

## IMAGES SEQUENCE MOVEMENT ESTIMATION. AN ARTIFICIAL VISION APPLICATION

### ESTIMACION DE MOVIMIENTO EN SECUENCIAS DE IMÁGENES. UNA APLICACIÓN DE LA VISION ARTIFICIAL

Oscar Ivan Higuera Martínez

Universidad Pedagógica y Tecnológica de Colombia

Calle 4S # 15-134

osiva18@yahoo.es

**Abstract:** This paper contains the project description that is carrying out as the part of the Signal Processing Group (DSP) in the UPTC, as the images processing research line, this project seeks to estimate and extract the movement of a body in an images sequence to its identification. An exploration is given to the algorithms until the moment developed to estimate the movement and the expectations, as well as the results reached at this moment.

**Resumen:** El presente articulo contiene la descripción del proyecto que se esta llevando a cabo por parte del grupo de procesamiento de señales (DSP) de la UPTC, en su línea de investigación de procesamiento de imágenes, este proyecto pretende estimar y extraer un cuerpo en movimiento en una secuencia de imágenes para posteriormente realizar su identificación. Se da una exploración a los algoritmos hasta el momento desarrollado para estimar el movimiento y las expectativas que se tienen, así como los resultados hasta el momento alcanzados.

**Keywords:** Images sequence, Images processing, Optical flux, Estimation algorithm.

## 1. INTRODUCCIÓN

El desarrollo de algoritmos eficientes para la estimación de movimiento en secuencias de imágenes es un área de la visión artificial que a tenido avances significativos en las últimas décadas y sus diversas aplicaciones aumentan día a día, entre ellas se encuentran la detección de movimiento aplicado a seguridad (Reconocimiento de huellas o rostros, vigilancia), Detección y localización de objetos en movimiento donde se puede incluir también la detección de trayectorias y predicción de futuras trayectorias, también esta la obtención de propiedades 3-D de objetos a partir

del movimiento, y la estimación de movimiento en problemas de compresión de vídeo, control de calidad, biomedicina (Diagnostico medico). Todas estas aplicaciones están fundamentadas básicamente en la determinación del flujo óptico.

Hace algunos años solo se operaban sobre imágenes estáticas, en la actualidad los algoritmos se pueden emplear en el procesamiento de secuencias de imágenes, la extracción de información en secuencia de imágenes es una tarea muy importante y clave para muchos sistemas de visión, el conocimiento en esta área es de vital importancia en proyectos que se desarrollan en el

grupo de procesamiento de señales (DSP) de la UPTC en su línea de investigación en procesamiento de imágenes, donde se pretende en un futuro realizar el reconocimiento de objetos en movimiento.

## 2. ESPECIFICACIONES DEL PROBLEMA.

El proyecto que se está desarrollando consiste en la detección y extracción de cuerpos en movimiento en una secuencia de imágenes para realizar su posterior identificación, este tipo de imágenes serán tratadas en formato escala de grises y su escala óptica es variable, la escala de la imagen incide en la velocidad percibida, de igual manera la resolución y el algoritmo empleado afectan la eficiencia computacional de la aplicación

## 3. ETAPAS DEL PROCESAMIENTO.

Dentro del procesamiento digital de imágenes a utilizarse se tienen varias etapas. El objetivo general es producir un resultado a partir de solucionar un problema, como en este caso es detectar y extraer cuerpos en movimiento, la primera etapa es realizar la adquisición de la secuencia de imágenes, la siguiente etapa es la adecuación de imágenes en un pre-procesamiento de la imagen (Filtrado, Umbralización, etc.) cuya misión es mejorar las características de la señal de forma que aumenten las posibilidades de cumplir los objetivos, posteriormente se realiza la detección del objeto en movimiento con esta etapa se pretende determinar el movimiento y así extraer el objeto en movimiento y finalmente realizar la caracterización.

A continuación se presenta un diagrama de bloques de dichas etapas del proceso.

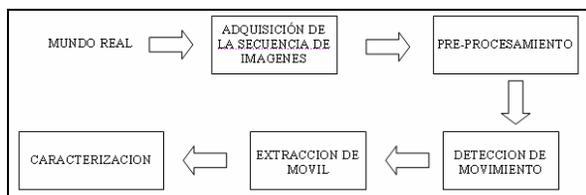


Fig. 1. Diagrama de bloques, etapas a seguir en el sistema a implementar.

