

FUSIÓN DE ALGORITMOS PARA DETECTAR OBJETOS EN MOVIMIENTO

Juan Vázquez¹, Manuel Mazo², Member IEEE, Jose L. Lázaro², Member IEEE, Carlos A. Luna¹, Jorge Cabello², Jesús Ureña², Juan J. García²

¹Departamento de telecomunicaciones, Universidad de Oriente, Cuba, jfvglez@yahoo.es

²Departamento de Electrónica, Universidad de Alcalá, España, mazo@depeca.uah.es

RESUMEN

La detección de objetos en movimiento es una tarea común en aplicaciones de visión computacional. En este trabajo se presenta un estudio sobre la fusión de dos algoritmos de detección de movimiento. Esta fusión se realiza, aplicando a las imágenes un algoritmo de detección de movimiento que utiliza el método del Análisis de Componentes Principales, y el resultado de éste algoritmo se usa como entrada para el algoritmo de detección de movimiento basado en un test de relación de probabilidades de orden reducido. Los resultados alcanzados muestran que la segmentación y la umbralización es menos afectada por el ruido, cuando se fusionan los algoritmos, esto queda demostrado en las gráficas y los resultados cuantitativos que se incluyen en este trabajo.

Palabras claves— análisis de las componentes principales, detección de objetos en movimiento, transporte ferroviario, vigilancia mediante visión, Visión computacional.

1. INTRODUCTION

La detección de movimiento es una tarea común en aplicaciones de visión, una de las estrategias más sencillas es realizando diferencias de imágenes sucesivas [1],[2]. Sin embargo el ruido de las imágenes puede dar grandes diferencias donde no hay movimiento, otro problema es que si el obstáculo en movimiento se desplaza muy lentamente en relación a la frecuencia con que se captan las imágenes, la diferencia resultante puede ser tan pequeña que no se detecte el movimiento. Una alternativa es comparar las imágenes mediante un procesamiento por regiones [3], en vez de hacerlo entre píxeles individuales, lo cual reduce la influencia del ruido. Otra forma es mediante la detección de bordes [4] que es más robusto ante los cambios de iluminación. El flujo óptico es otro método de estimación de movimiento [5]

que puede ser implementado mediante correlación, técnicas de gradiente o filtros sensibles a la velocidad. Este método solo es efectivo si el movimiento es significativo entre imágenes. También se emplea el método de comparar cada nueva imagen con una de referencia que contenga sólo la parte estática de la escena [6].

En este trabajo se evalúa la fusión de dos algoritmos con el objetivo de obtener una segmentación y umbralización de las imágenes procesadas, más robusta ante la influencia del ruido.

En ésta propuesta se utiliza el análisis de las componentes principales (PCA) para implementar uno de los algoritmos. El método del PCA ha sido utilizado ampliamente para el reconocimiento de rostros [7], siendo válido también para aplicaciones de visión donde se requiera modelar las zonas que no cambian en una secuencia de imágenes. Otra virtud del método es que permite reducir la dimensión del espacio donde se procesan las imágenes con lo cual se reduce la carga computacional.

El otro algoritmo que se utiliza, esta basado en un test de relación de probabilidades de orden reducido, cuya descripción detallada se puede encontrar en [8].

Este trabajo está estructurado como sigue: En la sección 2 se describe el algoritmo de detección de movimiento que utiliza el método basado en el PCA, la sección 3 explica el algoritmo de detección de movimiento que utiliza el método basado en el test de relación de probabilidades de orden reducido, la sección 4 describe como se fusionan los algoritmos descritos en las secciones 2 y 3, la sección 5 relaciona los resultados experimentales obtenidos y la sección 6 resume las principales conclusiones del trabajo.

2. ALGORITMO BASADO EN EL PCA

El algoritmo de detección de objetos en movimiento basado en el PCA que se utiliza en este trabajo, es el que se propone en [9].

Se parte de una muestra de M imágenes I_1, I_2, \dots, I_M , expresadas como vectores columnas de dimensión $N^2 \times 1$.

La media de esas imágenes es un vector columna definido como:

$$\Psi = \frac{1}{M} \sum_{i=1}^M I_i \quad (1)$$

Y la diferencia entre cada imagen y la media es:

$$\Phi_i = I_i - \Psi \quad (2)$$

El conjunto de las M imágenes se agrupan en una matriz $A=[\Phi_1, \Phi_2, \dots, \Phi_M]$ de dimensión $N^2 \times M$ a la que se le aplica el método de las componentes principales, obteniéndose un conjunto de vectores ortogonales (U) que describen a las M imágenes. Estos son los autovectores de la matriz de covarianza (C) de A .

