CAPÍTULO 4: RESULTADOS

4. RESULTADOS

En el desarrollo del proyecto se han utilizado nueve vídeos de anuncios de televisión, todos con un tamaño de imagen de 320x240 píxeles y con una frecuencia de cuadro de 25 cuadros/s. Los vídeos han sido grabados de tres cadenas de televisión distintas y obedecen la nomenclatura *xx_yy*; donde *xx* se refiere al número de vídeo e *yy* a la cadena de televisión. Las tablas 4 y 5 indican el significado de estos códigos:

Tabla 4: Códigos de cadenas de televisión

Código	Cadena
01	TVE
02	Antena 3
03	Tele 5

Tabla 5: Códigos de anuncios

Código	Anuncio
01	Ford
02	MasVital
03	Kinder

Los vídeos se han codificado en MPEG-4 con el programa *Movavi Video Converter*. *Movavi Video Converter* es un nuevo programa revolucionario para convertir vídeo y archivos DVD. Movavi lo fundó en 2004 un grupo de aficionados decididos a simplificar la edición y el procesamiento de vídeo. El producto fue bien recibido y revisado por World, CNet, MacNN, Yahoo!Tech, VideoMaker y muchas otras publicaciones. A día de hoy más de cuatro millones de personas y empresas usan este software para el trabajo relacionado con la edición de vídeo. Sus principales características son las siguientes:

- Convertir archivos de vídeo y DVD en todos los formatos y para cualquier dispositivo.
- Extraer los DVD sin protección de su colección con un único clic.
- Ver vídeos y DVD en cualquier dispositivo móvil: iPod, iPhone, PSP, Zune, BlackBerry, Creative Zen, Archos, teléfonos y muchos otros.
- Ajustes de formato de salida totalmente configurables para usuarios avanzados.
- Prestaciones de edición únicas: dividir, fusionar, recortar, rotar, ajustar y añadir filigranas, así como galería fotográfica.
- Tecnología de conversión de vídeo muy rápida con funciones multihilo y conversión por lotes.

Ilustración 35: Movavi Video Converter

Este programa se eligió para la codificación de los vídeos por su interfaz sencilla, por la multitud de opciones de diseño y por realizar la codificación de vídeos de forma simple y acorde con la norma ISO/IEC 14496 Part 2: MPEG-4 Visual [30].

La decodificación parcial de los vídeos y el desarrollo de los algoritmos se ha realizado con ***MATLAB® - The Language of Technical Computing*.** MATLAB es la abreviatura de MATrix LABoratory. Fue creado por The MathWorks en 1984, al ver la necesidad en ingeniería de un entorno de programación más potente y eficaz que lenguajes como Fortran y C. En respuesta a esa necesidad, los creadores de Mathworks, combinaron su experiencia en matemáticas, ingeniería e informática para desarrollar MATLAB. MATLAB combina las matemáticas y funciones gráficas con un lenguaje de programación de alto nivel muy potente. Las aplicaciones principales de MATLAB incluyen: el desarrollo de algoritmos, visualización y análisis de datos, cálculo numérico, procesamiento de señales e imágenes y modelos financieros. A día de hoy más de un millón de usuarios lo utilizan en 100 países, repartidos entre empresas, laboratorios de investigación, instituciones financieras y más de 3500 universidades [29].

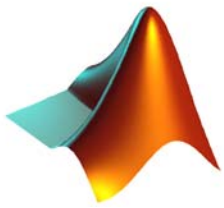

Ilustración 36: Logo de Matlab

Antes de presentar las gráficas obtenidas para cada algoritmo y vídeo, resumimos en las tablas 5, 6 y 7 las características principales y los cambios de plano de los vídeos:

Tabla 5: Características vídeo 1

Formato imagen	320x240píxeles										
Duración	19.96s										
Número cuadros	499										
Codificación	572kbps										
Cambios	45	81	121	144	173	190	228	250	272	299	451

Tabla 6: Características vídeo 2

Formato imagen	320x240píxeles						
Duración	20						
Número cuadros	500						
Codificación	572kbps						
Cambios	110	236	281	330	366	416	485