En este momento se está estudiando la detección y estimación de movimiento, y los diferentes modelos para determinar el movimiento, en base a esto se han trabajado los algoritmos mencionados

más adelante, a continuación se da un breve resumen de los fundamentos que se han trabajado hasta el momento en estimación de movimiento.

## 4. VISION Y MOVIMIENTO

Para entender la visión por computador se ha dividido en dos niveles principalmente, estos son visión de bajo nivel y visión de alto nivel, la figura 2 nos ilustra esta división ampliamente.

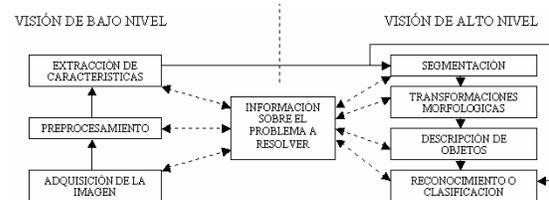


Fig. 2. Aplicación basada en visión por computador

La estimación del movimiento se encuentra estrechamente relacionada con los cambios temporales y espaciales de los valores de gris, el movimiento de una imagen a otra puede ser descrito por un vector desplazamiento, el cual puede ir desde un punto del contorno de la primera imagen a algún punto del contorno de la segunda, existen múltiples modelos de movimiento.

El análisis del movimiento a partir de secuencias de imágenes, busca extraer parámetros que caractericen el movimiento de los objetos. Por lo general, los parámetros son coeficientes de ecuaciones que rigen el comportamiento dinámico de los objetos. Según esto, en el procesamiento de imágenes en secuencia se han utilizado ecuaciones de diverso tipo como lineales, no lineales y polinomiales. Sin embargo, las ecuaciones más utilizadas son las lineales, como la detallada a continuación. Suponiendo que  $R1$  es la región proyectada por cierto objeto en el plano de la imagen en un tiempo  $t1$ , y  $R2$  es la región proyectada por el mismo objeto en el tiempo  $t2$  después de una traslación. Se puede considerar que la transformación que relaciona ambas regiones es:

$$T(x, y) = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a1 & a2 \\ a3 & a4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b1 \\ b2 \end{bmatrix} \quad (1)$$

O escrita en forma vectorial:

$$T = A X + B \quad (2)$$

El vector  $[b1, b2]$  tiene en cuenta la traslación espacial, mientras la matriz  $[a1, a2, a3, a4]$  tiene en cuenta la rotación y el cambio de escala. En otras palabras, un punto de la imagen ubicado en  $(x, y)$  se desplaza a una posición  $(x', y')$  según la transformación  $T(x, y)$ .

Ahora bien, también es posible expresar el cambio de posición como una velocidad, para lo cual se puede escribir la ecuación vectorial de velocidad:

$$V = \Omega \times r + V_L \quad (3)$$

Donde:

$\Omega = (w_x, w_y, w_z)$  es el vector velocidad angular

$r = (x, y, z)$  es la posición del punto en la imagen

$V = (v_x, v_y, v_z)$  es el vector velocidad en el punto

Se han propuesto una gran cantidad de algoritmos para la detección del movimiento en imágenes de vídeo. Cada autor establece algunas mejoras o nuevas formas de resolver el problema que supone la extracción del flujo óptico de una serie de imágenes. Sin embargo, básicamente se utilizan dos formas de acometer el problema, con diversas variaciones sobre ellos.

El primero y quizás el más empleado es el método de los gradientes espacio-temporales el cual consiste en determinar los cambios (gradientes) espaciales y temporales del patrón de grises de la imagen y a partir de ellos obtener el flujo óptico. Este método tiene como base el hecho de que en los cambios de intensidad de la imagen está contenida la información del movimiento de los objetos.

Tiene la ventaja de que genera un campo vectorial muy denso (un vector por píxel) que resulta muy útil en algunas aplicaciones. Sin embargo adolece del problema de ser muy sensible al ruido en las imágenes y no puede determinar el flujo bajo ciertas circunstancias.

