

Filtros adaptativos para el tratamiento de la oclusión en el seguimiento de objetos

E. PARRILLA¹, D. GINESTAR², J.L. HUESO², J. RIERA¹, J. R. TORREGROSA²

¹ *Dpto. Física Aplicada, Universidad Politécnica de Valencia, Camino de Vera s/n, Ed. 8L, E-46022 Valencia. E-mails: edparber@fis.upv.es, jrriera@fis.upv.es.*

² *Dpto. de Matemática Aplicada, Universidad Politécnica de Valencia, Camino de Vera s/n, E-46022 Valencia. E-mails: dginesta@mat.upv.es, jlhueso@mat.upv.es, jrtorre@mat.upv.es.*

Palabras clave: Seguimiento de objetos, video, oclusión, flujo óptico, filtros adaptativos

Resumen

En este artículo, se estudia un algoritmo rápido y sencillo para el seguimiento de objetos, basado en el cálculo del flujo óptico de un conjunto poco denso de puntos en los distintos fotogramas de un vídeo. Los métodos de este tipo no son capaces de tratar la oclusión de los objetos en movimiento. Para mejorar este comportamiento, se propone el uso de filtros adaptativos para predecir las velocidades instantáneas esperadas de los objetos. Estas velocidades esperadas se comparan con las calculadas mediante el flujo óptico y se usan como indicadores del buen funcionamiento del método.

1. Introducción

Una de las formas de realizar el seguimiento de un objeto en una secuencia de vídeo se basa en el cálculo del flujo óptico asociado al objeto. El flujo óptico se puede considerar como el campo de velocidades efectivas 2-D, resultado de la proyección de las velocidades de los puntos tridimensionales, correspondientes a las superficies de una escena, sobre el plano visual del sensor. En este trabajo, para determinar el flujo óptico, se ha usado el algoritmo de Lucas y Kanade. Mediante esta técnica se han obtenido buenos resultados en diferentes aplicaciones [1, 6, 7].

Uno de los problemas que plantea el cálculo del flujo óptico es la presencia de obstáculos que oculten temporalmente el objeto que se está siguiendo. Este problema se conoce como un problema de oclusión y da lugar a errores en el seguimiento del objeto y su consecuente

pérdida. Para solucionar este problema, haremos uso de filtros adaptativos [2, 4], que nos proporcionarán una predicción de la velocidad del objeto a seguir, que nos servirá de indicador del funcionamiento del algoritmo de seguimiento.

2. Cálculo del flujo óptico

Para calcular el flujo óptico es necesario asumir que las variaciones de intensidad de la imagen son causadas únicamente por el desplazamiento de los objetos, sin considerar los cambios de iluminación. Esta hipótesis, propuesta inicialmente por Horn y Schunck [3], supone que las estructuras de intensidad presentes en la imagen, a nivel local, se mantienen constantes a lo largo del tiempo, al menos durante intervalos temporales pequeños. Si se pretende seguir un punto de la imagen (x, y) que tiene una velocidad $\vec{v} = (v_x, v_y)$, e $I(x, y, t)$ es la intensidad de la imagen en el punto (x, y) y en el instante t , ha de cumplirse transcurrido un cierto intervalo de tiempo Δt que

$$I(x, y, t) \approx I(x + v_x \Delta t, y + v_y \Delta t, t + \Delta t) . \quad (1)$$

Haciendo un desarrollo en serie de Taylor y quedándose a primer orden en Δt , se obtiene la ecuación del flujo óptico

$$\partial_x I v_x + \partial_y I v_y + \partial_t I = 0 , \quad (2)$$

donde $\partial_x I$, $\partial_y I$, $\partial_t I$ son las derivadas parciales de la imagen y $\vec{v} = (v_x, v_y)$ representa el vector de flujo en cada punto.