Tabla 7: Característica vídeo 3

Formato imagen	320x240píxeles																	
Duración	30																	
Número cuadros	750																	
Codificación	572kbps																	
Cambios	23	42	69	114	191	262	291	332	361	393	420	436	484	532	551	583	604	627

4.1. Detección mediante diferencias de histogramas de color

En este algoritmo se utilizan las diferencias entre histogramas de color de las imágenes DC. Se calculan estas distancias por separado para utilizarlas de forma conjunta en la detección. La condición para marcar un plano como nuevo es que la distancia para él sea la máxima en una ventana de 15 cuadros y mayor que el triple del valor medio de las distancias de las ventanas. La ilustración 37 muestra de forma gráfica las condiciones para la detección de un cambio de plano. Las ilustraciones 38, 39, 40 muestran los resultados obtenidos con este algoritmo.

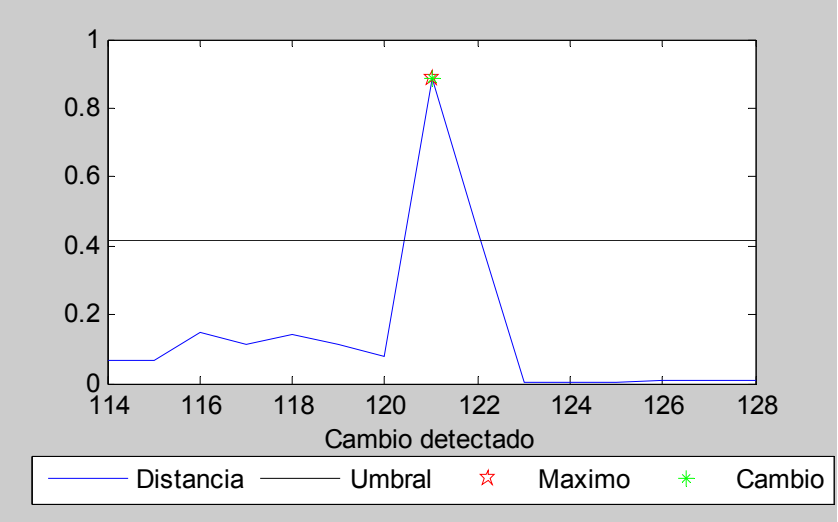

Ilustración 37: Condiciones de detección

4.1.1. Vídeo 1

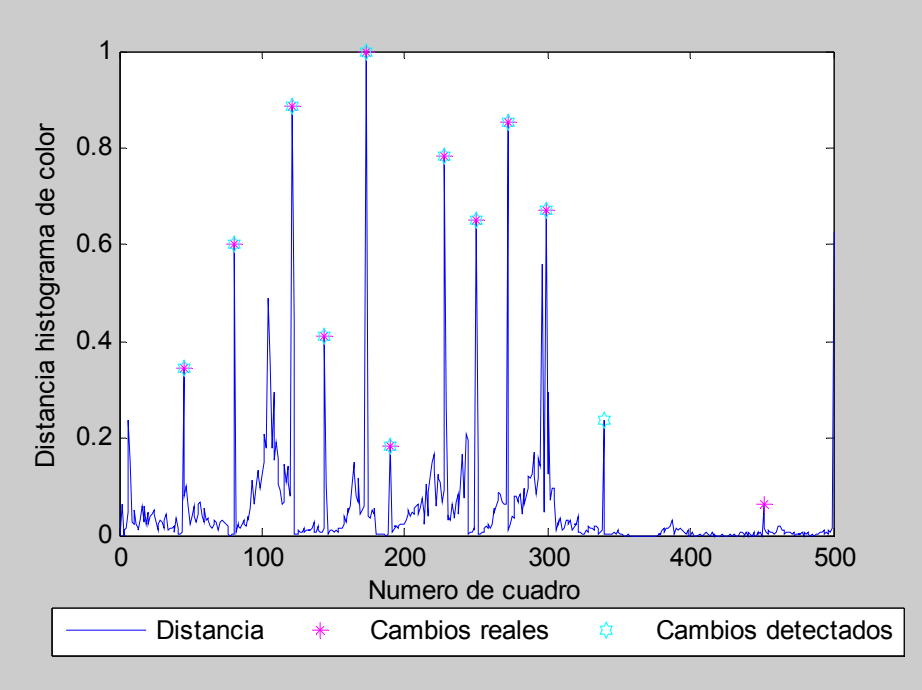

Ilustración 38: Histogramas de color en vídeo 1

4.1.2. Vídeo 2

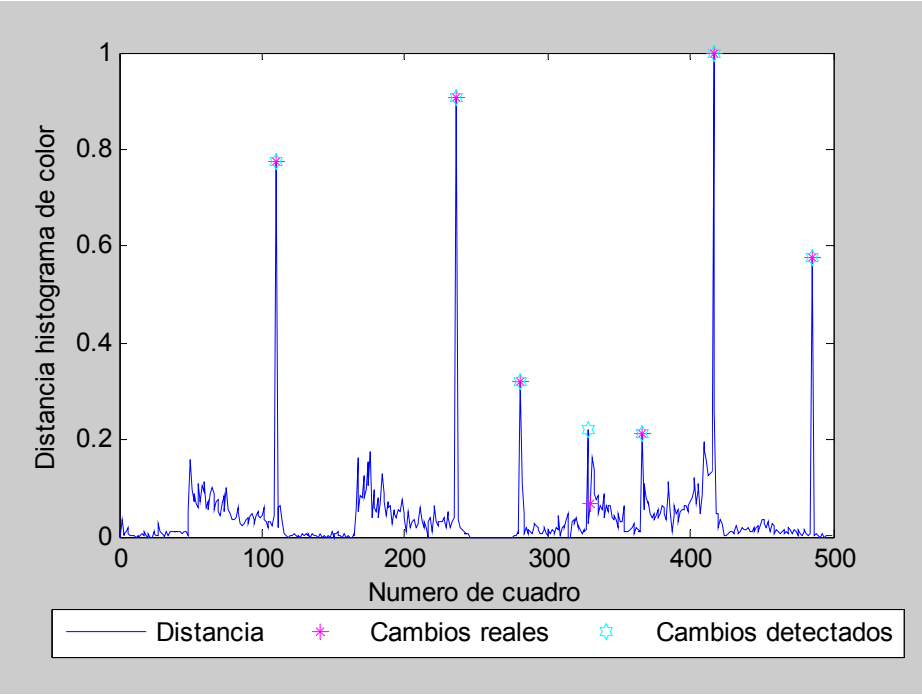

Ilustración 39: Detección con histogramas de color en vídeo 2

4.1.3. Vídeo 3

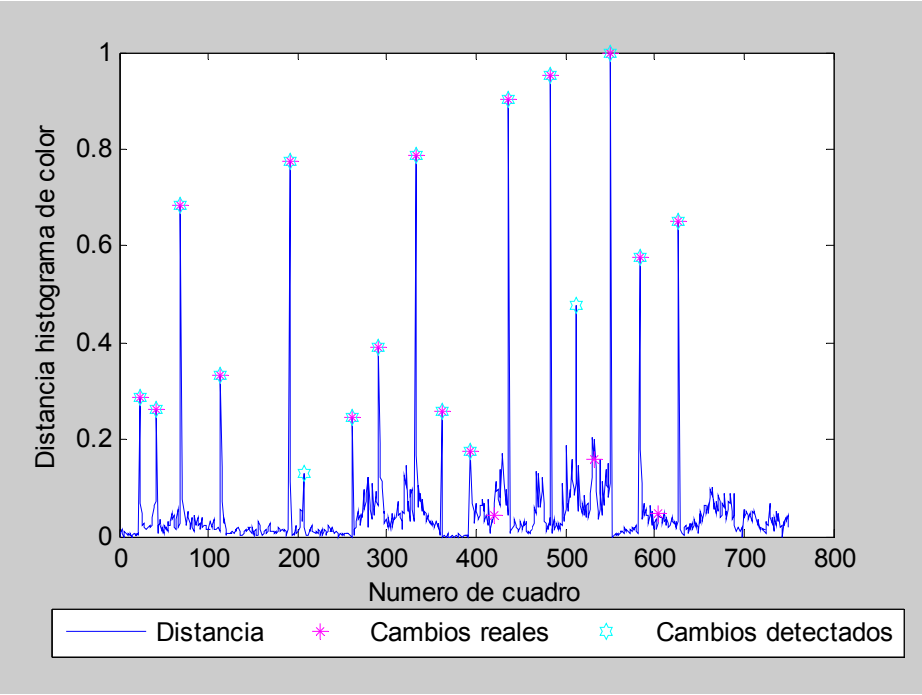

Ilustración 40: Detección con histogramas de color en vídeo 3

4.2. Detección con ventana deslizante de imágenes DC

