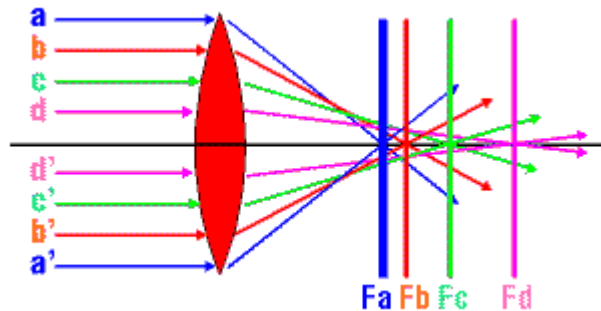


Figura 2. 32. Aberración esférica

así esta aberración.

Coma. Hasta ahora se ha supuesto que los rayos son paraxiales, pero cuando inciden con un cierto ángulo sobre el eje, éstos no son enfocados sobre el plano de imagen, sino que forman una imagen similar a un cometa con una cola. La cola tiene el aspecto de círculos que permanecen siempre tangentes a dos rectas que forman entre sí un ángulo de 60° . Esta cola puede apuntar en la dirección al

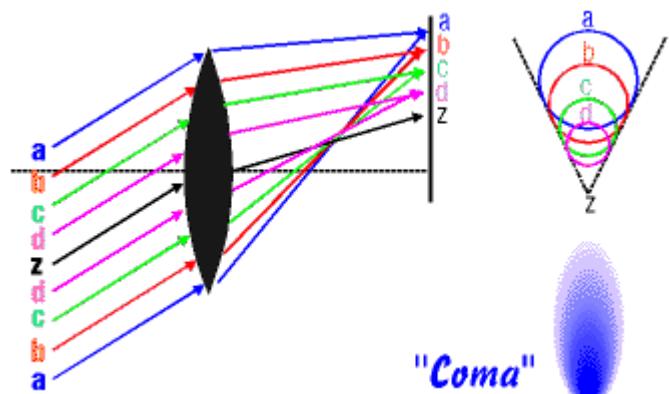


Figura 2. 33. Aberración de coma

centro de la imagen o en la opuesta, originando la coma entrante o la coma saliente. Un objeto que contenga curvas concéntricas al eje óptico sufrirá un gran desenfoque a lo largo de un borde de cada curva concéntrica. Si una lente no está corregida contra la coma, el contraste en los bordes de la imagen será malo. La coma puede ser corregida dando una menor abertura al diafragma.

Astigmatismo. En algunos objetivos ocurre que resulta imposible enfocar con nitidez simultáneamente los objetos verticales y horizontales, incluso estando ambos en el mismo plano. Este defecto se comprueba fácilmente fotografiando unas rejillas metálicas y examinando atentamente los bordes. Si el plano de enfoque se desplaza hasta encontrar las líneas verticales de forma nítida, las horizontales aparecerán borrosas y viceversa. Disminuyendo la abertura del diafragma, y por tanto aumentando la profundidad de campo, se consigue disminuir el astigmatismo hasta un cierto punto, pero no se elimina completamente.

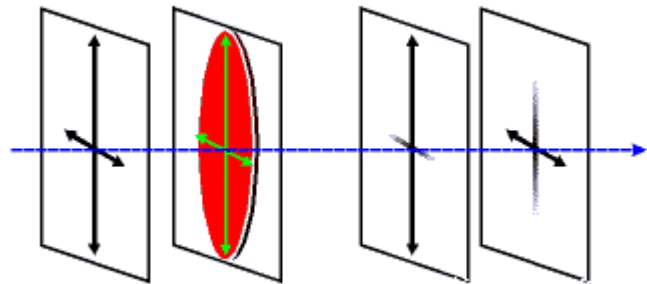


Figura 2. 34. Astigmatismo

Curvatura de campo. Representa la incapacidad de una lente para enfocar un objeto plano sobre el plano sensor. Realmente, el plano focal de un objetivo no es totalmente plano, sino que forma una superficie cóncava hacia el objetivo. Como la zona sensible está totalmente plana, es difícil enfocar a la vez y exactamente, el centro y los bordes de la imagen. Los fabricantes solucionan esta degradación, parcialmente, variando el diseño de las lentes. Como el astigmatismo, la curvatura de campo puede ser corregida disminuyendo la abertura del diafragma para aumentar la profundidad de campo.

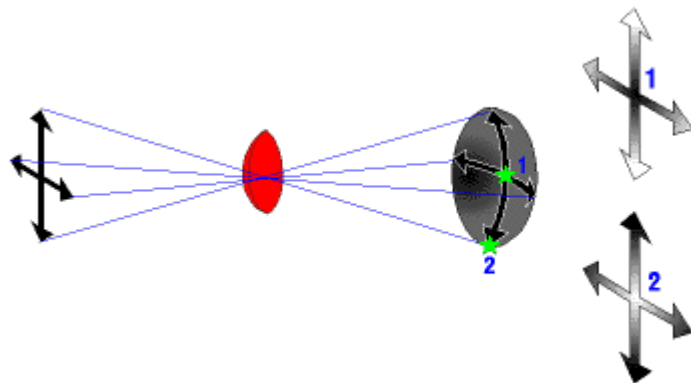


Figura 2. 35. Curvatura de campo

Distorsión. Mientras que las aberraciones mencionadas hasta ahora se relacionan con la nitidez del enfoque de la imagen, la distorsión afecta a la forma que toma los objetos proyectados en la imagen. Se acostumbra a expresar la distorsión como un porcentaje entre la altura ideal de la imagen y

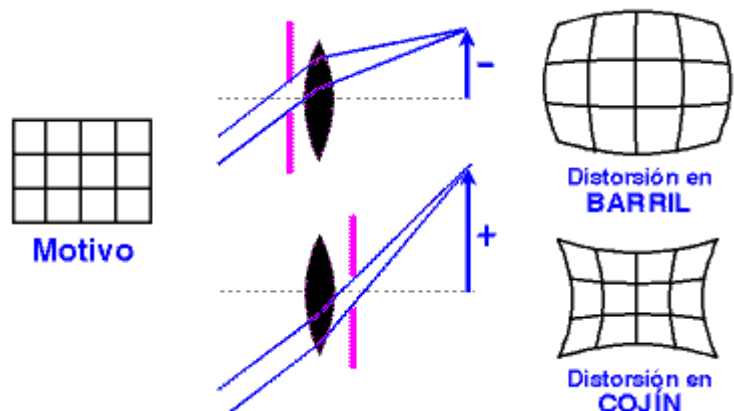


Figura 2. 36. Modelo de distorsión

la distancia que los puntos se desvían de su posición en la imagen ideal. La distorsión destruye la perfecta semejanza entre el objeto y su imagen, dando o bien un efecto de cojín (distorsión positiva) o bien un efecto de barril (distorsión negativa).

Las lentes de amplio ángulo visual, i.e. de baja distancia focal, deforman las líneas rectas tanto vertical como horizontalmente. Al fotografiar una cuadrícula, ésta puede deformarse de forma cóncava o convexa, resultando así las llamadas distorsiones de "cojín" y "barrilete".

En el modelo de la distorsión cabe considerar dos componentes: una radial y otra tangencial. Cada una de ellas se modela como una serie de infinitos términos. Sin embargo, para la mayoría de las aplicaciones de visión es suficiente con modelar sólo la distorsión radial y, de ésta, vale con los dos primeros términos de la serie. De esta forma, la relación entre los píxeles con y sin distorsión sería:

$$\left. \begin{aligned} x_d(1 + k_1 r^2 + k_2 r^4) &\cong x_u \\ y_d(1 + k_1 r^2 + k_2 r^4) &\cong y_u \end{aligned} \right\} r^2 = x_d^2 + y_d^2 \quad (2.6)$$

Donde (x_d, y_d) son las coordenadas del píxel distorsionado y (x_u, y_u) del píxel sin distorsión. El grado de distorsión viene dado por los valores de k_i . En la figura de abajo se muestra los modelos de distorsión tanto radial como tangencial, así como la combinación de ambos efectos, procedentes de la cámara empleada en las prácticas de la asignatura.

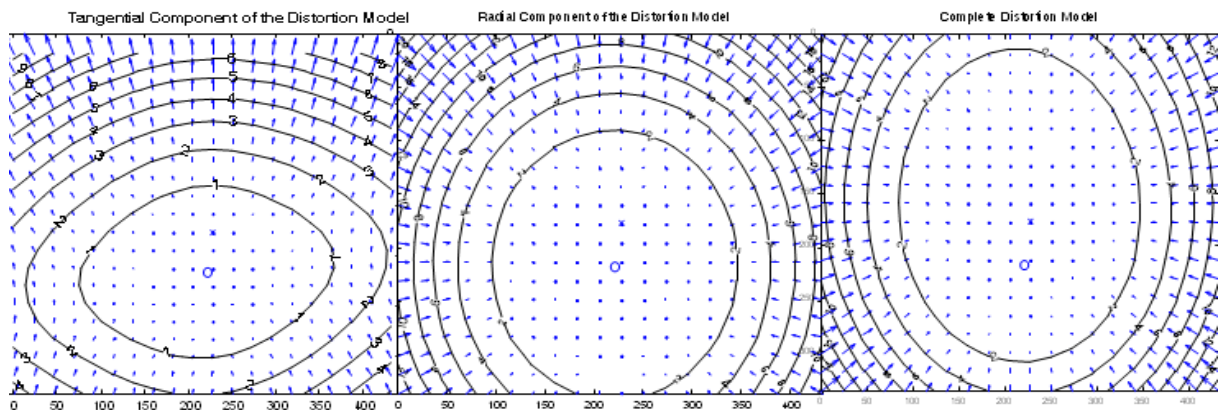


Figura 2. 37. Modelo de distorsión de la óptica de la cámara de prácticas

2.3.7.3 Efecto del diafragma sobre las aberraciones

Según se acaba de comentar, los defectos provocados por las aberraciones ópticas aumentan con el tamaño efectivo de la lente, puesto que los rayos incidentes en los extremos de la lente sufren mayores desviaciones de foco y distorsión que los rayos centrales. Como consecuencia de esta conclusión, cuanto más cerrado esté el diafragma y por tanto menor sea la abertura efectiva de la lente, tanto menor serán los defectos apreciables de la lente. El efecto adverso será una disminución de la intensidad luminosa que atraviesa la cámara, cuya consecuencia se traduce en una necesidad de

aumentar la potencia de la fuente luminosa para conservar el nivel de excitación radiométrica del píxel o bien aumentar el tiempo de exposición.

2.3.8 Función de punto extendido (PSF y OTF)

Por muy pequeño que sea el orificio de una cámara real, por él siempre pasarán varios rayos pertenecientes al haz de rayos divergentes que se originan de la reflexión de la luz sobre cualquier superficie. Este hecho implica que en vez de tener un punto luminoso en el plano sensor, se tendrá un círculo proyectivo. El efecto anterior provoca una degradación en la imagen. La luz incidente en cada punto del plano sensor proviene de un conjunto de puntos próximos del objeto y no de un solo punto.

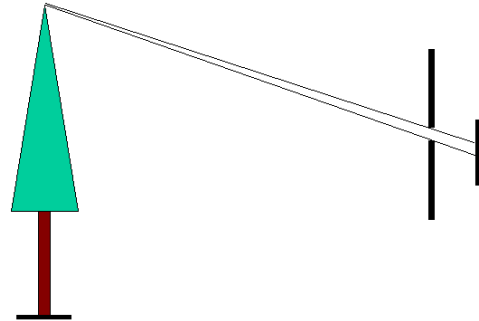


Figura 2. 38. Degradación del pulso de dirac

Cuando se emite un pulso de luz, tipo dirac, éste es procesado por el sistema óptico, formando sobre el plano sensor una mancha difusa. La distribución de radiación sobre el plano sensor del pulso de luz, cuando el nivel de radiación es la unidad, se define como la función de punto extendido (*Point Spread Function*, PSF). A la transformada de Fourier de la función de punto extendido se la denomina función de transferencia óptica (OTF).