2.1. Algoritmo de Lucas y Kanade

Uno de los principales problemas que existen en el seguimiento de un punto a lo largo de una secuencia de imágenes es la imposibilidad de seguir un único píxel, a menos que éste tenga una intensidad muy diferente al resto de píxeles vecinos. Aun así, el ruido de la imagen puede hacer cambiar el valor de dicho píxel y confundirlo con los puntos más cercanos. Debido a esto, el algoritmo de Lucas y Kanade [9] no sigue píxeles aislados sino que trabaja con ventanas de píxeles con la suficiente textura que se supone que se mueve a una cierta velocidad. Desafortunadamente, los puntos de una misma ventana pueden tener comportamientos diferentes, es decir, pueden tener distintas velocidades y la solución tomada por Lucas y Kanade consiste en estimar la velocidad promedio.

Para formalizar el algoritmo, consideremos una ventana de la imagen, W , que se desplaza a una velocidad (v_x, v_y) , desconocida. Esta velocidad se estima suponiendo que es la que hace mínima la función error

$$\varepsilon^2 = \int_W (\partial_x I v_x + \partial_y I v_y + \partial_t I)^2 \omega(x, y) dx dy ,$$

donde $\omega(x, y)$ es una función peso que depende de la posición del píxel en la ventana. Las ecuaciones normales asociadas a este problema son

$$\begin{aligned} \int_W \left((\partial_x I)^2 v_x + \partial_x I \partial_y I v_y + \partial_x I \partial_t I \right) \omega dx dy &= 0 , \\ \int_W \left(\partial_y I \partial_x I v_x + (\partial_y I)^2 v_y + \partial_y I \partial_t I \right) \omega dx dy &= 0 . \end{aligned} \quad (3)$$

Una vez elegidas una función peso ω y unas discretizaciones para las derivadas de I sobre los píxeles de la imagen y de la integral, el sistema de ecuaciones (3) permite estimar la velocidad de la ventana. Para mejorar el cálculo del flujo óptico de la ventana se realiza un proceso iterativo: dada una estimación inicial de las velocidades, se deshace el desplazamiento entre los dos fotogramas considerados y se vuelve a calcular el flujo óptico entre las dos imágenes obtenidas. Este proceso se repite hasta obtener un flujo óptico por debajo de una tolerancia. El valor final de flujo óptico se calcula como la suma de todos los desplazamientos efectuados. Este proceso iterativo se combina con un algoritmo piramidal, que hace uso de las imágenes con distintos grados de resolución, para mejorar el cálculo del flujo óptico.

3. Filtros adaptativos

Con el objeto de obtener un indicador del funcionamiento del algoritmo para el seguimiento de un objeto basado en el cálculo del flujo óptico, cada una de las componentes de las velocidades estimadas a partir del flujo óptico para los distintos fotogramas del vídeo, se consideran como una serie temporal. Se utilizarán filtros adaptativos para obtener una predicción de las mismas, que se pueda comparar con las velocidades obtenida a partir del flujo óptico. Se estudiarán dos filtros, el filtro RLS y el filtro de Kalman.

3.1. Filtro RLS

Dada una señal x_n , $n = 0, \dots, N_T$, se asume para dicha señal un modelo DAR(p), o sea,

$$x_k = -(a_1)_k x_{k-1} - (a_2)_k x_{k-2} - \dots - (a_p)_k x_{k-p} + \varepsilon_k, \quad (4)$$

donde $(a_1)_k, \dots, (a_p)_k$ son parámetros a determinar y ε_k un proceso de ruido blanco [4]. Asumiendo este modelo para $k = p, \dots, N$, se obtiene el siguiente sistema de ecuaciones

$$X_N = C_N a_N + \epsilon_N, \quad (5)$$

donde

$$X_N = \begin{bmatrix} x_p \\ x_{p+1} \\ \vdots \\ x_N \end{bmatrix}, \quad C_N = \begin{bmatrix} x_{p-1} & x_{p-2} & \cdots & x_0 \\ x_p & x_{p-1} & \cdots & x_1 \\ \vdots & & & \vdots \\ x_{N-1} & x_{N-2} & \cdots & x_{N-p} \end{bmatrix},$$

$$a_N = \begin{bmatrix} -(a_1)_N \\ -(a_2)_N \\ \vdots \\ -(a_p)_N \end{bmatrix}, \quad \epsilon_N = \begin{bmatrix} \varepsilon_p \\ \varepsilon_{p+1} \\ \vdots \\ \varepsilon_N \end{bmatrix}.$$