El segundo método es la segmentación de las imágenes, consiste en identificar una serie de características locales (objetos, bloques, segmentos, etc.) en una imagen y tratar de obtener las correspondencias con la otra imagen, y por último determinar los movimientos. Estas técnicas adolecen del problema de que no generan un campo vectorial muy denso y de que su complejidad crece con el número de particularidades con el que cuenta la imagen.

## 1. ESTIMACION DE MOVIMIENTO

Para la estimación del movimiento en una primera aproximación, es necesario tener por lo menos un par de imágenes consecutivas, sin embargo, el uso de más imágenes aporta una mayor información y, por ello, puede obtenerse una estimación más real del movimiento. Para determinar el movimiento algo muy utilizado es la ecuación restringida del flujo óptico, en la cual se basan la gran mayoría de los algoritmos que determinan el flujo a partir del gradiente espacio temporal.

La recuperación del campo de movimiento parece ser la tarea esencial de cualquier sistema de visión artificial que extraiga información a partir de una secuencia de imágenes, sin embargo el único dato disponible es la variación espacial y temporal del patrón de brillo de la imagen. De ellas es posible obtener una aproximación del campo de movimiento denominado flujo óptico. El campo de movimiento y el flujo óptico son iguales sólo en el caso en que las variaciones espaciales del patrón de brillo correspondan a características estructurales de las superficies. La estimación de flujo óptico, es de vital importancia para el modelo que se desea aplicar, consiste en estimar la correspondencia entre dos *frames*, si las componentes de desplazamiento o velocidad de cada píxel son tratadas como variables independientes, entonces el número de ecuaciones es igual al número de píxeles, donde cada vector tiene dos componentes.

La ecuación restringida del flujo óptico (Optical Flow Constraint Equation OFCE), tiene como condiciones: la iluminación de los objetos presentes en las imágenes es uniforme, la reflectancia de los objetos varía con suavidad y sin discontinuidades espaciales, no existen objetos que se oculten unos a otros y no hay nuevos objetos en la secuencia de imágenes, bajo estas condiciones se puede afirmar que la intensidad de cada punto en particular no cambia, lo que cambia es su posición:

$$E(x, y, t) = E(x + \delta x, y + \delta y, t + \delta t) \quad (4)$$

En las cercanías del punto  $(x, y, t)$  se tiene:

$$E(x, y, t) = E(x, y, t) + \delta x \frac{\partial E}{\partial x} + \delta y \frac{\partial E}{\partial y} + \delta t \frac{\partial E}{\partial t} + TOS \quad (5)$$

Si se sustrae  $E(x, y, t)$  en ambos lados y se divide en  $\delta t$ , y no se tienen en cuenta los TOS (términos de orden superior), y haciendo que el límite  $\delta t$  tienda a 0, se obtiene la ecuación restringida del flujo óptico:

$$\frac{dx}{dt} \frac{\partial E}{\partial x} + \frac{dy}{dt} \frac{\partial E}{\partial y} + \frac{\partial E}{\partial t} = 0 \quad (6)$$

Siendo

$$V_x E_x + V_y E_y + E_t = 0 \quad (7)$$

Esta ecuación nos relaciona la velocidad en los puntos en la imagen y el cambio en los niveles de gris de la imagen.

$$\vec{\nabla} E \bullet \vec{V} + \frac{\partial E}{\partial t} = 0 \quad (\text{OFCE}) \quad (8)$$

Siendo  $\vec{\nabla} E$  el vector gradiente del patrón de grises  $\left( \frac{\partial E}{\partial x}, \frac{\partial E}{\partial y} \right)$ , y  $\vec{V}$  el vector velocidad  $(V_x, V_y)$ .

Para determinar el vector velocidad se emplean diferentes métodos entre los que se encuentran el método del gradiente, el método Nagel, entre otros.

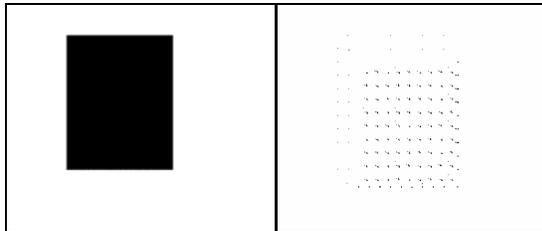


Fig. 3. Flujo óptico de la imagen (Desplazamiento diagonal)