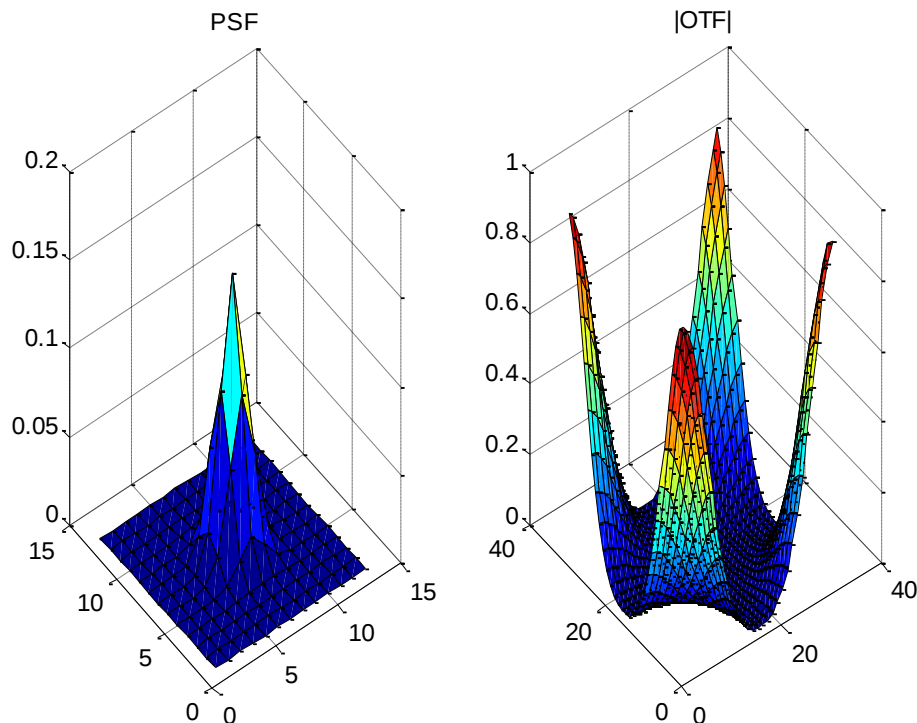


Figura 2. 39. Distribución gaussiana de la PSF y su función de transferencia óptica

2.3.9 Lentes comerciales.

Para evitar las degradaciones de la lente delgada, las ópticas o lentes comerciales se componen de varias lentes simples para intentar disminuir las aberraciones. La diferencia fundamental entre las ópticas utilizadas en fotografía y las usadas en las cámaras de vídeo es el tamaño del plano de la imagen. Mientras las fotográficas emplean una imagen de 35 mm, las de vídeo utilizan un plano sensor según el tamaño del sensor empleado; los más usuales son los de 1/3 y 1/2 de pulgada, aunque existen sensores desde 1/7 hasta 1 pulgada. El tamaño del sensor se refiere a su diagonal y la relación entre su anchura y altura suele ser de 3/4.

2.3.10 Características de las ópticas.

Enfoque. Como se ha visto anteriormente, actuando sobre el enfoque se puede conseguir una imagen más o menos nítida. Las ópticas comerciales se pueden clasificar por el enfoque en: ópticas fijas, ópticas variables y ópticas motorizadas. Estas últimas disponen de un motor de forma que se puede enfocar automáticamente o de manera teleoperada.

Iris o diafragma. El dispositivo de abertura está formado por láminas muy delgadas dispuestas de tal manera que en el centro se forma un círculo por el que puede entrar la luz. Abriendo o cerrando el iris entrará más o menos luz, haciendo que la imagen sea más clara o más oscura. Como se ha visto anteriormente cerrando el iris también se aumenta la profundidad de campo y se disminuyen las aberraciones de la lente, sin embargo se tendrá que iluminar más la escena. El iris puede ser manual, automático (autoiris) o motorizado. Las ópticas con autoiris tienen un motor, al igual que las motorizadas, pero ésta no se controla por un usuario a distancia, sino por la propia cámara. Las cámaras de vídeo tienen una salida de control para este propósito que se conecta con un cable a las ópticas con autoiris. La cámara regulará el autoiris abriéndolo o cerrándolo en función de la cantidad de luz que esté recibiendo.

Zoom. Las ópticas con zoom permiten, mediante dos conjuntos de lentes que se pueden acercar o separar, variar la distancia focal. De esta manera se puede fijar en detalles que a simple vista no se aprecian. El efecto de aumentar el zoom es la disminución del ángulo de vista. Las ópticas pueden ser con o sin zoom y éstas pueden ser manuales o motorizadas.

2.3.11 Parámetros de las ópticas.

A la hora de comprar una óptica en el mercado para un sistema de visión hay que tener claro las necesidades del proyecto. El precio de la óptica puede variar entre unos pocos euros hasta cientos de euros, en función de sus características y calidades. Los principales parámetros en la elección de una óptica son:

Distancia focal f . La óptica comercial tiene dos puntos principales, primario y secundario. La distancia en milímetros entre el punto principal-secundario y el sensor de imagen es la distancia focal. En ópticas de enfoque variable, se permitirá un desplazamiento del segundo punto principal para modificar el foco respecto al valor nominal. En el caso de ópticas motorizadas, se dará el rango de enfoque. Es el principal

parámetro a tener en cuenta, ya que éste va a delimitar a que distancia se encuentra la cámara de la escena.

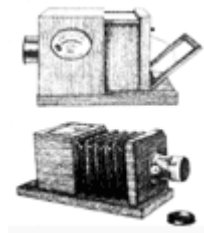
Número F. Es el índice de cantidad de luz que pasa a través de la lente. Cuanto más pequeño sea el número F, mayor será la cantidad de luz. Se puede calcular su valor como la relación entre la distancia focal y diámetro efectivo de abertura de la lente.

Angulo de vista. Es el ángulo que se forma con las dos líneas que parten desde el punto principal secundario a los extremos de la diagonal del sensor de imagen. A un mayor ángulo de vista se obtendrá un mayor campo de la escena en la imagen.

Rosca. La rosca es la forma con la que la óptica se sujeta a la cámara. Las más típicas en ópticas de vídeo son la rosca C y la rosca CS, habiendo incluso adaptadores de uno a otro. En cámaras fotográficas se suele emplear la rosca de bayoneta. Existen otros tipos de rosca como la rosca K y otras que emplean algunos fabricantes particulares como la rosca Wattec.

2.4 Sensores de vídeo

Básicamente, la cámara fotográfica del siglo XIX se constituía por una caja estanca en la atravesaba la luz a través de una lente. Un dispositivo mecánico dejaba pasar, durante un pequeño lapsus de tiempo, la radiación incidente sobre el interior de la caja. La escena se proyecta sobre la zona sensible, la cual estaba formada por cristales de halogenuros de plata, muy sensibles a la luz. La tecnología de la fotografía se ha modificado espectacularmente hasta hoy día, pero el principio sigue siendo el mismo, i.e. de excitación con fotones a un área sensible dentro de un habitáculo cerrado.



En 1897 se inventó el tubo de rayos catódicos, ofreciendo la oportunidad de imágenes no fotográficas, pero no fue hasta 1923 cuando se perfeccionó para poder adquirir imágenes. El sistema usaba un tubo de vacío, una pantalla sensible a la luz y un haz de electrones para el escaneado de la señal. La llegada en 1930 de la televisión comercial hizo que las cámaras de vídeo alcanzaran una importancia trascendental.

En 1970 se inventó el primer sensor de estado sólido. La imagen era formada por la conversión de los fotones a cargas eléctricas. Esta transformación se realiza por un elemento semiconductor llamado píxel que se disponía en forma matricial. Las cargas eléctricas eran leídas posteriormente por filas y por columnas, transformando la señal de vídeo en información eléctrica. Se inició entonces una batalla por la dominación del mercado entre las cámaras de estado sólido y las de tubo de vacío. Ésta fue ganada por las cámaras de estado sólido antes sus innumerables ventajas, como espacio, inmunidad a vibraciones, coste de fabricación, etc. Algo similar se libra hoy en día entre las clásicas cámaras fotográficas analógicas y las digitales.

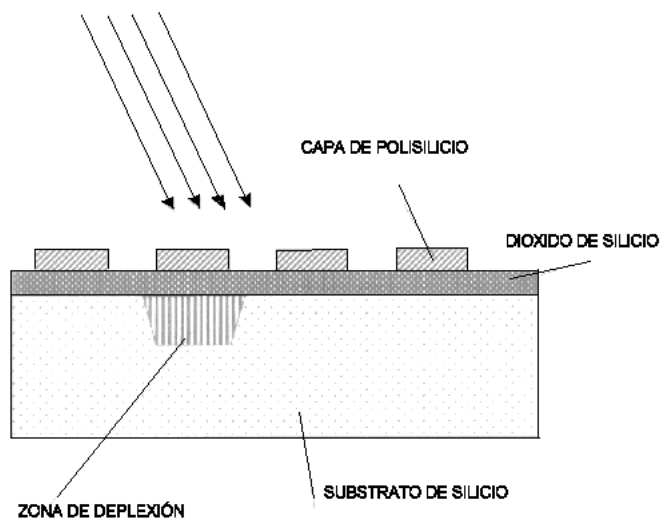


2.4.1 Cámaras de estado sólido

La mayoría de las cámaras de estado sólido están basadas en dispositivos de cargas acopladas o CCD (*Charge Coupled Device*), si bien existen otras tecnologías, como son la de inyección de carga o CID (*Charge Injection Device*) o las cámaras de integración o TDI (*Time Delay Integration*). Otra tecnología pujante es la CMOS. Aun así, el funcionamiento básico en las distintas tecnologías es similar y está fundado en el efecto fotoeléctrico. Los fotones de la luz pasan por una estructura cristalina de silicio y son absorbidos creando pares electrón-hueco. La cantidad de pares electrón-hueco será proporcional a la intensidad de luz que ha incidido sobre el elemento sensor.

2.4.2 Fundamentos del fotosensor.

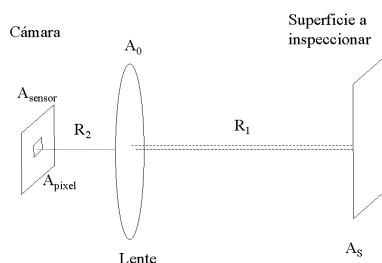
El elemento básico de las células fotosensibles son los píxeles. Éstos están constituidos por condensadores MOS (Metal Oxido Semiconductor). Sobre un sustrato de silicio de tipo P, se monta una zona de tipo N y se deposita una película de dióxido de silicio, SiO_2 , formando una capa de aislante o dieléctrico.



Encima de esta capa se sitúa un electrodo conductor de polisilicio transparente. La radiación incidente de la luz visible pasa a través del conductor transparente y crean pares electrón-hueco en el semiconductor. Los huecos libres se combinan con los electrones del sustrato, mientras que la acumulación de los electrones libres produce una carga eléctrica proporcional a la intensidad de la luz y al tiempo de integración.

Figura 2. 40 Estructura del píxel

2.4.3 Modelo radiométrico del píxel



Si el sistema de adquisición de imágenes, lente más cámara, se encuentra a una distancia de R_1 , desde la superficie de observación, el nivel energético de fotones incidentes dentro del sistema óptico, de área A_0 y durante el tiempo de integración, t_i , es de:

Figura 2. 41 Modelo radiométrico del píxel

$$d\phi \cong L_r \cdot S_{aparente} \cdot d\omega_r$$

$$n_{lente} = L_r A_s \frac{A_0}{R_1^2} \tau_A t_I \quad (2.7)$$

donde L_r es la radiación refleja en la dirección del sensor y τ_A es el coeficiente de transmisión atmosférica. El número de fotones que alcanza el sensor será:

$$n_{sensor} = L_r A_s \frac{A_0}{R_1^2} \tau_A \tau_O t_I \quad (2.8)$$

τ_O es la transmitancia del sistema óptico. El número de fotones incidentes en un píxel, es una relación de áreas que depende del factor de relleno:

$$n_{pixel} = n_{sensor} \frac{A_{pixel}}{A_{sensor}} \quad (2.9)$$

usando la aproximación de ángulo pequeño para rayos axiales:

$$\frac{A_s}{A_{sensor}} = \frac{R_1^2}{R_2^2} \quad (2.10)$$

e introduciendo los conceptos de magnificación óptica, M , y la relación del modelo de lente delgada:

$$\frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{f} \quad (2.11)$$

el número de fotones que recibe el píxel será:

$$n_{pixel} \cong L_r \frac{A_0 A_{pixel}}{f^2 (1+M)^2} \tau_A \tau_O t_I \quad (2.12)$$

asumiendo una abertura circular e introduciendo el concepto de F - número, quedará:

$$n_{pixel} = \frac{\pi}{4} \frac{L_r A_{pixel}}{F^2 (1+M)^2} \tau_A \tau_O t_I \quad (2.13)$$

Todas las variables son función de la longitud de onda y de la adquisición de la imagen. El número de foto-electrones generados en un detector de estado sólido es:

$$n_{pe} = \int_{\lambda_1}^{\lambda_2} R_r(\lambda) n_{pixel}(\lambda) d\lambda$$

$$n_{pe} = \int_{\lambda_1}^{\lambda_2} R_r(\lambda) \frac{\pi}{4} \frac{L_r(\lambda) A_{pixel}}{F^2 M^2} \tau_A(\lambda) \tau_O(\lambda) t_I d\lambda \quad (2.14)$$

siendo $R_r(\lambda)$ la sensibilidad del píxel ante esta longitud de onda. Como se observa, la conversión de la radiación a señal eléctrica es sólo función de la luminancia. No discrimina el color. Para adquirir el color se necesitará que la integración se extienda sólo en una franja determinada del espectro. Se suele emplear la teoría del triestímulo, separando la radiación en tres zonas espectrales, las correspondientes al modelo RGB. Por otro lado, la conversión de fotones a carga eléctrica no es un proceso determinista sino probabilístico, debido al carácter cuántico del fenómeno fotoeléctrico.