Los coeficientes del modelo, a_N , se obtienen como aquellos coeficientes que minimizan la función de error

$$J_N = \frac{1}{2} \sum_{k=p}^N \lambda^{N-k} \varepsilon_k \varepsilon_k = \frac{1}{2} (X_N - C_N a_N)^T \Lambda_N (X_N - C_N a_N). \quad (6)$$

Se ha introducido la matriz de pesos

$$\Lambda_N = \begin{bmatrix} \lambda^{N-p} & 0 & \cdots & 0 \\ 0 & \lambda^{N-p-1} & & \vdots \\ \vdots & \vdots & & \\ 0 & \cdots & & 1 \end{bmatrix},$$

donde $0 < \lambda \leq 1$ es una constante que controla la importancia de las muestras antiguas de la señal en la función de error J_N , y que se conoce como factor de olvido del método. Los coeficientes a_N son las soluciones de la ecuación

$$\frac{\partial J_N}{\partial a_N} = 0,$$

es decir,

$$a_N = [C_N^T \Lambda_N C_N]^{-1} C_N^T \Lambda_N X_N. \quad (7)$$

Consideremos ahora $N + 1$ muestras de la señal. El sistema (5) se convierte en

$$\begin{bmatrix} X_N \\ x_{N+1} \end{bmatrix} = \begin{bmatrix} C_N \\ c_{N+1}^T \end{bmatrix} a_{N+1} + \epsilon_{N+1}, \quad (8)$$

donde

$$c_{N+1}^T = [x_N, x_{N-1}, \dots, x_{N-p+1}].$$

a_{N+1} serán los coeficientes que minimizan la nueva función de error,

$$J_{N+1} = \frac{1}{2} \sum_{k=p}^{N+1} \lambda^{N+1-k} \epsilon_n \epsilon_N = \frac{1}{2} \epsilon_{N+1}^T \Lambda_{N+1} \epsilon_{N+1}.$$

La solución de este problema es

$$a_{N+1} = [C_{N+1}^T \Lambda_{N+1} C_{N+1}]^{-1} C_{N+1}^T \Lambda_{N+1} X_{N+1}. \quad (9)$$

La obtención de la inversa de la matriz $C_{N+1}^T \Lambda_{N+1} C_{N+1}$ es un proceso de un alto coste computacional, especialmente cuando se considera un número grande de muestras. Para resolver este problema, se tiene en cuenta que

$$\begin{aligned} C_{N+1}^T \Lambda_{N+1} C_{N+1} &= \begin{bmatrix} C_N^T & c_{N+1} \end{bmatrix} \begin{bmatrix} \lambda \Lambda_N & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} C_N \\ c_{N+1}^T \end{bmatrix} = \\ &= C_N^T \lambda \Lambda_N C_N + c_{N+1} c_{N+1}^T. \end{aligned}$$

Utilizando el lema de inversión [5]

$$(A + BD)^{-1} = A^{-1} - A^{-1}B(I + DA^{-1}B)^{-1}DA^{-1},$$

con

$$A = \lambda C_N^T \Lambda_N C_N, \quad B = c_{N+1}, \quad D = c_{N+1}^T,$$

se puede describir

$$[C_{N+1}^T \Lambda_{N+1} C_{N+1}]^{-1} = \frac{1}{\lambda} \left([C_N^T \Lambda_N C_N]^{-1} - \frac{[C_N^T \Lambda_N C_N]^{-1} c_{N+1} c_{N+1}^T [C_N^T \Lambda_N C_N]^{-1}}{\lambda + c_{N+1}^T [C_N^T \Lambda_N C_N]^{-1} c_{N+1}} \right). \quad (10)$$