### 5.1. Determinación del flujo óptico por gradientes.

Se observa que podemos determinar la componente del flujo óptico en la dirección del gradiente del brillo de la imagen  $(E_x, E_y)$ . La información disponible en cada punto de una secuencia de imágenes es sólo la componente del campo de movimiento en la dirección del vector gradiente del patrón de brillo. Esto se conoce con el nombre de problema de apertura, podemos expresarlo como:

$$\|\vec{\nabla} E\| V_{\perp} = -E_t \quad (9)$$

$$V_{\perp} = -\frac{E_t}{\|\vec{\nabla} E\|} = -\frac{E_t}{\sqrt{E_x^2 + E_y^2}} \quad (10)$$

Donde:  $\|\vec{\nabla} E\|$  es la norma del vector gradiente del brillo y  $V_{\perp}$  es la norma del vector velocidad en la dirección del gradiente del brillo, expresando en forma vectorial el vector velocidad en la dirección del gradiente de brillo se tiene:

$$\vec{V}_{\perp} = -\frac{E_t}{\sqrt{E_x^2 + E_y^2}} \hat{e} \quad (11)$$

Donde  $\hat{e}$  es el vector unitario en la dirección del gradiente de brillo.

Se concluye que no puede ser determinada la componente del movimiento perpendicular al gradiente del brillo, y tampoco puede determinarse en los casos de zonas de igual brillo (*iso-brillo*). La determinación del flujo óptico pasa, por tanto, por el paso previo de determinar los gradientes temporales y espaciales del patrón de grises de la imagen. Horn y Schunck proponen hacer esto con una convolución tridimensional con un cubo de  $2 \times 2 \times 2$  figura 2.

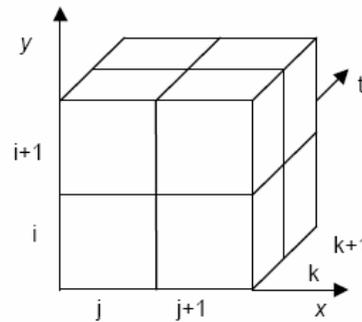


Fig. 4. Convoluciones  $2 \times 2$  para determinar las componentes del vector gradiente espacial y temporal del patrón de brillo.

Para determinar el valor del flujo óptico en las áreas donde falla la ecuación OFCE, Horn y Schunck agregan una condición adicional: los cambios en el flujo óptico deben ser suaves. En otras palabras, no hay cambios bruscos en el movimiento entre puntos cercanos de la imagen.

Adicionalmente, dado que se requiere calcular las derivadas espacio-temporales del patrón de grises de la imagen, es preciso que éste sea diferenciable. Por ello, casi todos los autores utilizan un filtro por convolución (Gausiano por lo general) que suaviza la imagen antes de proceder a calcular el gradiente.

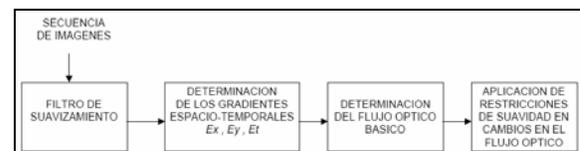


Fig. 5. Determinación del flujo óptico por Horn y Schunck.

### 5.2. Determinación del flujo óptico por el método de Nagel

Emplea derivadas de primer orden para medir el flujo óptico, como una alternativa Nagel sugiere un suavizamiento orientado al cambio exagerado en la intensidad del gradiente (bordes) en un intento por manejar la oclusión, el problema se formula como la minimización de la función:

$$\int \left[ (\nabla^T v + E)^2 + \frac{\alpha^2}{|\nabla E|^2 + 2\delta} \left[ (v_x E_x - v_x E_x)^2 + (v_y E_y - v_y E_y)^2 + \delta(v_x^2 + v_x^2 + v_x^2 + v_y^2) \right] \right] \quad (12)$$

Minimizando con respecto a  $V$  atenúa la variación del flujo en la dirección perpendicular al gradiente, como sugiere se usa  $\delta=1$  (valores inferiores de  $\delta$  producen inestabilidad numéricas.) y  $\alpha=0.5$ , con el uso de iteraciones Gauss-Seidel, la solución puede ser expresada como:

$$v_x^{k+1} = \xi(v_x^k) - \frac{E_x(E_x \xi(v_x^k) + E_y \xi(v_y^k) + E_t)}{E_x^2 + E_y^2 + \alpha^2} \quad (13)$$

$$v_y^{k+1} = \xi(v_y^k) - \frac{E_y(E_x \xi(v_x^k) + E_y \xi(v_y^k) + E_t)}{E_x^2 + E_y^2 + \alpha^2} \quad (14)$$