En este método se utiliza las diferencias de las imágenes DC de cuadros consecutivos. Para la detección se utiliza una ventana deslizante en la que, para declarar un cambio de plano en un punto la distancia en él debe cumplir las siguientes condiciones:

- Ser el máximo en la ventana.
- Superar el doble del valor medio de la distancia en la ventana.
- Superar en un 20% el valor el segundo máximo de la ventana.

La ilustración 41 muestra estas condiciones de forma gráfica:

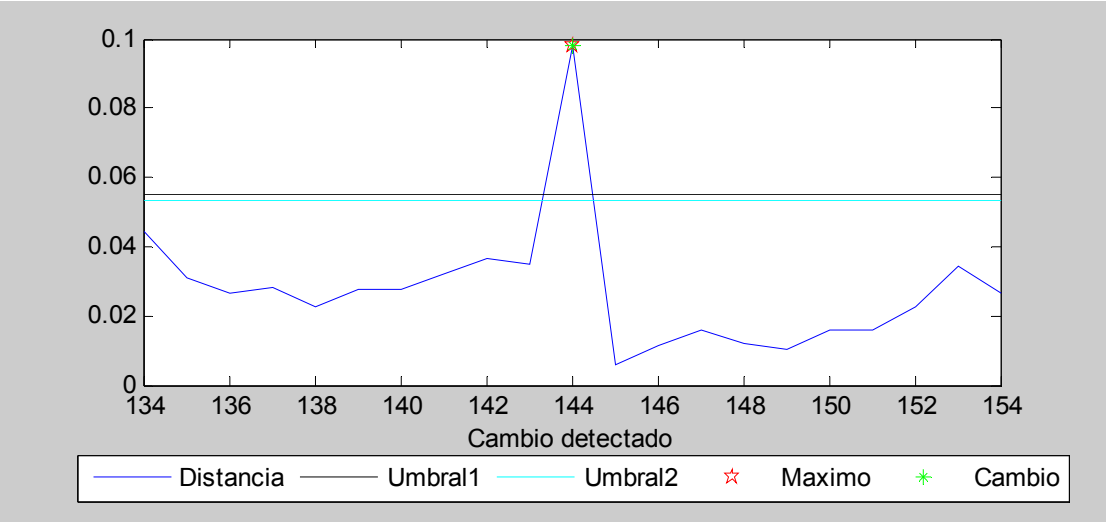

Ilustración 41: Condiciones de detección

Los resultados de este método están representados en las ilustraciones 42, 43 y 44:

4.2.1. Vídeo 1

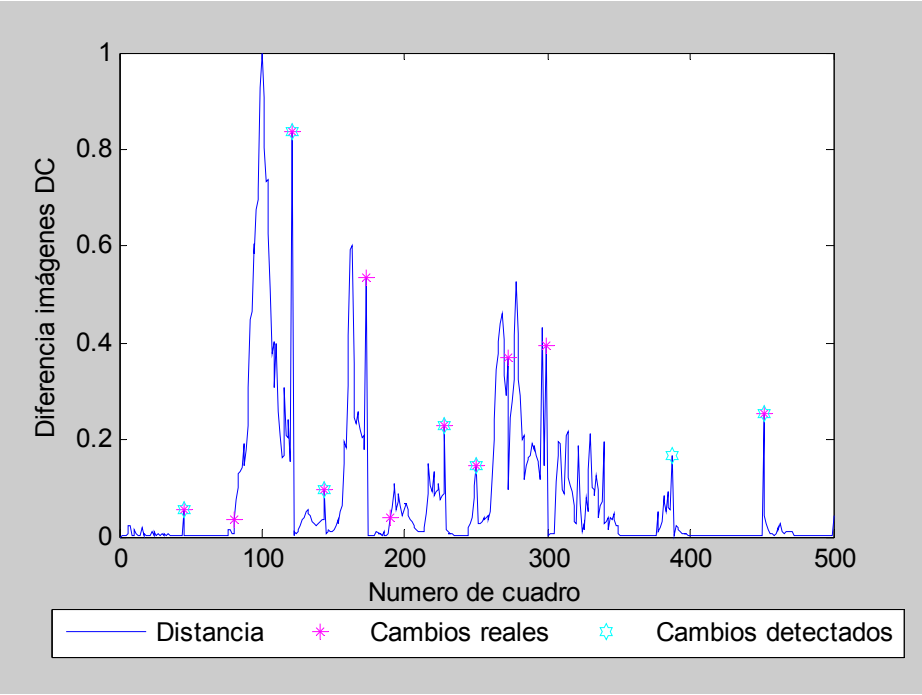

Ilustración 42: Detección con imágenes DC en vídeo 1

4.2.2. Vídeo 2

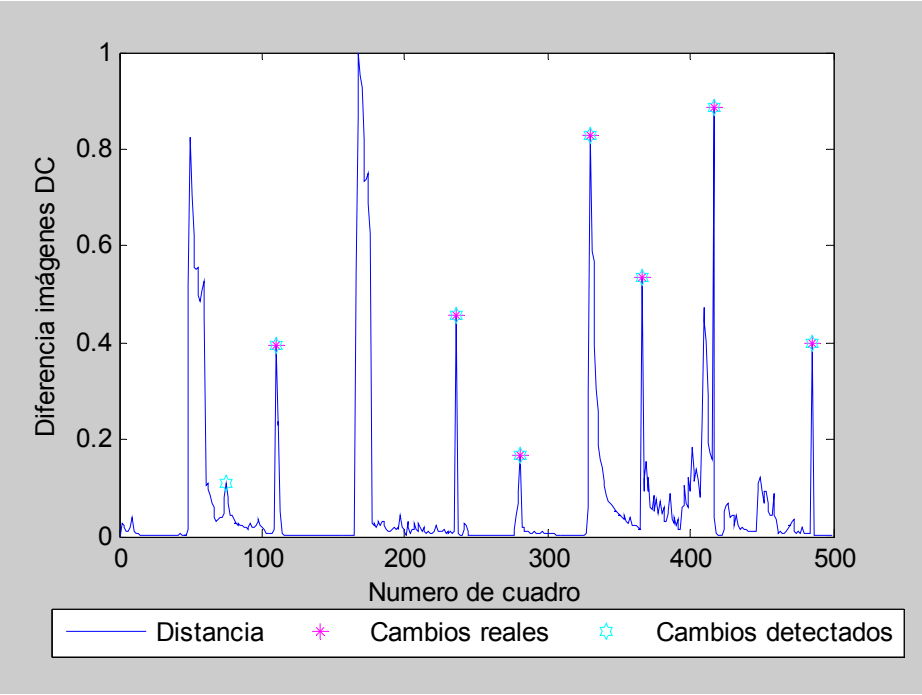

Ilustración 43: Detección con imágenes DC en vídeo 2

4.2.3. Vídeo 3

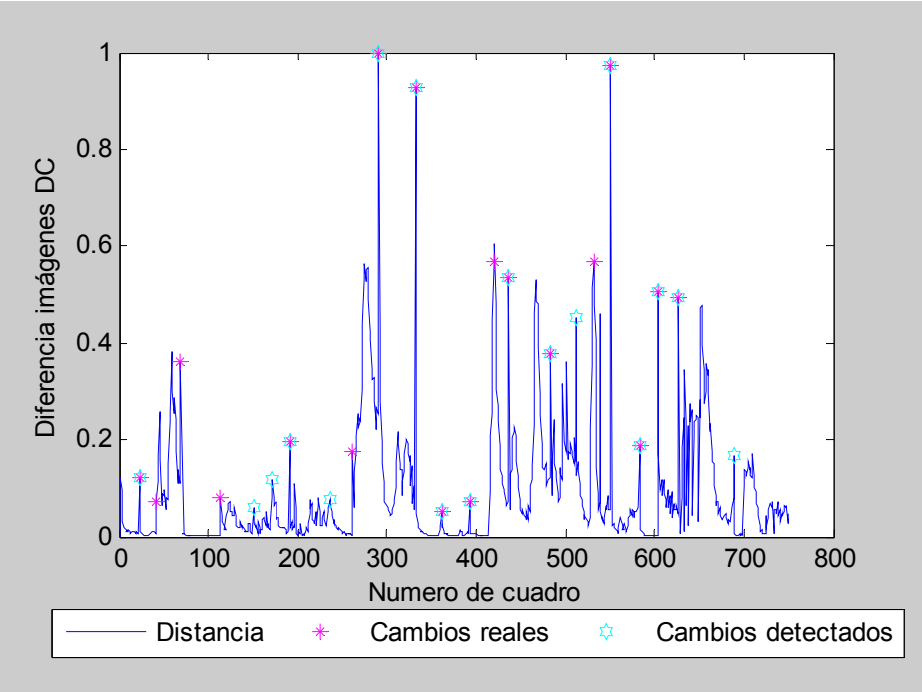

Ilustración 44: Detección con imágenes DC en vídeo 3

4.3. Detección con el método de amplificación de ventana doble

Al aumentar este método las diferencias entre cuadros consecutivos, el uso de un umbral numérico es suficiente para obtener unos buenos resultados en la detección. Las ilustraciones 45, 46 y 47 muestran el funcionamiento de esta técnica.