Introduciendo la notación

$$P_N = [C_N^T \Lambda_N C_N]^{-1}, \quad K_{N+1} = \frac{P_N c_{N+1}}{\lambda + c_{N+1}^T P_N c_{N+1}},$$

se obtiene la recurrencia

$$P_{N+1} = \frac{1}{\lambda} (P_N - K_{N+1} c_{N+1}^T P_N). \quad (11)$$

De esta forma, usando (9), se puede escribir [4]

$$a_{N+1} = a_N + K_{N+1} (x_{N+1} - c_{N+1}^T a_N)$$

Para utilizar el método, primero se debe realizar un proceso de inicialización, considerando N_{in} muestras de la señal y calculando

$$P_{N_{in}} = [C_{N_{in}}^T \Lambda_N C_{N_{in}}]^{-1}, \quad a_{N_{in}} = P_N C_{N_{in}}^T \Lambda_{N_{in}} X_{N_{in}}.$$

Después de este paso, se utiliza el siguiente proceso recursivo

$$\begin{aligned} K_{N+1} &= \frac{P_N c_{N+1}}{\lambda + c_{N+1}^T P_N c_{N+1}}, \\ P_{N+1} &= \frac{1}{\lambda} (P_N - K_{N+1} c_{N+1}^T P_N), \\ a_{N+1} &= a_N + K_{N+1} (x_{N+1} - c_{N+1}^T a_N). \end{aligned} \quad (12)$$

Finalmente, una vez obtenido a_{N+1} se estima

$$x_{N+1} = C_{N+1}^T a_{N+1}. \quad (13)$$

3.2. Filtro de Kalman

El filtro de Kalman es una técnica de filtrado muy general que se puede aplicar a la solución de muchos problemas como estimación óptima, predicción, filtrado del ruido y control óptimo estocástico [5]. En este apartado, se resumirán los aspectos básicos de la técnica de filtrado de Kalman utilizado como método adaptativo. Al igual que en el método anterior, se asumirá un modelo DAR(p) para la señal que, para la muestra k de la señal se puede escribir de la forma

$$x_k = -(a_1)_k x_{k-1} - (a_2)_k x_{k-2} - \dots - (a_p)_k x_{k-p} + \varepsilon_k = c_k^T a_k + \varepsilon_k \quad (14)$$

Además, para el vector de coeficientes se supone una evolución temporal siguiendo la ecuación

$$a_{k+1} = a_k + w_k \quad (15)$$

donde w_k es el proceso de ruido del sistema y ε_k es la medida de la secuencia de ruido. El filtro de Kalman asume una iteración de la forma

$$\tilde{a}_{N+1} = \tilde{a}_N + K_{eN} (x_N - \tilde{x}_N) = \tilde{a}_N + K_{eN} (x_N - c_N^T \tilde{a}_N) \quad (16)$$

Para obtener un método con similares propiedades de memoria que el método RLS con el factor de olvido comentado anteriormente [8], se empieza a partir de las ecuaciones generales del filtro de Kalman, (14) y (15) y se asume que

$$E [w_j w_k^T] = Q_k \delta_{jk} = \lambda^2 \delta_{jk}, \quad E [\varepsilon_j \varepsilon_k^T] = R_k \delta_{jk} = 1,$$

y se determinan los valores óptimos de los parámetros \tilde{a}_{N+1} de forma que la matriz de covarianza

$$P_{N+1} = E [e_{N+1} e_{N+1}^T] \quad (17)$$

es mínima. Para inicializar el filtro de Kalman, se toman la matrices

$$P_0 = 0_{p \times p}, \quad a_0 = 0_{p \times 1}$$

y para el resto de las muestras de la señal se utilizan las siguientes ecuaciones recursivas (ver [5] para la deducción de estas expresiones)

$$\begin{aligned} K_{eN} &= \frac{P_N c_N}{1 + c_N^T P_N c_N}, \\ a_{N+1} &= a_N + K_{eN} (x_N - c_N^T a_N), \\ P_{N+1} &= P_N - \frac{P_N c_N c_N^T P_N}{1 + c_N^T P_N c_N} + \lambda^2 I, \end{aligned} \quad (18)$$

donde se han identificado los parámetros del modelo a_N de la ecuación (14) con la estimación de estos parámetros \tilde{a}_N de la ecuación (16).