En estas ecuaciones,  $k$  representa el número de iteraciones, y  $\xi(v_x^k), \xi(v_y^k)$  son conseguidos por:

$$\xi(v_x^k) = v_x^{-k} - 2E_x E_y v_{x,yy} - q^T (\nabla v_x^k) \quad (15)$$

$$\xi(v_y^k) = v_y^{-k} - 2E_x E_y v_{y,xy} - q^T (\nabla v_y^k) \quad (16)$$

Donde

$$q = \frac{1}{E_x^2 + E_y^2 + 2\delta} \nabla E^T \left[ \begin{pmatrix} E_{yy} & -E_{xy} \\ -E_{xy} & E_{xx} \end{pmatrix} + 2 \begin{pmatrix} E_{xx} & E_{xy} \\ E_{xy} & E_{yy} \end{pmatrix} W \right] \quad (17)$$

Donde  $v_x^{-k}$  y  $v_y^{-k}$  son el promedio de los vecinos de  $v_x^k$  y  $v_y^k$ , y  $W$  es el peso de la matriz.

$$W = (E_x^2 + E_y^2 + 2\delta)^{-1} \begin{pmatrix} E_y^2 + \delta & -E_x E_y \\ -E_x E_y & E_x^2 + \delta \end{pmatrix} \quad (18)$$

Para este método se emplea la velocidad inicial a cero, un filtro Guassiano de desviación estándar de 1.5 pix/space, derivadas de segundo orden usando 4-puntos centrales para la operación, para las derivadas de primer orden 2 puntos  $\frac{1}{2}(1,0,-1)$ , se emplearon 100 iteraciones;

### 5.3. Determinación del movimiento a partir de las derivadas de segundo orden.

Se emplea el Hessian de la imagen para determinar el movimiento, la ecuación para este método esta determinada por las derivadas de segundo orden y el sistema:

$$\begin{bmatrix} E_{xx} & E_{yx} \\ E_{xy} & E_{yy} \end{bmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix} + \begin{pmatrix} E_{tx} \\ E_{ty} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (19)$$

La ecuación (19) se deriva de (4) o de la conservación de  $\nabla E$ ,  $d\nabla E = 0$ . Estrictamente hablando, la conservación de  $\nabla E$  implica que no deben presentarse ni rotación ni dilatación, esto es algo fuertemente restrictivo. Para medir la velocidad en la imagen, asumiendo que  $d\nabla E = 0$ , la restricción puede ser usada para aislar o simultáneamente con (6) para ceder sobre un determinado sistema de ecuaciones. Las derivadas de segundo orden no es usualmente una medida suficientemente exacta para determinar la componente de la velocidad tangencial de  $v$ . Como consecuencia, la velocidad estimada de los métodos de segundo orden es frecuentemente asumida como una medida dispersa y menos exacta que las estimadas con métodos de derivadas de primer orden.

### 5.4. Determinación del movimiento a partir de los gradientes.

Yamamoto introduce un método para determinar el movimiento con una mayor exactitud, sin necesidad de calcular previamente el flujo óptico, Yamamoto parte de la ecuación de OFCE (8) y la sustituye en la ecuación de velocidad (3), de ellas obtiene una representación para tres casos diferentes, el primero de ellos es el caso tridimensional (3-D), con la siguiente ecuación

$$E_x v_x + E_y v_y - E_y \omega_x z + E_x \omega_y z + (E_y x - E_x y) \omega_z + E_t = 0 \quad (20)$$

Donde las incógnitas son  $(v_x, v_y, \omega_x, \omega_y, \omega_z)$ , y se supone que  $z$  es conocida; En caso que el movimiento este restringido al plano perpendicular al eje óptico (2-D),  $\omega_x = \omega_y = 0$ , la ecuación se simplifica a:

$$E_x v_x + E_y v_y + (E_y x - E_x y) \omega_z + E_t = 0 \quad (21)$$

Y en caso que no este permitida la rotación (1-D)  $\omega_z = 0$ , obteniéndose la misma ecuación OFCE.