Una vez obtenida la cuantificación de la radiación incidente hay que transmitirla. Esta operación depende de la arquitectura de la cámara y de la tecnología a emplear. Básicamente se hace a través de los vecinos de forma secuencial, de modo similar a cómo se comporta un registro de desplazamiento. El mecanismo se basa en aplicar unas señales de reloj sobre los electrodos de cada píxel. Al aplicar un potencial positivo en el vecino de un píxel, cuyo pozo de electrones libres es proporcional a luz recibida, estos se desplazarán a la ubicación del vecino, transfiriéndose la carga y produciendo un registro de desplazamiento analógico. Este tipo de transferencia se conoce como CCD. En la figura se observa un CCD con dos señales de desfase, V_1 y V_2 , que son utilizadas para transferir cargas desde los fotosensores al registro de lectura

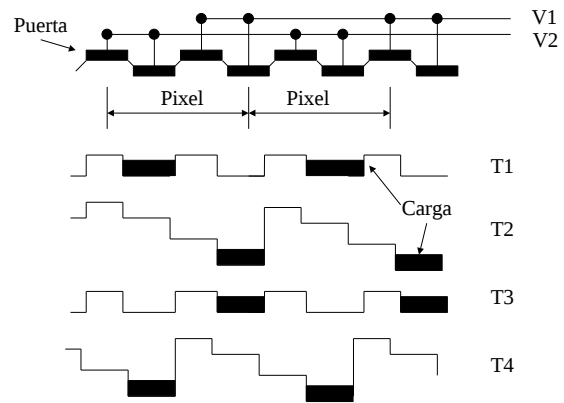


Figura 2. 42. Transferencia de cargas CCD

2.4.4 Arquitectura de las cámaras

Según la disposición de los píxeles, las imágenes pueden ser:

- Unidimensionales (cámaras lineales, asociadas a un vector de sensores)
- Bidimensionales (cámaras matriciales, asociadas a una matriz de sensores).

2.4.4.1 Cámaras lineales

Si bien es posible encontrar cámaras matriciales de hasta 5120 x 5120, éstas exigen anchos de bandas muy elevados. Por ejemplo, si se desea 25 cuadros por segundo, el volumen de

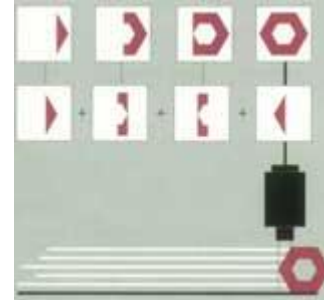


información es de 640 Mpixel/s, haciéndose prohibitivo para la mayoría de las aplicaciones industriales. Con el fin de evitar este exceso de información y si se desea alta resolución, la solución hay que encontrarla en las cámaras lineales.

La tecnología de cámaras lineales hace mucho tiempo que fue desarrollada para aplicaciones de inspección de materiales fabricados en continuo, como papel, tela, planchas metálicas, etc. Sin embargo, en la actualidad se está imponiendo en muchos otros procesos productivos que requieren de alta resolución y / o alta velocidad a un precio competitivo.

Las cámaras lineales utilizan sensores lineales que acostumbran a tener desde los 512 hasta 8192 elementos (píxeles), obteniendo grandes valores de SNR.

El hecho de construir una imagen de alta calidad a partir de líneas individuales requiere un alto grado de precisión. La alineación y los sincronismos del sistema son críticos si se quiere obtener una imagen correcta del objeto a analizar.



Constructivamente las cámaras lineales carecen de distancia interpíxel, i.e. no existen zonas no sensibles. Ello es debido a que el registro de desplazamiento se encuentra a uno de los lados de los fotosensores, y aunque éste también es sensible a la luz, se cubre con una capa protectora. Pudiéndose en este caso doblar la resolución de la cámara lineal.

2.4.4.1.1 Cámaras lineales TDI

En cualquier cámara convencional, el máximo tiempo de exposición está limitado a la velocidad de la escena. Como consecuencia de la adquisición línea a línea, el tiempo de integración es pequeño y se necesita mucha más iluminación que en una cámara matricial ya que la matricial funciona con tiempo de integración mayores. Por tanto, aplicaciones de alta velocidad en las escenas y baja iluminación no pueden ser adquiridas por las cámaras lineales.

La tecnología TDI (*Time Delay Integration*) utiliza el movimiento sincronizado para capturar múltiples tomas de la misma línea y sumarlas, obteniendo una línea con sensibilidad amplificada. TDI es el nombre dado al método de combinar muchas líneas individuales a la vez, con el fin de incrementar la sensibilidad de la cámara. La ventaja de este método reside en la mejora de la relación señal - ruido.

Las cámaras más modernas incorporan hasta 96 líneas de sensores, de forma que pueden alcanzar hasta 96 veces más sensibilidad que una cámara lineal convencional.

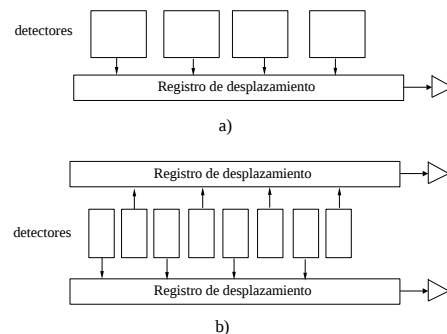


Figura 2. 43. a) un único registro de desplazamiento b) doble registro

El dispositivo se presenta físicamente como un array de sensores en disposición bidimensional, con un aspecto idéntico al de cualquier cámara matricial. El matiz que lo hace diferente es la forma de capturar la imagen y la electrónica que lleva integrada para que esto se pueda realizar. Mientras que en una cámara matricial se hace un muestreo de la información del array completo, en las cámaras TDI sólo se aprovecha una única línea.

Para entender el modo de operar de una cámara TDI, se sigue la figura 2.44. En el tiempo T1, la imagen se sitúa en el primer detector y se crea una carga. Llegado el tiempo T2, la imagen se ha desplazado al segundo detector. Simultáneamente, el reloj del píxel ha movido la carga del primero al segundo detector. En este punto, se crea una imagen formada por la carga anterior añadiéndose la nueva creada en el segundo detector y así sucesivamente. La carga (señal) se incrementa linealmente con el número de líneas, también lo hace el ruido pero no proporcionalmente sino con la raíz cuadrada del número de líneas TDI, N_{TDI} . El resultado es una mejora de la relación señal ruido en $\sqrt{N_{TDI}}$.

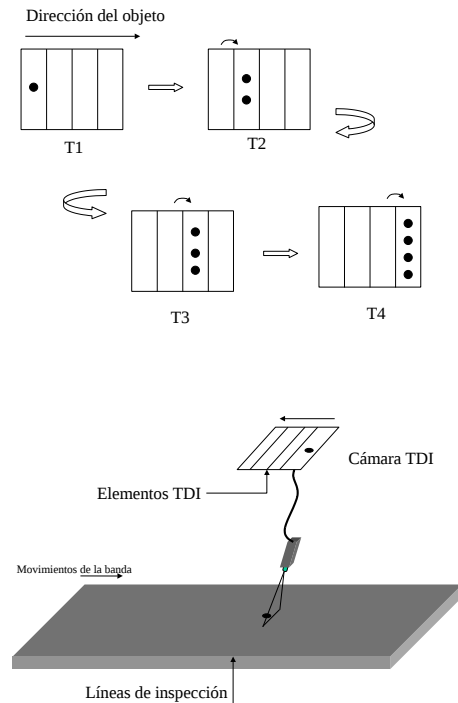


Figura 2. 44 Efecto TDI

Con las cámaras TDI, por tanto, se puede alcanzar más velocidad con baja iluminación. Sin embargo, hay que tener en cuenta que este tipo de cámaras deben estar perfectamente sincronizadas con la dirección del movimiento de la aplicación y que la velocidad de la aplicación debe ser bastante constante.

Entre sus aplicaciones, a parte de las industriales, se encuentran las capturas de imágenes procedentes de los satélites. Las imágenes de satélites se caracterizan por tener una gran variabilidad de iluminación, de ser adquiridas a velocidad constante y de requerir alta resolución. Por tanto, las cámaras TDI son ideales para estas tareas.

2.4.5 Cámaras matriciales

Una cámara matricial produce una imagen de un área, normalmente con una relación de aspecto de 4 a 3. Esta relación viene de los tiempos de las cámaras Vidicon y de los formatos de cine y televisión. Actualmente existen muchas cámaras que ya no mantienen esta relación y que no siguen los formatos de la televisión analógica.



Los sensores de cámaras modernos son mayoritariamente de tipo CCD o CMOS. Millones de píxeles se posicionan de forma muy precisa en una matriz y bien los registros de desplazamiento transfieren las cargas de los píxeles (CCD) para formar la señal de video o las cargas son convertidas en potenciales y transmitidas a la salida de vídeo (CMOS).

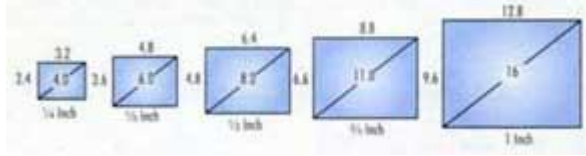


Figura 2. 45. Tamaños de los CCD

Los tamaños de las cámaras matriciales están definidos en pulgadas. Los formatos más comunes actualmente son de 1/3", 1/2", y 2/3".

2.4.5.1 Características de las cámaras matriciales

2.4.5.1.1 Factor de relleno

El factor de relleno es el porcentaje del área de píxel que es sensible a la luz. El caso ideal es 100%, i.e. cuando los píxeles activos ocupan el 100% del área del sensor. Sin embargo, circuitos como los registros de lectura y los circuitos anti-*blooming* reducen este factor, en algunas ocasiones hasta en un 20%. El efecto de esta reducción se traduce en una menor sensibilidad y en problemas de muestreo. Para evitar estas desventajas, muchos sensores con bajo factor de relleno (normalmente CCD con transferencia interlínea) utilizan microlentes que cubren cada uno de los píxeles incrementando la efectividad del factor de relleno.

2.4.5.1.2 Transferencia de las cámaras CCD

Uno de los mayores inconvenientes de una matriz de fotosensores es la transmisión de las señales recogidas. Dependiendo de la tecnología de la cámara, existen diferentes estructuras para la adquisición y transmisión de la información. Cualquier cámara, en principio, requeriría un área fotosensora que fuese capaz de captar toda una imagen enfocada sobre ella mediante un sistema óptico. Esta matriz de elementos fotosensores es la encargada de recoger la información que presenta la escena que se desee captar. Tras un cierto tiempo (denominado tiempo de integración), la información luminosa de la escena se encuentra recogida en esta matriz de elementos. Para extraer esta información de los elementos se utiliza la técnica de registro de desplazamiento analógico. De esta forma, la carga generada en cada elemento sensor se desplaza hasta la salida del dispositivo en su conjunto. Existen distintas configuraciones posibles que darán lugar a diferentes posibilidades tanto en la extracción como en la transmisión de la información adquirida. Las transferencias más empleadas son: transferencia de cuadro, transferencia interlinea y transferencia de cuadro interlínea.