3.3. Combinación del flujo óptico y el filtro adaptativo

Para la utilización de los filtros adaptativos en combinación con el algoritmo de cálculo del flujo óptico, se seleccionan los puntos a seguir y el tamaño de las ventanas. Se calcula el flujo óptico, mediante el algoritmo de Lukas y Kanade para N_{in} fotogramas, obteniendo las series $(v_{x0}, v_{y0}), \dots, (v_{xN_{in}}, v_{yN_{in}})$, que se utilizan para iniciar los coeficientes del filtro. Posteriormente, para el fotograma N , se calculan las velocidades v_{xN} y v_{yN} realizando los siguientes pasos:

- 1) Con el algoritmo de flujo óptico, se estiman las velocidades v_{xN}^{fo} y v_{yN}^{fo} .
- 2) Con el filtro adaptativo, a partir de las velocidades anteriores, se estiman las velocidades v_{xN}^{fa} y v_{yN}^{fa} .
- 3) Si $\left| v_{x,yN}^{fo} - v_{x,yN}^{fa} \right| < \text{tol}$, entonces $v_{x,yN} = v_{x,yN}^{fo}$.
Si no $v_{x,yN} = v_{x,yN}^{fa}$.

Con esto se consigue que el movimiento del objeto que se sigue tenga una dinámica causal, basada en el movimiento que tenía en los fotogramas anteriores. Corrigiendo el algoritmo del flujo óptico, en algunos problemas de oclusión.

4. Resultados numéricos

El algoritmo expuesto en este trabajo se ha probado en diferentes secuencias de vídeo, estudiando vídeos sintéticos y vídeos reales de tráfico urbano. Se han hecho las pruebas utilizando ventanas de 7×7 píxeles, $N_{in} = 7$ fotogramas, un filtro autoregresivo de orden 2, una tolerancia de $tol = 1$ píxel y un factor de olvido $\lambda = 0,99$.

En la figura 1 se pueden apreciar distintos fotogramas de una secuencia sintética realizada mediante el programa 3D Studio Max [®] que consiste en un objeto que traza una curva, pasando en un determinado momento por debajo de un obstáculo.

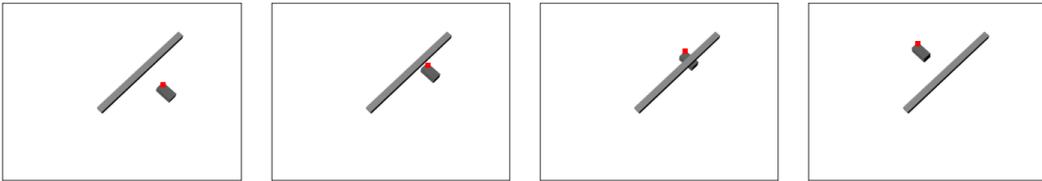


Figura 1: Seguimiento de un objeto en una secuencia sintética.

La figura 2 pertenece a una secuencia de vídeo de una vía urbana de Valencia. En ella se puede observar como la trayectoria de los vehículos queda parcialmente oculta por una farola.



Figura 2: Seguimiento de una secuencia de tráfico real.

En ambas secuencias, sin la utilización de filtros adaptativos, se produce un error en el seguimiento del objeto al llegar al obstáculo. Como se puede apreciar en las imágenes, mediante el uso de los filtros adaptativos, los objetos se siguen correctamente a lo largo de toda la secuencia.