$$E_x v_x + E_y v_y + E_t = 0 \quad (22)$$

Las ecuaciones pueden ser escritas para  $n$  puntos de la imagen, donde  $n$  es el número de incógnitas de la ecuación, adicionalmente las ecuaciones pueden representarse en forma matricial a resolver  $\mathbf{AX}=\mathbf{B}$

$$(23)$$

Donde la  $i$ -ésima fila de los vectores A, X, B esta determinada por:

$$A = (E_{xi}, E_{yi}, -E_{yi}z_i, E_{xi}z_i, E_{yi}x_i, E_{xi}y_i)$$

$$X = (v_x, v_y, \omega_x, \omega_y, \omega_z) \quad (24)$$

$$B = (E_{ii})$$

Se observa que para obtener los parámetros de movimiento hace falta determinar, los gradientes en alguna forma y luego tomar una serie de puntos de la imagen y aplicar la ecuación 14, para el caso 3-D,  $z$  debe ser obtenida por algún método.



Fig. 6. Determinación de los parámetros del movimiento por el método de Yamamoto

Para este método propuesto por Yamamoto se trabaja con las ecuaciones para 1-D, con lo que los vectores dados en (24) quedan representados por:

$$A = (E_{xi}, E_{yi})$$

$$X = (v_x, v_y) \quad (25)$$

$$B = (E_{ii})$$

Solucionando el sistema anterior se logra determinar la velocidad del movimiento., al igual que los métodos nombrados anteriores determinar los gradientes espacio temporales nos determinara en gran medida la exactitud de la estimación dada.

## 6. CONCLUSIONES

El trabajo presentado es uno de los primeros intentos de implementación de algoritmos matemáticos para visión artificial, relacionado con secuencias de imágenes y objetos en movimiento. Dentro del proceso de experimentación y validación pueden determinarse parámetros que incrementen la efectividad de los algoritmos desarrollados, tales como la resolución de las imágenes, los métodos de calculo del gradiente, los parámetros fijados en las ecuaciones, etc. También es posible aumentar la eficiencia del proceso reduciendo el esfuerzo de procesamiento, modificando el orden de las operaciones. La utilización de otros algoritmos desarrollados por el

grupo de procesamiento de señales puede mejorar la precisión de la estimación del movimiento. Aunque los algoritmos desarrollados solo determinan el movimiento traslacional, se puede con ellos determinar deformaciones suaves de los objetos.

Se espera en el futuro desarrollar mas algoritmos de este tipo que ayuden a aumentar la velocidad y exactitud que hasta el momento se han obtenido con estos algoritmos, para este efecto queda abierta la posibilidad de implementar algoritmos basados en la segmentación, así como nuevas teorías para el procesamiento de secuencias de imágenes.

Dependiente de la aplicación y tipo de movimiento que se desee determinar o estimar se debe acoger un modelo para el movimiento presente en la imagen, los métodos presentados no son los únicos existentes, pero brindan una introducción a la estimación del movimiento en secuencia de imágenes; El estudio de esta teoría requiere fundamentacion en procesamiento digital de imágenes, los resultados hasta el momento obtenidos nos incitan a seguir investigando en este campo hasta obtener los mejores resultados.

## REFERENCIAS

- Bravo V, Antonio. "Procesamiento Digital de Imágenes."  
<http://www.ing.ula.ve/~abravo/document/tutoria1/imagenes/indice.html#rb>
- Escalera Hueso, Arturo de la. (2001). "visión por computador, Fundamentos y métodos". Prentice Hall, Madrid España.
- Faundez Zanuy, Marcos, (2001). "Tratamiento digital de voz e imagen y aplicación a la multimedia)," Alfaomega Grupo Editor. México DF.
- Forsyth D. y J. Ponce. Tercer borrador de la versión online del libro Computer visión. A modern approach  
<http://www.etsi2.ugr.es/depar/ccia/mia/complementario.html#General>
- Nalwa, V.S. (1993), A guided tour of computer visión, Addison & Wesley.
- Sucar, L. Enrique. Giovanni Gómez. Procesamiento de imágenes y visión computacional, Departamento de Computación. Tecnológico de Monterrey. Campus Cuernavaca.  
<http://www.mor.itesm.mx/~vision/Libro.html>