Transferencia de cuadro (*Full Frame*).- Son los CCD que tienen una arquitectura más simple. Emplean un registro paralelo simple para exposición de los fotones, integración de la carga y transporte de la carga. Se utiliza un obturador mecánico para controlar la exposición. El área total del CCD está disponible para recibir los fotones durante el tiempo de exposición. El factor de relleno de estos tipos de CCD es del 100%.

Según este tipo de estructura el chip de silicio se divide en dos áreas. En la mitad superior del dispositivo aparece la sección donde va a incidir la luminosidad de la imagen que se desea captar en un momento determinado, mientras que en la mitad inferior, se encuentra el área de almacenamiento de cargas y el registro de salida. Esta mitad inferior, evidentemente estará recubierta por una placa opaca de forma que la luz no pueda alcanzarla y así no contamine la información recogida por el área de la zona superior.

Transferencia interlínea - Los CCD que incorporan esta tecnología son los más comunes y utilizan registros de desplazamiento que se encuentran entre las líneas de los píxeles y se encargan de almacenar y transferir los datos de la imagen. La principal ventaja de este tipo de CCD es la alta velocidad de obturación, pero no son tan sensibles y precisas como otros tipos de tecnologías. Tienen un bajo nivel de factor de relleno y un rango dinámico más bajo.

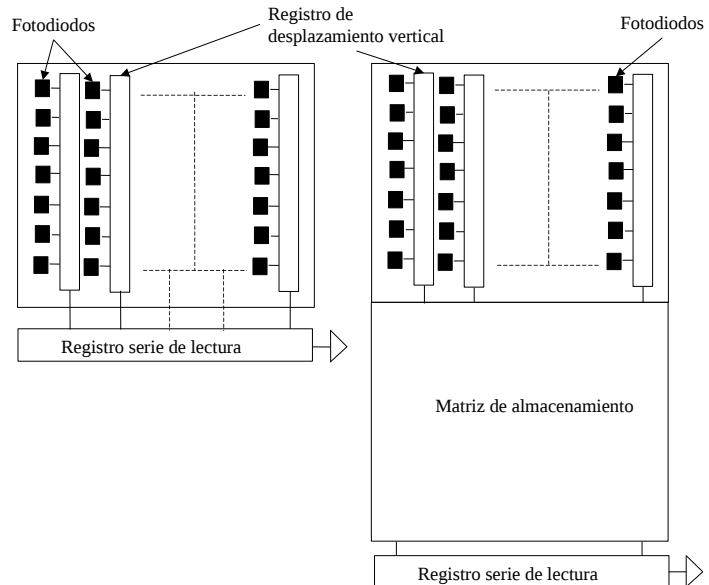


Figura 2. 46 a) Transferencia interlínea b) Transferencia interlínea cuadro

Este tipo de transferencia consiste en un array de fotodiodos separados por registros de transferencia vertical que son cubiertos por una capa opaca de metal. Después de la integración, las cargas generadas por los fotodiodos son transferidas al registro vertical CCD en un tiempo de $1\mu s$, minimizando los efectos de contaminación lumínica. La principal ventaja de esta transferencia es la rapidez, por lo que no hay necesidad de utilizar obturadores. La desventaja es que este tipo de salida implica una menor zona activa del sensor. El área efectiva puede estar por debajo del 20%. Aunque este factor puede ser ampliado mediante el uso de microlentes.

Transferencia de cuadro interlínea- Este tipo de estructuras incorpora características tanto de la transferencia interlínea como de cuadro completo. La parte superior de la estructura es exactamente igual a una interlineal. Las cargas se mueven horizontalmente desde los elementos fotosensores hacia los registros de desplazamientos verticales durante el periodo de borrado vertical. Una vez las cargas están en el registro de desplazamiento vertical, en lugar de ser leídas fila a fila a la frecuencia de línea, son enviadas a un array de almacenamiento. Este array de almacenamiento está, al igual que el registro de desplazamiento vertical, cubierto por una máscara de aluminio opaca a la luz. Es desde este array de almacenamiento donde las cargas pasan fila a fila al registro de salida. La diferencia es que en esta zona de almacenamiento los paquetes de información no corren ningún peligro de ser contaminados por excesos de carga que hayan penetrado en el registro de desplazamiento vertical, ya que en éste permanecen durante un período de tiempo muy

pequeño. Además, por tener el chip el registro de desplazamiento vertical separado de los elementos fotosensores y cubierto por una máscara opaca a la luz, no será necesario el empleo de obturadores mecánicos, como ocurría en los sensores de cuadro entero.

Esta estructura ofrece las mejores prestaciones de los CCDs actuales. Sin embargo, su estructura es compleja y requieren un área total mayor al tener la zona de almacenamiento separada. En concreto el área activa es alrededor del 22% del total de la zona sensora, lo cual ofrece una idea de la relación entre la zona activa y la zona sensora presentada en el proceso de captación de imágenes.

2.4.5.1.3 Integración de la señal de salida.

Las señales de vídeo fueron estandarizadas antes de que existiesen las cámaras de estado sólido. Los formatos de vídeo definen la frecuencia de la señal de video, el número de líneas por cuadro y el número de cuadros por segundo. Muchas de las cámaras de estado sólido siguen estos estándares, haciendo que la salida de la información sea de tipo entrelazado, como son las mayorías de las señales de televisión analógica. De esta forma, en los CCD con estructura de transferencia de campo, los mismos elementos fotosensores son utilizados en los campos pares e impares. En las estructuras de transferencia interlinea y de transferencia de cuadro, no es así, sino que se utilizan píxeles separados para cada campo, con el inherente

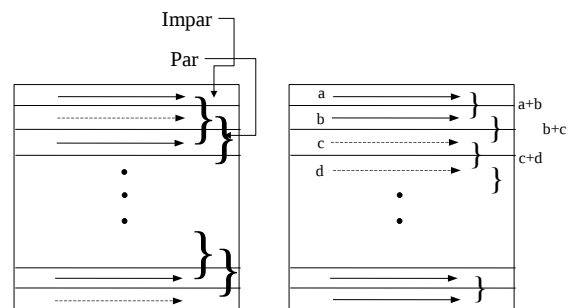


Figura 2. 47 Entrelazado a) Campo b) Cuadro

Existen dos posibles modos de lecturas de las cargas: Integración de cuadro e integración de campo.

En el modo de integración de cuadro cada elemento acumula cargas durante un cuadro (dos campos) antes de transferirlas al registro de desplazamiento vertical y desde él ser enviadas hacia la salida. Otra opción posible consiste en combinar las cargas de filas adyacentes del array de fotosensores. En este caso, los píxeles son leídos cada campo, denominándose este modo integración de campo. Como las señales adyacentes se promedian, la resolución vertical obtenida es menor que en el modo de Integración de Cuadro. Sin embargo, como la carga sólo se integra durante un campo, se obtiene mayor resolución dinámica que con el modo de Integración de Cuadro.

2.4.5.1.4 Cámaras progresivas

La mayor desventaja de las cámaras entrelazadas está en el desfase de los tiempos de integración, bien sea entre líneas o entre campos. Con el objeto de evitar parar el movimiento de los objetos al visualizarlas (acción requerida con las cámaras entrelazadas), aparecen las cámaras progresivas. Éstas “escanean” secuencialmente línea a línea las imágenes. La primera ventaja de las cámaras progresivas es la de tomar imágenes en un mismo instante, en contraste con los sistemas entrelazados que adquieren la imagen en el distintos tiempo de integración. Así cualquier movimiento vertical producido entre campos, provocará ‘manchas’ en la imagen, mientras movimientos horizontales originan cortes en las líneas verticales. Aún en el caso de utilizar luces estroboscópicas para

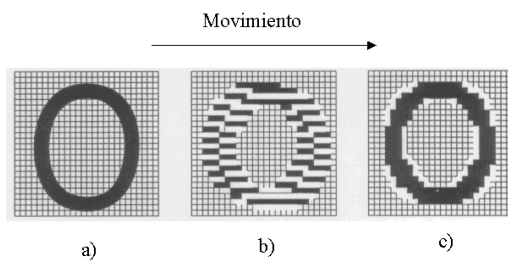


Figura 2. 48. Efecto del movimiento a) carácter en movimiento horizontal b) Discontinuidad de la imagen debido al sistema entrelazado c) salida de una cámara progresiva

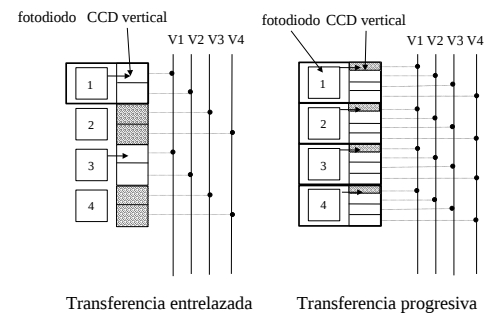


Figura 2. 49. Diferencias constructivas entre entrelazada y progresiva

congelar el movimiento, se tienen imágenes continuas sólo en el campo que fue activado el pulso de luz. Cuando hay movimiento, sólo un campo puede ser usado para el procesamiento de las imágenes, reduciendo en un 50% la resolución. En la figura se aprecia la discontinuidad de las cámaras entrelazadas y la mejora de las progresivas, aumentando la resolución vertical respecto a las entrelazadas.

La diferencia entre progresivas y entrelazadas está en su aspecto constructivo. Obsérvese en la figura como las cámaras progresivas, de cuatro fases, difieren en su diseño, permitiendo almacenar cargas en el mismo instante de tiempo.

2.4.5.1.5 CCD versus CMOS

Mientras que las cámaras CCD presentan una conversión de fotones a cargas eléctricas y éstas son desplazadas según los distintos tipos de transferencias analizados; en las cámaras CMOS, cada píxel se transforman directamente de fotones a señal eléctrica y estos pulsos eléctricos son los desplazados para ser transmitidos.

Las cámaras CCD tienen una mejor calidad de las imágenes y flexibilidad a expensas de un mayor tamaño. Sus aplicaciones son en fotografía digital, aplicaciones industriales de altas prestaciones, aplicaciones científicas y médicas. Mientras las cámaras CMOS tiene una integración mayor, menor consumo y tamaño más reducido, a expensas de una menor calidad de imagen y menor flexibilidad. Para grandes volúmenes con pocos requerimientos de calidad es la opción tecnológica. Sus

aplicaciones principales son las cámaras de seguridad, webcam, código de barras, scanners, juguetes, teléfonos móviles,....

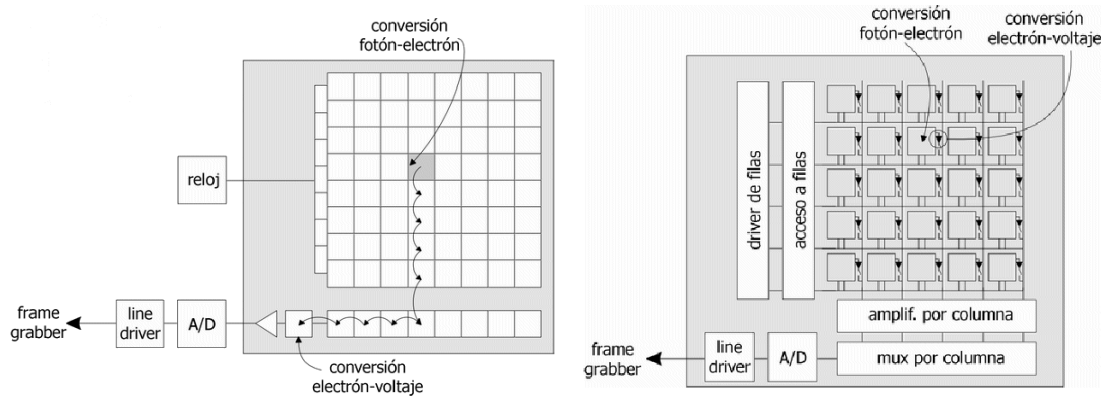


Figura 2. 50 a) Tecnología CCD b) Tecnología CMOS

2.4.5.1.6 Cámaras en color

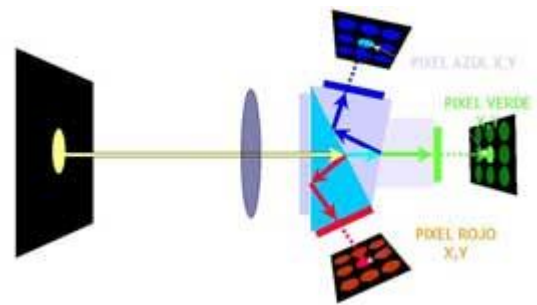
Aunque el proceso de adquisición de las imágenes color resulta ser más complejo, este tipo de cámaras pueden proporcionar más información que las cámaras de sólo luminancia. Hay dos tipos de arquitectura: 1CCD o 3CCD.