Para efectuar las pruebas se ha seleccionado un punto del objeto a seguir mediante el criterio de selección utilizado por el algoritmo de Lucas y Kanade, que elige los puntos óptimos para realizar el seguimiento. El cálculo del flujo óptico en los primeros 7 fotogramas de la secuencia se utiliza para inicializar los coeficientes del filtro adaptativo. Cuando el punto que se está siguiendo queda oculto por el obstáculo, se produce un error en los valores del flujo óptico que provoca que la diferencia de dichos valores con los calculados mediante el filtro adaptativo sea mayor que la tolerancia por lo que se toman como válidos los valores predichos y, de esta forma, se puede predecir el movimiento del punto en los fotogramas en los que queda oculto. Una vez pasado el obstáculo, los valores de flujo óptico

vuelven a ser semejantes a los predichos por el filtro. De esta forma se toman de nuevo como los valores fiables del movimiento y se utilizan para actualizar los coeficientes del filtro adaptativo.

Se ha comprobado que un factor crítico para el buen funcionamiento del método son los valores dados a los distintos parámetros usados en el algoritmo. De esta forma, un valor de tolerancia muy pequeño provocará que se tomen como válidos los valores predichos aun cuando no exista ningún obstáculo. Por el contrario, un valor demasiado grande conllevará que no se detecte la presencia de ninguna oclusión.

El valor del factor de olvido λ también es determinante ya que un valor demasiado pequeño puede provocar que el filtro adaptativo no se ajuste a la trayectoria del objeto a seguir por lo que se puede producir la pérdida del objeto cuando reaparece después de la oclusión.

5. Conclusiones

En esta comunicación se ha analizado el problema que presenta la oclusión en el seguimiento de objetos en secuencias de imágenes. Se ha propuesto el uso de filtros adaptativos para predecir el movimiento del blanco y poder aproximar su trayectoria durante el periodo de tiempo que queda oculto.

Por otro lado, se ha estudiado el efecto que producen los distintos parámetros en el comportamiento del método, hasta llegar a unos valores óptimos. Finalmente, se han mostrado dos ejemplos que verifican el buen funcionamiento del algoritmo.

Referencias

- [1] Fabrice Bourel, Claude C. Chibelushi y Adrian A. Low, *Recognition of facial expressions in the presence of occlusion*, Proceedings of the Twelfth British Machine Vision Conference, Vol. 1 pp. 213-222, 2001.
- [2] S. Haykin. *Adaptive Filter Theory*. Prentice-Hall, Englewood Cliffs NJ, 1986.
- [3] Horn B. K. P. y Schunck B. G., *Determining optical flow*, Artificial Intelligence. 17(1-3) pp. 185-203, 1981.
- [4] L. Ljung. *System Identification. Theory for the User*. Prentice-Hall, Upper Saddle River NJ., 1999.
- [5] Ogata, K., *Discrete-Time Control Systems*. Prentice Hall. Upper Saddle River, NJ, 1987.
- [6] E. Parrilla, J. Riera, Marcos H. Giménez, Juan R. Torregrosa y José L. Hueso, *Vehicle tracking in urban routes by means of Lucas&Kanade algorithm*, Proceedings of the International Conference on Computational and Mathematical Methods in Science and Engineering pp. 438-445, 2005.
- [7] E. Parrilla, J. Riera, Marcos H. Giménez, Juan R. Torregrosa y José L. Hueso, *Cálculo de velocidades mediante un sistema estereoscópico y algoritmos de flujo óptico*, Congreso de Ecuaciones Diferenciales y Aplicaciones, 2005.
- [8] Schlögl, A., Flotzinger, D., Pfurtsceller, G., *Adaptative Autoregressive Modeling used for Single-trial EEG Classification*, Biomedizinische Technik, 42, 162-167, 1997.
- [9] Lukas, B.D., Kanade, T., *An iterative image registration technique with an application to stereovision*. Proceedings of Imaging Understanding Workshop, pp 121-130, 1981.
- [10] Young, P.C., *Nonstationary time series analysis and forecasting*. Progress in Environmental Science, 1, 1, 3-48, 1999.