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_i \Phi_i^T = \frac{1}{M} A A^T \quad (3)$$

El número de autovectores de U puede reducirse si existen autovalores cuyo valor sea despreciable, de manera que sólo se tendrán en cuenta aquellos autovectores cuyos autovalores asociados (λ) sean significativos. Lo anterior se puede expresar cuantitativamente mediante (4) donde D es un valor menor que 1 (por ejemplo 0.9).

$$\frac{\sum_{j=1}^K \lambda_j}{\sum_{j=1}^M \lambda_j} \leq D \quad \text{Siendo } \lambda_1 > \lambda_2 > \dots > \lambda_M \quad (4)$$

La ecuación anterior significa que se utilizarán los K mayores autovectores asociados a los K mayores autovalores que cumplan con (4), obteniéndose una matriz de $N^2 \times K$ ($K < M$). Esta matriz permite representar cualquier imagen en un espacio cuya dimensión es mucho menor (K) que la dimensión del espacio original (N^2) y describe satisfactoriamente la parte estática (fondo) de las imágenes.

Una vez obtenida la matriz de autovectores U , la proyección de una nueva imagen I_i en el espacio transformado se realiza mediante.

$$w = U^T \Phi_i \quad (5)$$

Estos coeficientes w modelarán la parte estática de la imagen y no las zonas que contengan movimiento. Por tanto, la diferencia entre la imagen Φ_i y su proyección recuperada Φ_{if} (6), delatará la presencia de los objetos que hayan entrado en la escena.

$$\Phi_{if} = U w \quad (6)$$

3. ALGORITMO BASADO EN EL TEST DE RELACIÓN DE PROBABILIDADES

El algoritmo de detección de objetos en movimiento basado en el test de relación de probabilidades, que se usa en este trabajo, es el que se propone en [8].

Se parte de considerar que una imagen contaminada con ruido de una secuencia de imágenes puede describirse como:

$$r_m(x) = S_m(x) + n_m(x) \quad (7)$$

donde $S_m(x)$ es la intensidad de la imagen sin ruido en la posición x y $n_m(x)$ es la contribución del ruido. De la misma manera, la diferencia entre dos imágenes sucesivas es:

$$\bar{r}_m(x) = S_m(x) - S_{m-1}(x) + \bar{n}_m(x) \quad (8)$$

donde $\bar{n}_m(x)$ es la diferencia de la intensidad de ruido entre dos imágenes sucesivas, el subíndice $m-1$ indica la imagen anterior y m la imagen actual.

La imagen actual puede describirse como:

$$S_m(x) = S_{m-1}(x - d_m(x)) \quad (9)$$

Donde $d_m(x)$ representa el vector movimiento

Linealizando (9), (8) puede describirse como:

$$\bar{r}_m(x) = \nabla_s^T(x) d_m(x) + \bar{n}_m(x) \quad (10)$$

Para clasificar cada píxel como perteneciente a un objeto en movimiento o como perteneciente al fondo, se usa una prueba de hipótesis binaria sobre la imagen diferencia. Analizando la ecuación (10) se llega a las siguientes hipótesis para cada píxel.

$$H_1 : \bar{r}_m(x) = \nabla_s^T(x) d_m(x) + \bar{n}_m(x)$$

$$(11)$$

$$H_0 : \bar{r}_m(x) = \bar{n}_m(x)$$

La principal causa de la alta carga computacional es el producto indicado en (10).

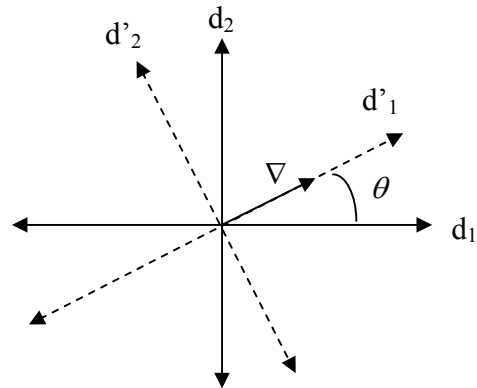


Fig. 1. Rotación de las bases.

Esto se puede reducir apreciablemente, rotando las bases del vector de movimiento, de manera que una de las bases del vector de movimiento coincida con el gradiente, como se muestra en la figura 1, resultando:

$\nabla^T(x)d = \|\nabla(x)\|d'_1$ donde d'_1 es la componente del movimiento en la dirección del vector gradiente.