Las cámaras color de 1 CCD incorporan un sensor con un filtro en forma de mosaico que incorpora los colores primarios RGB. Esta máscara de color es conocido como filtro Bayer. De hecho, el sensor sólo integra la radiación al que se le ha superpuesto el filtro de color. La forma en que se disponen los colores R, G y B es como se muestra en la figura. Como se puede ver hay el doble de píxeles con filtro verde que con filtro azul o rojo. Efecto éste necesario para hacer más semejante la percepción del sensor a la visión humana que es más sensible al verde.



Figura 2. 51. 1CCD con filtro de Bayer

Debido al carácter del propio filtro es evidente que en los píxeles donde se sitúa el filtro rojo, no tienen señal ni de verde ni de azul. Para subsanar la falta de estos colores en estos píxeles, se construye una señal RGB a partir de los píxeles adyacentes de cada color. Este cálculo se realiza en el interior de la cámara mediante un DSP específico, que permite realizar la operación en tiempo real y dependiendo de la cámara permite obtener una señal analógica o digital en RGB en cada caso.



Las cámaras de color 3CCD incorporan un prisma y tres sensores. La luz procedente del objeto pasa a través de la **Figura 2. 52 3CCD**

óptica y se divide en tres direcciones al llegar al prisma. En cada una de los tres extremos del prisma se encuentra un filtro de color (rojo, verde y azul) y un sensor que captura la luz de cada color que viene del exterior. Internamente la cámara combina los colores y genera una señal RGB similar a la que ve el ojo humano. La fidelidad de las imágenes de las cámaras de 3CCD es muy superior a las de las cámaras de 1 CCD, pero hay un par de inconvenientes inherentes al sistema. Por una parte, este tipo de cámaras requieren más luz debido a que el prisma hace que sea menor la cantidad de iluminación que incide sobre los sensores y por otra, se genera un efecto de aberración cromática debida a la propia estructura del prisma. Este efecto puede ser subsanado colocando las ópticas diseñadas específicamente para este tipo de cámaras.

2.4.6 Elección de una cámara

Muchos son los aspectos que puedan influir en la elección de una cámara u otra. Sin embargo, algunas consideraciones deben de tenerse en cuenta para no arrepentirse de la compra o para no gastarse más dinero en el proyecto del que realmente hace falta. Normalmente, la elección de la cámara va íntimamente unida a la de la óptica.

Las características técnicas son las más importantes y las que siempre se tendrá presente a la hora de adquirir una cámara. A continuación se enumeran algunas de ellas.

2.4.6.1 Formato de Vídeo

Este tema será tratado con mayor profundidad en el siguiente apartado. El formato de la señal es importantísimo ya que tanto las cámaras como las digitalizadoras y, a veces, los monitores deben ser del mismo formato. El formato cambia de un país a otro. En Europa se tiene el sistema PAL o CCIR que será siempre al que se haga referencia. En los países francófonos se emplea el sistema SECAM y en el continente Americano y en Japón el sistema NTSC. Otro tipo de cámaras que no sigue ninguno de estos formatos son las llamadas cámaras digitales que envían la información al ordenador en señales digitales codificadas según el formato de cada fabricante, aunque se está extendiendo los protocolos IEEE 1394, USB 2.0 o GigaEthernet. Si se adquiere una cámara de este tipo hay que tener en cuenta las ventajas e inconvenientes de video digital respecto al analógico (anchos de banda, mezcla de video procedente de varias cámaras, ...).

2.4.6.2 Resolución

Una imagen en el sistema CCIR está formada por 576 filas y 768 columnas, sin embargo, el tamaño del área sensora puede ser mayor o menor. Como norma, cuantos más píxeles disponga el sensor mayor será la calidad de la imagen. Otro factor a tener en cuenta en la resolución, es el número de bits con que se digitaliza el píxel. Lo normal en las cámaras digitales es emplear 8 bits, aunque existen cámaras de hasta 16 bits por píxel.

2.4.6.3 Tipo de Rosca

La rosca es la forma con la que la óptica se sujeta a la cámara. Las más típicas en ópticas de vídeo son la rosca C y la rosca CS, habiendo incluso adaptadores de uno a otro. En cámaras fotográficas se suele emplear la rosca de bayoneta. Existen otros tipos de rosca como la rosca K, y otras que emplean algunos fabricantes particulares como la rosca Watec. Normalmente las cámaras tienen rosca C, añadiendo normalmente una arandela que convierte de C a CS. Se debe prestar especial atención a este hecho, en caso contrario no se podrá acoplar la óptica a la cámara.

2.4.6.4 Señales adicionales

Todas las cámaras poseen un conector, normalmente tipo BNC si es analógica, donde se consigue la señal de vídeo, además del de alimentación. Esto es lo mínimo necesario que se puede ofertar. Sin embargo, algunas cámaras ofrecen además las señales de sincronismo generadas en la propia cámara, que pueden ser empleadas en la sincronización de dispositivos de iluminación externos. También se suele incluir una señal de entrada para sincronizar la cámara otras cámaras si se está haciendo visión estéreo. Si se emplean varias cámaras en el sistema de visión, se puede emplear un generador de sincronismo externo; este generador hará que el tiempo de integración de todas las cámaras se inicie y termine exactamente en el mismo instante de tiempo. Todas estas señales se suelen agrupar en un conector denominado EXT. Otra señal que suelen tener algunas cámaras es la del control del autoiris que se conecta a la óptica. De esta manera, el iris se abrirá o cerrará automáticamente dependiendo de la cantidad de luz recibida en el CCD.

2.4.6.5 Controles adicionales

El fabricante normalmente ofrece varios ajustes que pueden ser modificados a voluntad. Algunos de ellos pueden ser imprescindibles para el máximo aprovechamiento del sistema de visión, por lo que se deberá prestar especial atención a sí la cámara que se desea adquirir ofrece uno u otro ajuste. A continuación se enumeran algunos de estos controles:

- *Control de ganancia automática (CAG):* Hace las veces de un autoiris electrónico. Si está activado, se aumentará o disminuirá la sensibilidad del CCD en función de la intensidad luminosa que se esté recibiendo. Muchas aplicaciones deberán tener desactivado este control si no se quiere tener efectos impredecibles.
- *Tiempo de integración:* Como ya se explicó anteriormente es el tiempo en que la luz incide sobre el CCD. Equivaldría al tiempo de obturación en términos fotográficos. Es una característica interesante en una cámara que se pueda modificar el tiempo de integración. Si se está iluminando la escena con luz fluorescente, como la red eléctrica funciona a 50 Hz, generará unas fluctuaciones en la luminosidad, de carácter senoidal y también de 50 Hz. Si la cámara tiene un tiempo de integración pequeño, por ejemplo de 2ms, la imagen se oscurezca y aclare lentamente, pudiendo afectar negativamente en el sistema de visión. Si se quiere eliminar este efecto se deberá cambiar el tiempo de integración lo más próximo a 20 ms.

2.4.6.6 Filtros de infrarrojos, IR

Los CCD tienen una buena respuesta por encima del espectro no visible, desde los 700nm hasta los 850nm aproximadamente. Con esta sensibilidad espectral haría que a luz ambiente la imagen apareciera muy saturada. Para evitarlo se suele intercalar antes del CCD un filtro de corte de infrarrojos. Este filtro debe tenerse muy en cuenta si se va a utilizar luz estructurada con longitudes de onda de infrarrojo (IR), en este caso se deberá abrir la cámara y eliminar este filtro. Algunas ópticas también incluyen este filtro y en la mayoría de los casos no es posible su eliminación.

2.5 Señal de vídeo

La información recogida en el sensor de imagen es convertida a una señal vídeo analógica o digital según un formato estándar o no. La transmisión de imágenes de vídeo ha estado impuesta por la aplicación principal para las que fueron concebidas las cámaras, la televisión. La señal analógica de vídeo es una señal eléctrica continua que contiene toda la información de la imagen, así como información adicional o sincronismos para que la imagen pueda volver a ser reconstruida mediante un monitor. Los tiempos y la estructura de la señal fueron determinadas en los primeros días en que la televisión vio la luz. Posteriormente, se han hecho modificaciones o añadidos a la señal de vídeo para que se pudiera transmitir color, sonido estéreo e información adicional (teletexto), siendo estas nuevas señales compatibles con el formato original.

El primer estándar para transmisión de televisión en niveles de grises fue conocido como RS-170 en Estados Unidos en 1941. Posteriormente, el formato NTSC, compatible con RS170, se definió para la transmisión de señal de vídeo en color. Se usa actualmente en 32 países (USA, Canadá, Japón, ...). En 1950, en Europa se estableció el estándar CCIR para la transmisión de señales de televisión en blanco y negro y PAL para el color. Las diferencias esenciales entre los sistemas monocromos RS-170 y CCIR son el número de líneas en la imagen y la transmisión de frecuencias, lo cual les hace incompatibles en un nivel básico.

Formato	RS-170	CCIR
Fre. Cuadro (Hz)	30	25
Entrelazado	2:1	2:1
Fre. Campo (Hz)	60	50
N. líneas horizontales	525	625
N. líneas horizont. efectivas	480	576
Fre. Línea (Hz)	15.750	15.625
Tiempo activo línea	52,5 μ s	52 μ s
Píxeles por línea	752	744
Relación de aspecto	4:3	4:3
Tamaño de imagen (píxeles)	360.960	428.544

En ambos modelos, el formato de vídeo usa el *principio de entrelazado* para reducir la frecuencia requerida en la transmisión y corregir el efecto de parpadeo al reconstruir la imagen. Este principio consiste en transmitir dos campos de líneas alternas y transmitir las secuencialmente. Un campo puede transmitirse en la mitad de

tiempo que cuadro completo, dando la impresión de continuidad, a pesar del hecho que sólo la mitad de la información está siendo actualizada en cada ciclo.

Aunque los sistemas no entrelazados son comunes para los monitores de ordenadores, el entrelazar campo par e impar es una característica común a todos los sistemas de televisión. El campo par contiene todas las líneas pares, comenzando con media línea y terminando con una completa. El campo impar contiene las líneas impares comenzando con una línea completa y terminando con la mitad de una línea. Los dos campos puestos juntos constituyen un cuadro completo.

La frecuencia de cuadro y de campo son las constantes fundamentales en los dos estándares de vídeo. Como se puede ver en la tabla, el formato RS-170 tiene una frecuencia de cuadro más alta, pero las frecuencias de líneas de los dos estándares son aproximadamente las mismas. En ambos casos, el número final de líneas efectivas se ve reducido por el tiempo de borrado de campo, el cual se usa en los monitores de TV para que el pincel electrónico vuelva al principio. Además, en el borrado de campo se envían señales para sincronizar los osciladores de barrido vertical y horizontal, así como información adicional (teletexto).

Un esquema completo de la señal de TV de un cuadro completo puede ser estudiado en la figura. En ella se incluyen las señales de sincronismos empleadas para especificar el comienzo de un campo (sincronismo vertical o *Vsync*) y el comienzo de una línea (sincronismo horizontal o *Hsync*).

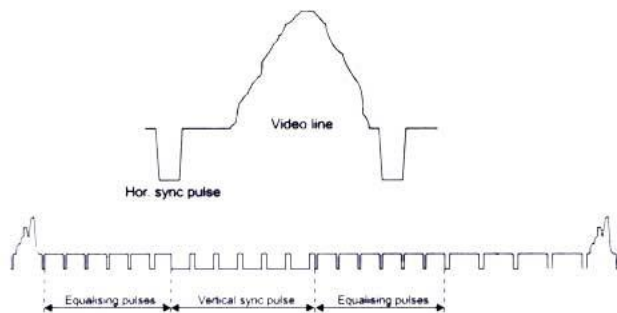


Figura 2. 53. Sincronismo vertical

Si las señales de sincronismo están combinadas con la señal de vídeo, la resultante es denominada señal de vídeo compuesto (*composite video*). Una señal que contenga *Vsync* y *Hsync* se denomina señal de sincronismo compuesto (*composite sync*). En CCIR habrá 625 pulsos de sincronismo horizontal por cuadro y 50 pulsos de sincronismo vertical por segundo.

En la figura se muestra un diagrama de la señal de una única línea con el sincronismo horizontal y los niveles de voltaje. El flanco de bajada del pulso de sincronismo es el comienzo de una nueva línea del campo y el nivel de borrado (*blanking level*) indica con cero voltios lo que sería negro en la imagen. El máximo nivel de blanco es equivalente a 0,7 voltios, mientras que el nivel de sincronismo está fijado a -0,3 V.

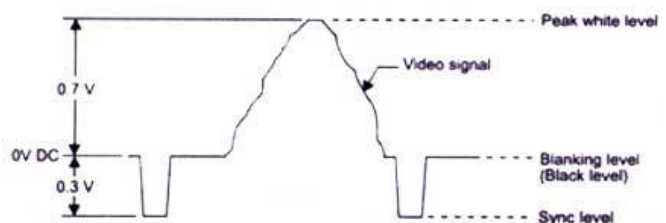


Figura 2. 54 Sincronismo de línea

Ningún estándar especifica la cantidad de píxeles en horizontal que debe tener un CCD, ya que los formatos nacieron para cámaras de tubo de vacío, en los que la señal analógica se formaba de manera continua. Sin embargo, los CCD hacen un muestreo de manera discreta. Cada fabricante de cámara ha elegido el número de píxeles que quiere incluir, aunque muchos de ellos han adoptado un estándar para RS170 de 752 píxeles. La resolución vertical del sensor suele ser de 484 líneas, aunque sólo 480 son activas.

2.5.1 Vídeo analógico no estándar

Las señales de video no estándares, pueden ser una combinación de: frecuencias no estándar, resolución no estándar o barrido progresivo. Las señales de video que no se ajustan a uno de los estándares descritos anteriormente no pueden ser interpretadas por un *frame grabber* convencional con circuitos de codificación para estándares. Las cámaras que tienen este tipo de señal son las cámaras de alta velocidad o alta resolución o las cámaras que se activan mediante trigger o captura asíncrona.

2.5.2 Señal de vídeo digital

Las cámaras digitales utilizan la misma tecnología que las analógicas, sin embargo, el convertidor analógico-digital está dentro de la cámara y digitaliza el video directamente y proporciona una señal de salida digital que puede ser de distintos tipos. El hecho de que la señal se digitalice en la propia cámara hace que la calidad de la señal sea mejor, pero tiene el inconveniente de que los cables entre la cámara y el *frame grabber* no pueden ser muy largos. La salida digital paralela permite conectar la cámara con un *frame grabber* mediante un cable multi hilos. Esta salida normalmente tiene el formato RS-422 o LVDS (EIA-644). No obstante, se está imponiendo la conexión de vídeo digital serie USB 2.0, FireWire o IEEE 1394 y GigaEthernet. A continuación se presentan los sistemas más habituales para la conexión directa entre las cámaras digitales y el ordenador.

2.5.2.1 USB (Universal Serial Bus)

USB es un sistema maestro (controladora USB) - esclavo (periférico) para las comunicaciones punto a punto, diseñado para reemplazar de manera universal los diversos protocolos de E/S tanto en serie como en paralelo. El dispositivo esclavo sólo puede comunicarse con el maestro (normalmente un ordenador) pero no con otros esclavos.

Con USB 2.0 (2000) hay un aumento de velocidad hasta 480 Mbps (casi 40 veces la velocidad de USB 1.1) con una diferencia de coste casi mínimo. Básicamente, USB 2.0 incluye todo lo que ofrece USB 1.1 y añade el modo de alta velocidad. USB 2.0 también usa el mismo tipo de cables y conectores para conectar los dispositivos de alta velocidad. La alta velocidad del USB le permite competir con la IEEE-1394a (Firewire) de 400 Megabits/ seg., si bien la IEEE-1394b sigue siendo más rápida, con 3,2 Gigabits/ seg.

Con respecto a la distancia, USB fue diseñada como una extensión del bus del ordenador personal, con la idea de que los periféricos estuviesen próximos, de forma que cada cable no midiese más de 5 metros. Otras interfaces, entre las que se incluyen la RS-232, la RS-485, la IEEE-1394b, y Ethernet permiten el uso de cables de longitud mayor. Si bien, la longitud de un vínculo USB se puede aumentar hasta los 30 metros usando cables que unan 5 hubs y un dispositivo.

2.5.2.2 IEEE-1394 (Firewire)

La tecnología FireWire (llamada así por su alta velocidad) la desarrolló Apple en 1986 y surgió de la necesidad de tener una conexión entre videocámaras digitales y los ordenadores personales, de bajo coste y orientada al consumidor. Es uno de los estándares de periféricos más rápidos desarrollados hasta el momento y resulta adecuado para el uso de periféricos multimedia.

En 1995, IEEE la adoptó como estándar bajo el nombre de IEEE 1394. Desde entonces se han desarrollado nuevas versiones que han añadido nuevas posibilidades y características. Sony es una de las compañías que ha adoptado este estándar desarrollando su propia versión conocida como i.Link que se puede encontrar en muchos de sus ordenadores y videocámaras digitales.

2.5.2.3 Ethernet

Ethernet fue desarrollado a principios de los 70 y estandarizándose en los 80 bajo la tutela del comité IEEE 802. Fue diseñado como un sistema de ‘*broadcast*’, de tal manera que todos los elementos conectados a la red pudieran enviar mensajes cuando quisieran y a quién quisieran, siendo recibidos los mensajes por todos los equipos de la red, pero atendiéndolos sólo el destinatario.

De la palabra Ethernet, “Ether” hace alusión al hecho de que el sistema no está restringido por el medio físico que utiliza para establecer la comunicación, ya sea cables de cobre, de fibra óptica o por ondas de radio. Existen varias versiones de Ethernet que se diferencian en el tipo de cable empleado y ha habido una evolución en las velocidades:

1. Fast Ethernet: también conocida como 100Base-T o IEEE 802.3u, emplea el mecanismo CSMA/CD como Ethernet pero tiene una tasa de transferencia mayor, de 100Mbps.

2. Ethernet Gigabit: es compatible tanto con Ethernet como Fast Ethernet, pero su tasa de transferencia es mayor, 1000Mbps.

2.5.2.4 Comparación de los distintos sistemas

En general, la IEEE-1394 puede ser más rápida y flexible que USB, pero su ejecución/aplicación es más cara. Con USB, una única tarjeta controla las comunicaciones con varios dispositivos. El ordenador se encarga de las tareas más complejas, de forma que los dispositivos pueden ser relativamente simples y baratos.

Por su parte, los dispositivos IEEE-1394 pueden comunicarse directamente entre ellos y una sola comunicación puede dirigirse a múltiples receptores. Esta característica tiene como resultado una interfaz más flexible, pero los dispositivos son más complejos y caros. La IEEE-1394 es más adecuada para aplicaciones que requieran comunicaciones extremadamente rápidas o para la emisión hacia múltiples receptores. Mientras que la USB es idónea para periféricos comunes como teclados, impresoras, escáneres, así como aplicaciones de velocidad baja a moderada y sensibles al coste.

Ethernet presenta como ventajas la posibilidad de usar cables muy largos, la capacidad de multidifusión, y la optimización de protocolos de Internet sobre Ethernet. Si bien, al igual que ocurría con la IEEE-1394, el hardware que se necesita para Ethernet es más complejo y caro que el hardware periférico habitual de la USB, por lo que le resta versatilidad.

Interfaz	Número de dispositivos (máximo)	Distancia (max. en metros)	Velocidad (bits/Segundo)	Aplicaciones típicas
USB	127	5 m y ampliable a 30m con 5 hubs.	1.5M, 12M, 480M	Ratón, teclado, audio, impresora y otros periféricos estándar
IEEE-1394b (FireWire 800)	64	91	3.2G	Vídeo y almacenamiento
Ethernet	-	488	10G	Comunicaciones de red

2.6 Tarjetas digitalizadoras

El nombre tarjeta digitalizadora no es correcto del todo, ya que ésta incluye operaciones de procesamiento e incluso permiten recibir señal de vídeo digital. Sin embargo, este término es ampliamente aceptado por razones históricas, pues al principio estas tarjetas no hacían más que digitalizar la imagen y era el computador el encargado de procesarlas. Estas tarjetas se podrían clasificar en tres tipos: a) Digitalizadoras de bajo coste, b) Digitalizadoras con características multicanal y c) Digitalizadoras con capacidad de procesamiento.

Una tarjeta digitalizadora o *frame grabber* tiene como propósito principal muestrear, en tiempo real, la salida estándar de una cámara de vídeo, pasando la información al computador. La información es almacenada temporalmente y una vez allí la información es procesada siguiendo las directrices de un programa. Dependiendo de la aplicación, la tarjeta puede ser programada para capturar y almacenar cuadros individuales, recoger una cierta secuencia de cuadros o quizás leer de manera continua la imagen de vídeo y realizar una tarea determinada en tiempo real. Algunas tarjetas tienen características avanzadas como puede ser captura síncrona de las imágenes.

Las digitalizadoras con capacidades para el procesamiento en tiempo real derivan sus prestaciones de los procesadores digitales de señal (DSP). Éstos pueden realizar procesamiento básicos sobre la señal de una manera encadenada y en tiempo real.

2.6.1 Funcionamiento de la adquisición de la señal de vídeo analógico

El funcionamiento en la adquisición de vídeo analógico se hace, primeramente, una adaptación de impedancias de la señal de vídeo entrante y es separada por un lado para obtener las señales de sincronismo y por otro el contenido de información de la imagen. Seguidamente, se acondiciona el nivel de continua a un nivel estándar (*DC restoration*), ya que la señal enviada por cada cámara puede variar ligeramente y este proceso ajusta el nivel de negro a una tensión constante. Posteriormente, se hace pasar la señal por un filtro paso bajo que elimina las altas frecuencias parásitas (*low pass filter*) y se corrigen las variaciones que se hayan producido en el nivel de continua y la ganancia (*Offset and Gain*). El fabricante normalmente ofrece la posibilidad de modificar los parámetros de estos circuitos mediante el programa que se ejecuta en el computador, de esta manera se puede conseguir una señal de mayor calidad. Debido a la alta frecuencia de la señal de vídeo, es muy importante que las impedancias de cable coaxial que se utilice y las de la cámara y entrada del digitalizador estén adaptadas. En caso contrario, se traducirá en una degradación de la señal de vídeo.

2.6.1.1 Sincronización.

Ya que las señales de sincronismo vienen junto a la de vídeo, tal y como se ha explicado, éstas deben separarse de la información de la imagen. La precisión de las señales de sincronismo es fundamental para la fidelidad de las imágenes capturadas por el digitalizador. El primer paso es separar la información de vídeo de los impulsos de sincronismo usando un circuito separador de sincronismo (*sync stripper*). La resultante se denomina sincronismo compuesto, ya que contiene tantos los impulsos verticales como los horizontales. Después es introducida como señal de mando a un bucle de control en cadena cerrada que en su salida tiene un oscilador generador de sincronismo (*Hsync Generator*), que es controlado mediante una señal de error producida por un circuito comparador de fase (*PLL Circuit*), que detecta las diferencias entre la señal de sincronismo que se está generando y la que debería ser. La señal obtenida es la que sincronizará en el convertidor analógico digital que es el encargado de pasar al formato digital la información de las imágenes.

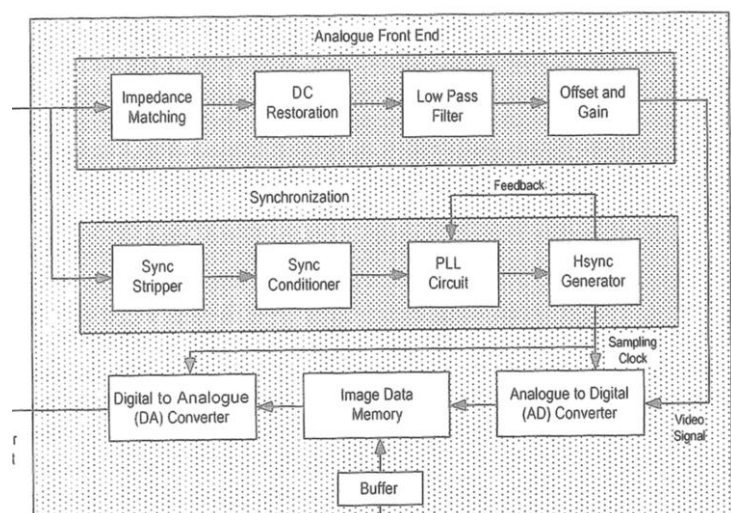


Figura 2. 55. Diagrama de bloques de una digitalizadora

El circuito generador de sincronismo horizontal fluctuará por algún tiempo hasta que el circuito PLL alcance un estado estable. Imprecisiones en la detección de la fase producirá el efecto denominado temblor de líneas o *line jitter*, que causa desplazamientos horizontales de hasta 3 píxeles. En una tarjeta digitalizadora de calidad, el *temblor de líneas* será normalmente menor de 0.1 píxeles.