Rescribiendo (10) queda:

$$\bar{r}_m(x) = \|\nabla(x)\|d'_1(x) + \bar{n}_m(x) \quad (12)$$

Ahora las hipótesis planteadas en (12) quedan como:

$$H_1 : \bar{r}_m(x) = \|\nabla(x)\|d'_1(x) + \bar{n}_m(x) \quad (13)$$

$$H_0 : \bar{r}_m(x) = \bar{n}_m(x)$$

Si \bar{n} es una variable Gaussiana con media cero y varianza σ_n^2 , entonces $\bar{r}(x)$ es una variable Gaussiana con media cero y varianza $\|\nabla(x)\|\sigma_d^2 + \sigma_n^2$ cuando hay movimiento y σ_n^2 cuando el píxel pertenece al fondo. La función de probabilidades logarítmica para las hipótesis de la ecuación (13) es:

$$\Lambda(\bar{r}(x)) = \log\left(\frac{p(\bar{r}(x) | H_1)}{p(\bar{r}(x) | H_0)}\right) \quad (14)$$

donde $p(\bar{r}(x) | H_1)$ es la función de densidad de probabilidades de la diferencia de intensidades entre imágenes, condicionada a la hipótesis H_1 y $p(\bar{r}(x) | H_0)$ es la función de densidad de probabilidades de la diferencia de intensidades entre imágenes, condicionada a la hipótesis H_0

Sustituyendo las expresiones de las funciones de densidad de probabilidades (desarrolladas en [8]) $p(\bar{r}(x) | H_1)$ y $p(\bar{r}(x) | H_0)$ en (14) se llega al test de relación de probabilidades

$$\frac{\bar{r}^2(x)\|\nabla(x)\|^2\sigma_d^2}{2\sigma_n^2(\|\nabla(x)\|^2\sigma_d^2 + \sigma_n^2)} + \frac{1}{2} \log\left(\frac{\sigma_n^2}{\|\nabla(x)\|^2\sigma_d^2 + \sigma_n^2}\right) \underset{H_0}{\overset{H_1}{><}} \gamma \quad (15)$$

H_1 y H_0 son las hipótesis de que el píxel analizado contiene movimiento o no respectivamente. Si el valor de (15) es menor que el umbral γ entonces se cumple la hipótesis H_1 y el píxel analizado se clasifica como perteneciente a un objeto en movimiento, de lo contrario, se cumple la hipótesis H_0 , en cuyo caso, el píxel analizado se clasifica como perteneciente a la parte estática de la escena.

La imagen diferencia \bar{r} y el gradiente de la primera imagen $\nabla(x)$ se pueden obtener directamente de las propias imágenes. Para determinar la varianza del ruido σ_n^2 existen muchos métodos. En este caso se usó el método propuesto en [10].

Por último, la varianza del movimiento se determina como:

$$\sigma_d^2 = \frac{\sum_x \bar{r}^2(x) - N\sigma_n^2}{\sum_x \|\nabla(x)\|^2} \quad (16)$$

4. FUSIÓN DE ALGORITMOS

El objetivo de fusionar los algoritmos descritos en las secciones 2 y 3 es disminuir la influencia del ruido en la segmentación y umbralización de las imágenes. La idea que se propone consiste en:

- Captar una imagen I_i .
- Proyectar la imagen Φ_i usando (5), en el espacio transformado descrito por la matriz de autovectores U (sección 2)
- Obtener la imagen recuperada Φ_{if} usando (6)
- Captar la imagen I_{i+1} .
- Proyectar la imagen Φ_{i+1} usando (5).
- Obtener la imagen recuperada $\Phi_{(i+1)f}$ usando (6)
- Obtener las diferencias de las imágenes

$$\Delta\Phi_i = \Phi_i - \Phi_{if} \quad \text{y} \quad \Delta\Phi_{(i+1)} = \Phi_{(i+1)} - \Phi_{(i+1)f}$$

Las imágenes diferencias $\Delta\Phi_i$ y $\Delta\Phi_{(i+1)}$ serán ahora las imágenes de entradas para el algoritmo descrito en la sección 3 es decir, $\Delta\Phi_i$ corresponderá con $S_{m-1}(x)$ y $\Delta\Phi_{(i+1)}$ corresponderá con $S_m(x)$. Es lógico suponer que las imágenes diferencias tendrán un nivel de ruido mucho menor que las imágenes originales por lo que es de esperar que los resultados que se obtengan con el algoritmo del test de relación de probabilidades sean mejores.

5. RESULTADOS

Para evaluar la propuesta de la sección 4, se utilizó una variada gama de imágenes reales (de 240x320 píxeles) captadas en entornos exteriores, algunas de ellas en condiciones adversas como es el caso de imágenes en presencia de niebla. En las figuras 2 y 3 se muestra un ejemplo de los resultados alcanzados. Aunque en ambas figuras resalta la presencia del objeto, en la figura 2, se

observa que el nivel de ruido es mucho mayor que en caso de la figura 3, donde se usó la propuesta que se hace en este trabajo.