Para solucionar este problema, algunas cámaras permiten ser sincronizadas por un oscilador exterior. En este caso el digitalizador será el que genere las señales de sincronismo vertical y horizontal o bien el sincronismo compuesto, haciendo que la digitalización sea en el momento preciso y corrigiendo totalmente este problema. Se tiene que tener especial cuidado cuando la longitud del cable sea elevada, ya que la señal de sincronismo puede ser retrasada por la inductancia y capacidad del cable, haciendo que la imagen aparezca desplazada.

También ha de tenerse especial cuidado con el efecto de calentamiento de los circuitos electrónicos de la cámara y la digitalizadora, ya que también pueden producir variaciones en el comportamiento de estos y afectar directamente en la calidad de la imagen. Es un buen hábito esperar una o dos hora con el equipo encendido hasta que la cámara se estabilice térmicamente.

La digitalización se produce en el rango de vídeo de 0 a 0.7 voltios. El periodo de muestreo para señales de televisión varía entre los 20 y los 30 MHz. Las tarjetas de propósito general emplean 8 bits que es equivale a un byte de información por píxel.

2.7 El modelo proyectivo de la formación de imágenes

El proceso de adquisición de una imagen por el sistema de Visión Artificial se modela en dos pasos: la formación de la imagen en el elemento sensor CCD a través de los elementos ópticos y la digitalización de la señal proveniente de la cámara. Cada uno de estos procesos se puede describir de distintas formas, más o menos complejas.

El modelo *pin-hole* es el más simple para representar la formación de la imagen. Está basado en la proyección de perspectivas y reduce el sistema óptico a un punto llamado centro óptico, C . Cada punto del espacio euclídeo se proyecta por un único rayo luminoso que pasa por el centro óptico, situado a una distancia f (distancia focal), hasta el plano de la imagen. El eje óptico es la línea perpendicular al plano de la imagen que pasa por el centro óptico. En este modelo se carece de proceso de enfoque.

Con este modelo y con una secuencia de imágenes adquiridas consecutivamente en el tiempo se trata de obtener la percepción visual de los objetos proyectados en el plano sensor. Para diferenciar las imágenes unas de otras se indexarán los parámetros. Así la proyección de un punto en el espacio $M_{wi}(x_{wi}, y_{wi}, z_{wi})$, teniendo el eje de coordenadas sobre el punto focal de la posición indexada i , se transforma en un punto sobre el plano sensor, estableciéndose la relación no lineal de:

$$\frac{x_i}{f_i} = -\frac{x_{wi}}{z_{wi}} \quad \frac{y_i}{f_i} = -\frac{y_{wi}}{z_{wi}} \quad (2.15)$$

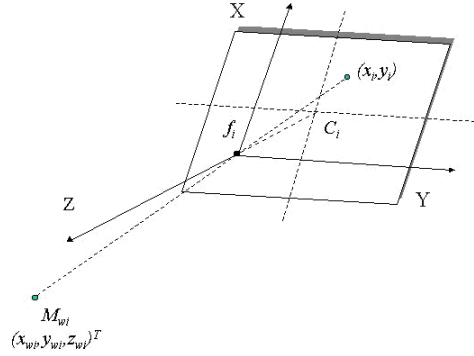


Figura 2. 56. Modelo *pin-hole* de la cámara

siendo x_i e y_i las coordenadas del punto proyectado. El signo menos refleja el carácter invertido del modelo. Obsérvese que estas proporciones están referenciadas en magnitudes de longitud, generalmente expresado en milímetros. Sin embargo, los algoritmos de extracción de las coordenadas del punto proyectado serán dados en píxeles. Habrá que proponer la conveniente conversión.

Para la transformación se requiere conocer el centro del plano sensor C_i (c_{xi} , c_{yi}); estos valores serán dados en píxeles desde el origen de coordenadas de la imagen digitalizada, es decir, la esquina superior izquierda. También es necesario conocer la distancia interpíxel de la cámara tanto en filas como en columnas (d_x , d_y). Hay que destacar que C_i es uno de los parámetros intrínsecos de la cámara, pues el desplazamiento del zoom o el enfoque de las ópticas hará variar este parámetro de una imagen a otra. En cambio, las distancias interpíxel se mantendrán al ser un parámetro de fabricación. No obstante, la diferencia entre el número de columnas sensoras de la cámara y la frecuencia de muestreo de la línea modificará el valor interpíxel de d_y . Además habrá de considerar el efecto de pérdida de sincronismo en la lectura de la línea, el llamado efecto temblor de línea que suele evaluarse por el factor de incertidumbre, k_y . Por tanto, se redefine d'_y como:

$$d'_y = d_y \frac{n_{sy}}{n_{py} k_x} \quad (2.16)$$

donde n_{sy} es el número de elementos sensores columna que hay en una cámara y n_{py} es el número de píxeles de la imagen digitalizada.

Del planteamiento presentado se deduce que la relación existente entre el punto proyectado sobre la cámara y su valor en píxeles sobre la imagen digitalizada resulta ser:

$$\begin{aligned} -x_i &= (p_{xi} - c_{xi})d_x \\ -y_i &= (p_{yi} - c_{yi})d'_y \end{aligned} \quad (2.17)$$

teniendo p_{xi} e p_{yi} que son las coordenadas en píxeles de la imagen sobre el punto proyectado en la cámara. Esta disposición de los signos hará eliminar el anteriormente mencionado efecto inversor del modelo *pin-hole*. Insertando las expresiones de la ec. 2.17 en la ec. 2.15 y presentándolo en coordenadas homogéneas y en notación matricial se tendrá:

$$\begin{pmatrix} p_{xi} \\ p_{yi} \\ 1 \end{pmatrix} = \begin{pmatrix} f_i/d_x & 0 & c_{xi} \\ 0 & f_i/d'_y & c_{yi} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{wi}/z_{wi} \\ y_{wi}/z_{wi} \\ 1 \end{pmatrix} \quad (2.18)$$

En 2.18 está implícito el carácter ortogonal de los ejes de proyección. En el caso de que no fuese así, el mapa de píxeles podría generalizarse a:

$$\begin{pmatrix} p_{xi} \\ p_{yi} \\ 1 \end{pmatrix} = \begin{pmatrix} f_i/d_x & -\rho f_i/d'_y & c_{xi} \\ 0 & f_i/d'_y & c_{yi} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{wi}/z_{wi} \\ y_{wi}/z_{wi} \\ 1 \end{pmatrix} \quad (2.19)$$

donde ρ es un escalar que mide el desplazamiento rotacional entre los dos ejes principales; si ρ es cero es que éstos son perpendiculares.

2.8 Calibración geométrica de las cámaras

La calibración de las cámaras se puede dar en un doble sentido: a) geométrica y b) radiométrica. La primera relaciona las transformaciones proyectivas desde la escena exterior 3D de la cámara con el plano sensor y la segunda con los aspectos de radiación en cada uno de los píxeles. El segundo aspecto fue tratado, someramente, en el apartado 2.4.3. Aquí se trata de determinar cual es el conjunto de transformaciones proyectivas que relaciona la escena con el plano sensor. Su utilidad resulta fundamental en tareas de percepción y de navegación visual. En todas estas aplicaciones, se trata de analizar las imágenes para obtener medidas geométricas del mundo exterior. Para

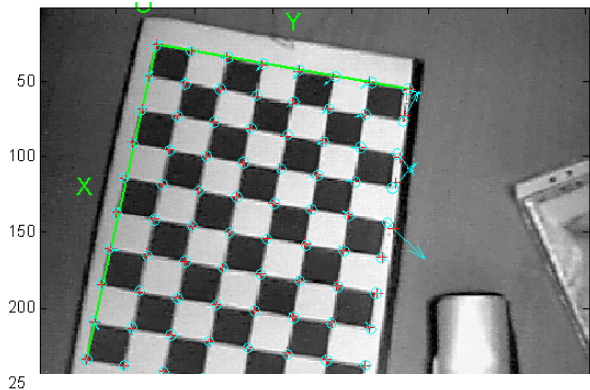


Figura 2. 57. Correspondencia entre los puntos 3D del espacio con los píxeles proyectivos

realizar estos algoritmos se exige de la calibración geométrica de las cámaras.

En función de lo comentado en el apartado anterior, existen dos tipos de parámetros en las transformaciones proyectivas: a) parámetros intrínsecos y b) parámetros extrínsecos. Los primeros son aquellos inherentes a la cámara y son: la distancia focal, el centro óptico y el modelo de deformación de la lente; mientras que los parámetros extrínsecos son los que definen la posición y orientación de la cámara respecto a un sistema de referencias 3D: vector de traslación y matriz de rotación.

2.8.1 Procedimiento general de calibración

Aunque existen distintos métodos para la calibración de una cámara, el procedimiento básico es el mismo en todos ellos:

1. Determinar con precisión un conjunto de puntos 3D del mundo exterior.
2. Fijar sus correspondencias con las proyecciones de estos puntos 3D sobre la imagen proyectada 2D.
3. Obtener mediante técnicas de optimización la mejor solución de la determinación de los parámetros intrínsecos y extrínsecos.

Los dos primeros ítems requieren conocer una serie de puntos 3D y sus correspondientes proyecciones en la imagen; son los que se denominan puntos de calibración. Dependiendo del algoritmo de calibración, los puntos pueden ser coplanares o no. Pero sea como fuere el tipo de plantilla de calibración, se sugiere un patrón de cuadros blancos y negros, a modo de tablero de ajedrez, tal que facilite el proceso de extracción de las esquinas con precisión subpixel y sea fácil de corresponder los puntos exteriores 3D con sus proyecciones 2D. Realizado los pasos de correspondencia se pasará a determinar los parámetros intrínsecos y extrínsecos mediante técnicas de optimización.

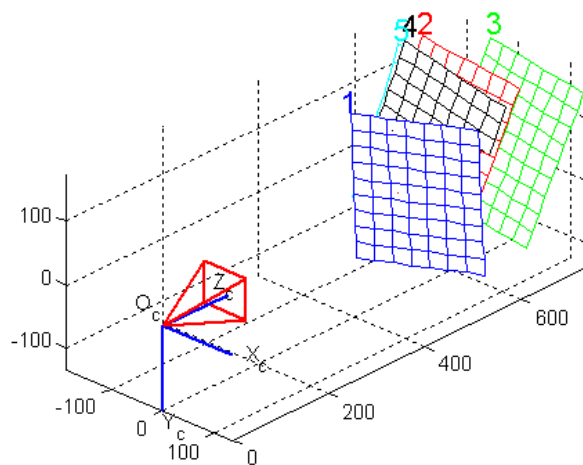


Figura 2. 58. Parámetros extrínsecos

2.9 Formatos gráficos.

Las imágenes digitales se pueden almacenar en distintos soportes físicos tales como: discos duros, CDs, DVDs o memoria flash. Para este objetivo se han inventado cientos de formatos con la función principal de comprimir la enorme cantidad de información que contiene una imagen digital. Una primera clasificación podría ser: formatos gráficos vectoriales y formatos gráficos en mapa de bits. Este último puede dividirse además en formatos con pérdida o no de la información.

2.9.1 Formatos gráficos vectoriales

Se almacenan las coordenadas de las figuras geométricas simples como segmentos, triángulos, rectángulos, circunferencias, etc., así como las órdenes necesarias para rellenar con los colores adecuados. Para dibujos simples el fichero resultante ocupará poco tamaño comparando con lo que ocuparía si estuviera en formato de mapa de bits. Tal es el caso de ficheros generados por programas de generación de imágenes 3D como AutoCAD, librerías de computación gráfica como VTK o programas de dibujos en general.