Desde el punto de vista cuantitativo la tasa de falsos positivos obtenidos cuando se utilizó solo el algoritmo del test de relación de probabilidades de orden reducido fue entre un 15 a un 24 % mayor para los diferentes bancos de imágenes respecto a la fusión de algoritmos propuesta en este trabajo

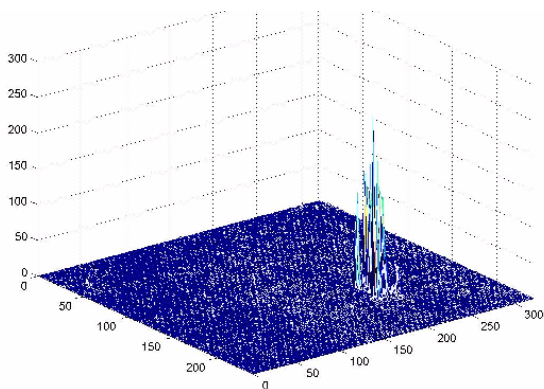


Fig. 2. Resultado de aplicar solo el algoritmo del test de relación de probabilidades de orden reducido.

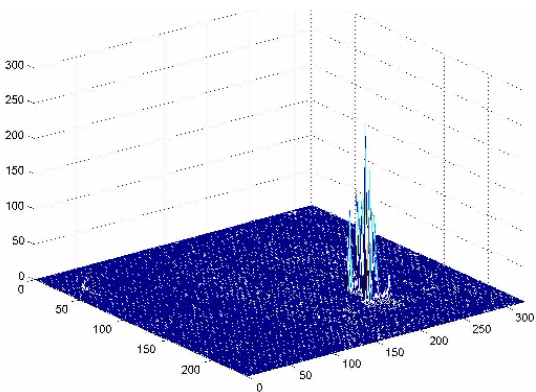


Fig. 3. Resultado de aplicar el algoritmo del PCA y posteriormente el algoritmo del test de relación de probabilidades de orden reducido.

6. CONCLUSIONES

Los resultados experimentales obtenidos demuestran que la fusión del algoritmo que utiliza el PCA con el algoritmo que utiliza el test de relación de probabilidades de orden reducido permite obtener una segmentación con menor nivel de ruido, lo que facilita la umbralización y disminuye la clasificación errónea de píxeles. Todo esto

contribuye a incrementar la robustez y la inmunidad al ruido.

AGRADECIMIENTOS

Este trabajo se ha llevado a cabo gracias al proyecto TELEVIA, financiado por el Ministerio de Ciencia y Tecnología (COO1999-AX049) y a las empresas RENFE y Logytel (Cátedra de control electrónico aplicada al transporte).

REFERENCIAS

- [1] R. Dahyot, P. Charbonnier, F. Heitz, "Unsupervised statistical detection of changing objects in camera-in-motion video", ICIIP-2001 International Conference on Image Processing, Thessaloniki, Grecia, Octubre, 2001.
- [2] D. Jang, G. Young, H. Choi, "Model-based tracking of moving object", Pattern recognition, Vol. 30, No. 6, pp 999-1008, 1997.
- [3] V. Zeljkovic, M. Popovic, "Detection of moving objects in video signal under fast changes of scene illumination", 5th International Conference on Telecommunications in Modern satellite, Cable and broadcasting Services. TELSIKS 2001, Yugoslavia, 19-21 Sept. 2001.
- [4] Q. Gao, A. Parlow, M. Tan, "Object motion detection based on perceptual edge tracking", Second international workshop on digital and computational video, February 08-09, 2001.
- [5] I. Stratmann, "Omnidirectional Imaging and optical flow", Third Workshop on omnidirectional vision (OMNIVIS'02), June 02, 2002, Copenhagen, Denmark.
- [6] N. Oliver, B. rosario, A. Pentland, "A bayesian computer vision system for modeling human interactions", *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, No. 8, august 2000.
- [7] M. turk, A. Pentland, "Eigenfaces for recognition", *Journal of cognitive neuroscience*, vol. 3, No.1, pp. 71-86, 1991.
- [8] Foster, G.J.; Namazi, N.M.; Detection of non-uniform motion in image sequences using a reduced order likelihood ratio test. *Image Processing, 2000. Proceedings. 2000 International Conference on*, Vol: 1, 10-13 Sept. 2000. Page(s): 248 -2511
- [9] J. Vázquez y otros: Detección de obstáculos en movimiento, Seminario Anual de Automática y Electrónica Industrial (SAAEI'03), Vigo, España, Septiembre 2003, ISBN: 84-688-3055-6.
- [10] K. Rank, M. Lendi and R. Unbehauen; Estimation of image noise variance; *Vision, Image and Signal Processing, IEEE Proceedings*, Volume: 146, Issue: 2, April 1999, pag. 80-84.