2.9.2 Formatos gráficos en mapa de bits

Este formato es el más intuitivo, ya que lo único que hace es almacenar secuencialmente los píxeles que componen la imagen en un fichero. El orden de almacenamiento se puede entender considerando la imagen como una gran matriz de píxeles. Así partiendo de la esquina superior izquierda, avanzando por las columnas de izquierda a derecha y por filas de arriba abajo.

Si el fichero se almacena así tal cual, se dice que está en formato crudo (*.raw), será necesario para su lectura saber cuantas filas y columnas tiene la imagen. El inconveniente del formato en crudo es la gran cantidad de espacio que se necesita. La mayoría de los formatos actuales llevan consigo una reducción sustancial del volumen informativo. La compresión de las imágenes se hace siguiendo algunas de estas tres estrategias:

1. Compresión en la codificación: en vez de emplear un número fijo de bits para cada nivel de gris o color, se emplea un código más corto a los niveles más probables y más largo a los menos probables.
2. Compresión en la representación: La similitud de intensidad en los vecinos hace emplear las rachas, las cuales indican el número de vecinos que se repiten. Se almacenan en vez de los valores de intensidad, el número de vecinos con la misma intensidad.
3. Compresión por redundancia visual: Las imágenes vistas por el ojo humano no es capaz de percibir el contraste de tonos muy próximos entre sí, por dicha razón en vez de almacenar los mapas de bits se aplican transformaciones matemáticas sobre la imagen, con el objeto de guardar

los principales coeficientes de la conversión. Evidentemente, la imagen resultante es de peor calidad que la entrante. Hay pérdidas de detalle. El ejemplo más característico es el algoritmo de compresión JPEG basadas en las transformadas discretas del coseno. También son ejemplos de compresión basada en redundancia visual los estándares de vídeo MPEG y MPEG2.

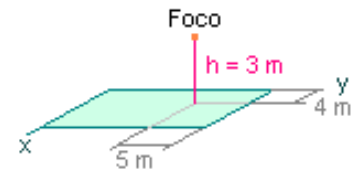
2.10 Cuestiones

1. Diferencias entre la Visión Artificial y la Computación Gráfica sobre la iluminación.
2. Factores que influyen en la formación de una imagen digital.
3. Calidad de una imagen.
4. Técnicas de iluminación.
5. Discrepancias entre el modelo de lente delgada y el modelo *pin-hole*.
6. Consecuencias de variar el diafragma.
7. Tipos de aberraciones y distorsiones ópticas.
8. Diferencias entre cámaras lineales y cámaras TDI
9. Modelado geométrico y radiométrico de una cámara.
10. Cámaras CCD versus cámaras CMOS.
11. Ventajas e inconvenientes de las distintas arquitecturas de transferencia: cuadro, interlínea, cuadro-interlínea.
12. Cámaras entrelazadas y cámaras progresivas.
13. Clasificación de las señales de vídeo.
14. Partes de una digitalizadora de vídeo de bajo costo.
15. Tipos de archivos de imágenes.

2.11 Problemas

Ejercicio 1

Una fuente luminosa puntual de 100 mW/sr se sitúa en un escenario de 10m x 20m, en X e Y respectivamente, tal cual se observa en la figura. Sabiendo que irradia por igual en todas las direcciones. Calcular la iluminación máxima y mínima sobre el suelo.



La iluminación máxima sobre el suelo se encontrará debajo del foco:

$$E_{\max} = \frac{I}{d_{\min}^2} = \frac{100}{9} \text{ mW/m}^2$$

La mínima se encontrará a la distancia más alejada entre el escenario y el foco:

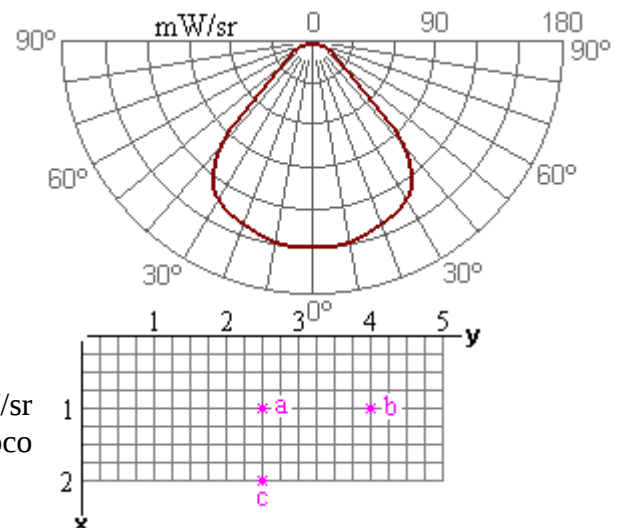
$$d_{\max}^2 = \sqrt{15^2 + 6^2} = 16.43 \text{ m}$$

El ángulo entre el foco y el extremo del escenario es de 79.4°. La iluminación total será:

$$E_{H,\min} = 0.068 \text{ mW/m}^2 \quad E_{V,\min} = 0.364 \text{ mW/m}^2 \quad |E_{\min}| = 0.37 \text{ mW/m}^2$$

Ejercicio 2

Una luminaria se sitúa en el centro de un escenario de 5m x 2m y a una altura de 3m. La intensidad luminosa es entregada por el fabricante según el diagrama polar adjuntado. Calcular la iluminación en los puntos a), b) y c).



En el punto A, la intensidad es de 150 mW/sr y el ángulo entre la normal de la superficie y el foco es nula:

$$E_{Ha} = \frac{I \cos^3 \alpha}{h^2} = 16.6 \text{ mW/m}^2$$

$$E_{Va} = E_H \tan \alpha = 0 \text{ mW/m}^2$$

$$|E_a| = 16.6 \text{ mW/m}^2$$

En el punto B, la intensidad es de 130 mW/sr y el ángulo entre la normal de la superficie y el foco es de 26.57°:

$$E_{Hb} = \frac{I \cos^3 \alpha}{h^2} = 10.3 \text{ mW} / \text{m}^2$$

$$E_{vb} = E_H \tan \alpha = 5.2 \text{ mW} / \text{m}^2$$

$$|E_b| = 11.5 \text{ mW} / \text{m}^2$$

En el punto C, la intensidad es de 140 mW/sr y el ángulo entre la normal de la superficie y el foco es de 18.43°:

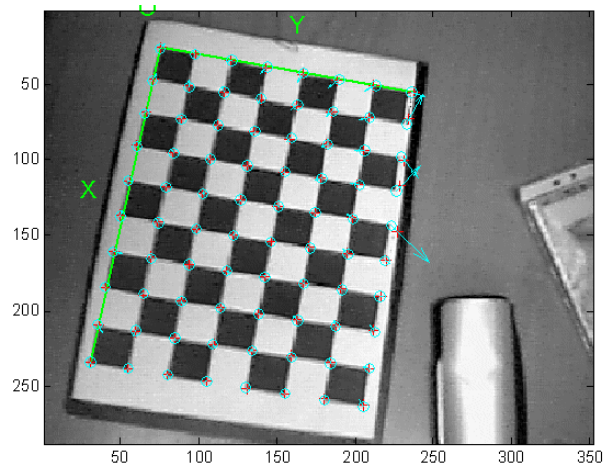
$$E_{Hc} = \frac{I \cos^3 \alpha}{h^2} = 13.3 \text{ mW} / \text{m}^2$$

$$E_{vc} = E_H \tan \alpha = 4.4 \text{ mW} / \text{m}^2$$

$$|E_c| = 14 \text{ mW} / \text{m}^2$$

Ejercicio 3

Se tiene una cámara de vídeo a calibrar. La óptica tiene una distancia focal de 3mm y el tamaño del píxel es de 5.6 μm x 5.6 μm. El número efectivo de píxeles son 357(H) x 293(V). ¿Cual es la distancia mínima que podrá ponerse una rejilla de calibración respecto de la cámara, si ésta se constituye por cuadros blancos y negros de 27mm de lado y según se observa en la figura, hay 7x9 cuadrados?



En la mínima distancia deberá de entrar la rejilla completa. Se parte de la hipótesis que los nueve cuadros deben de entrar en las 357 columnas, por tanto:

$$Z = f \frac{Y}{y} = 3 \cdot 10^{-3} \frac{9 \cdot 27 \cdot 10^{-3}}{357 \cdot 5.6 \cdot 10^{-6}} = 0.364 \text{ m}$$

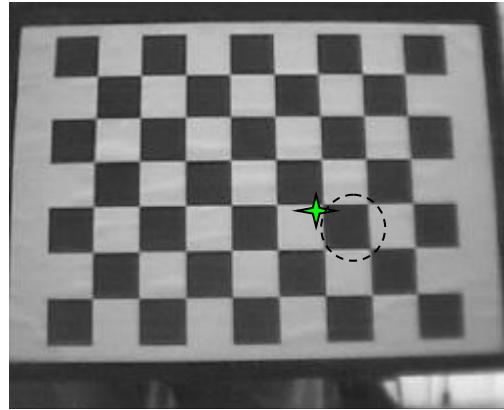
Habría que observar que en las filas entrar los siete cuadrados:

$$Z = f \frac{X}{x} = 3 \cdot 10^{-3} \frac{7 \cdot 27 \cdot 10^{-3}}{293 \cdot 5.6 \cdot 10^{-6}} = 0.345 \text{ m}$$

Por tanto, se verifica que cuando la distancia es de 0.364 m entra por completo la rejilla.

Ejercicio 4

Se emplea una cámara con una lente de 3mm y un sensor CMOS que tiene un pixel cuadrado de $5.6\mu\text{m} \times 5.6\mu\text{m}$. El número efectivo de píxeles es 357 (H) x 293 (V). Se ha colocado una plantilla de calibración a 500 mm respecto al eje de referencia en la cámara. Sabiendo que los lados del cuadrado son de 27mm y que las coordenadas XYZ del punto indicado en la rejilla son (50,50,500). Determinar en qué píxeles de la cámara se proyectan las cuatro esquinas del cuadrado seleccionado.



Considérese que no hay distorsión en la lente, el eje axial de la lente pasa por el centro del elemento sensor y que los ejes de proyección son ortonormales.

El modelo pin-hole de la cámara quedará definido por:

$$\begin{pmatrix} p_{xi} \\ p_{yi} \\ 1 \end{pmatrix} = \begin{pmatrix} f_i/d_x & 0 & c_{xi} \\ 0 & f_i/d'_y & c_{yi} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{wi}/z_{wi} \\ y_{wi}/z_{wi} \\ 1 \end{pmatrix} = \begin{pmatrix} 535.71 & 0 & 146.5 \\ 0 & 535.71 & 178.5 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{wi}/z_{wi} \\ y_{wi}/z_{wi} \\ 1 \end{pmatrix}$$

El primer punto quedará proyectado en el píxel (200,232) y los otros tres en (229,232), (200,261) y (229,261).

Ejercicio 5

La imagen aérea fue adquirida a una altura de 2500 m con una cámara con distancia focal de 125 mm. Sabiendo que la distancia entre las dos marcas es de 30m, determinar las coordenadas en píxeles de la segunda marca (la de la derecha), si la marca de la izquierda está en el píxel (250,500). Datos: Píxel cuadrado de $10\mu\text{m}$.



El factor de magnificación es: $M = \frac{f}{Z} = \frac{0.125}{2500}$, luego la proyección de las dos marcas sobre el elemento sensor es: $y = M \cdot Y = 0.0015\text{m}$, lo que da una separación en columnas de 150 píxeles. La coordenada del pixel de la derecha es (250,650).

Derecho de Autor © 2016 Carlos Platero Dueñas.

Permiso para copiar, distribuir y/o modificar este documento bajo los términos de la Licencia de Documentación Libre GNU, Versión 1.1 o cualquier otra versión posterior publicada por la Free Software Foundation; sin secciones invariantes, sin texto de la Cubierta Frontal, así como el texto de la Cubierta Posterior. Una copia de la licencia es incluida en la sección titulada "Licencia de Documentación Libre GNU".

La Licencia de documentación libre GNU (GNU Free Documentation License) es una licencia con [copyleft](http://www.gnu.org/copyleft/) para [contenidos abiertos](http://www.gnu.org/licenses/). Todos los contenidos de estos apuntes están cubiertos por esta licencia. La version 1.1 se encuentra en <http://www.gnu.org/copyleft/fdl.html>. La traducción (no oficial) al castellano de la versión 1.1 se encuentra en <http://www.es.gnu.org/Licencias/fdles.html>