4.3.1. Vídeo 1

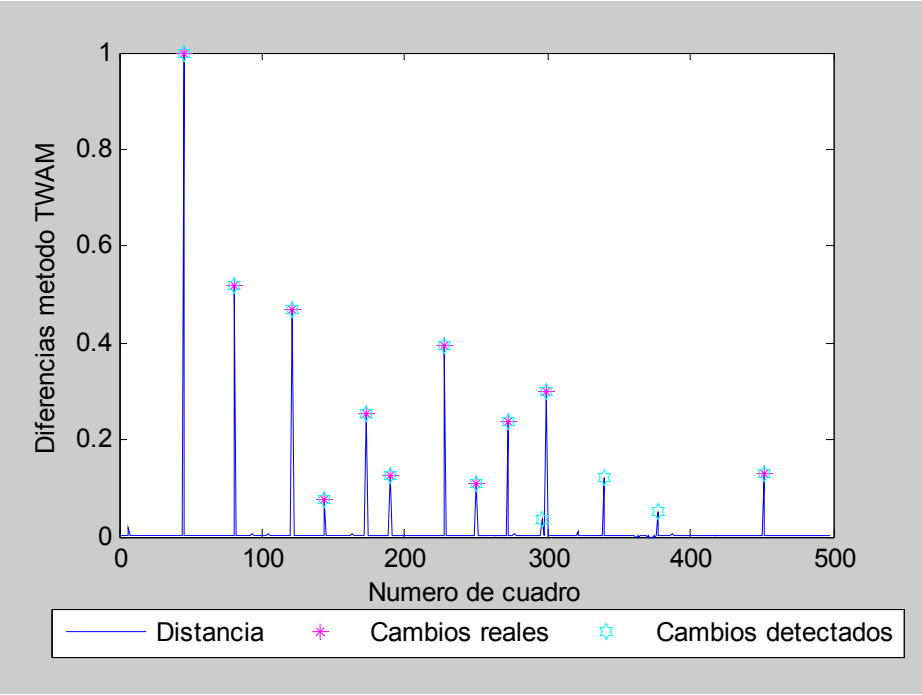

Ilustración 45: Detección con método TWAM en vídeo 1

4.3.2. Vídeo 2

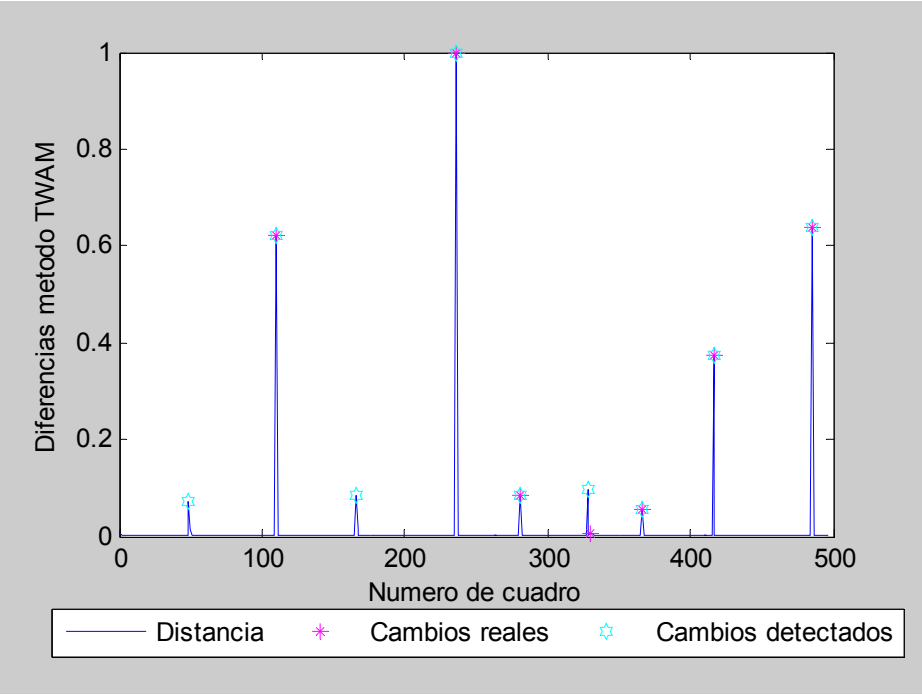

Ilustración 46: Detección con método TWAM en vídeo 2

4.3.3. Vídeo 3

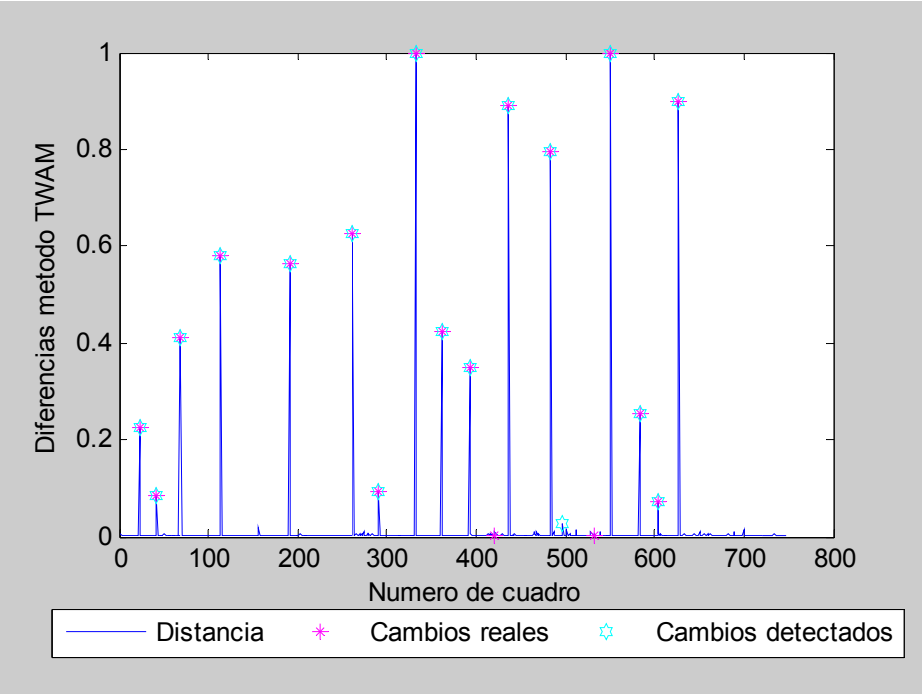

Ilustración 47: Detección por método TWAM en vídeo 3

4.4. Detección por tipos de VOP

Este algoritmo divide el proceso de detección en dos partes, según el tipo de VOP. Si el cuadro actual es de tipo I, se evalúa la distancia respecto al anterior VOP I, mientras que si es P se evalúa la proporción entre macrobloques tipo *intra* y tipo *inter*. Las condiciones que deben cumplirse son las siguientes:

- Cuadro tipo I:
  - la distancia debe ser la máxima en una ventana de 21 cuadros.
- Cuadro tipo P:
  - Debe ser el máximo en una ventana de 21 cuadros.
  - El valor debe superar en ocho veces el valor medio de dicha ventana.

La ilustración 48 muestra la detección de un cambio de plano en un cuadro I y en un cuadro P, al haberse cumplido las condiciones de detección.

Los resultados obtenidos con este método para los diferentes vídeos pueden verse en las ilustraciones 49, 50 y 51.

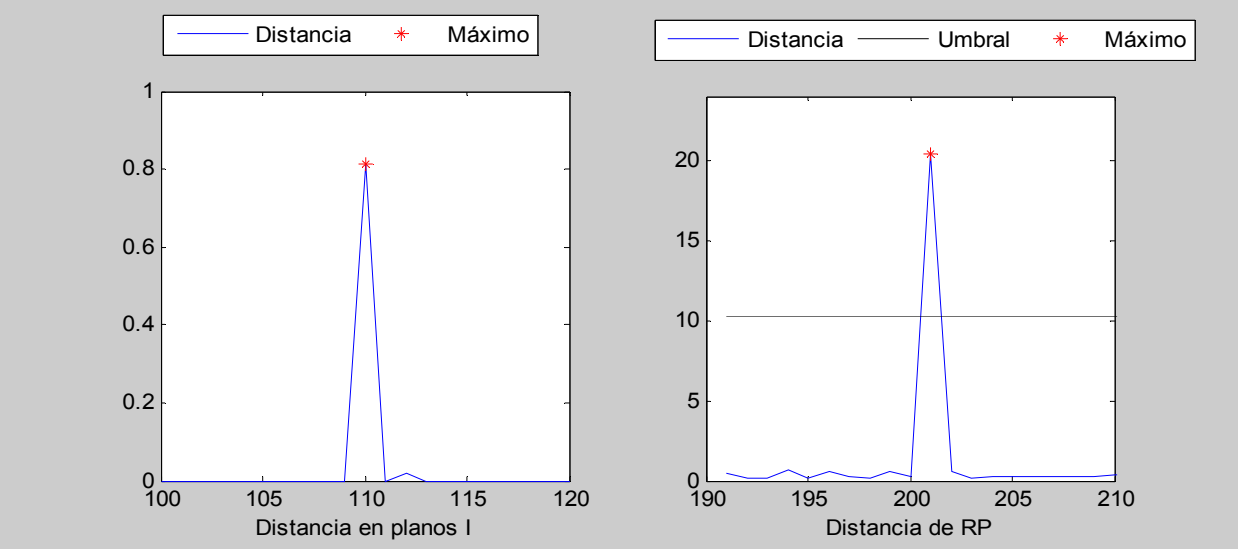

Ilustración 48: Detección por tipos de VOP

4.4.1. Vídeo 1

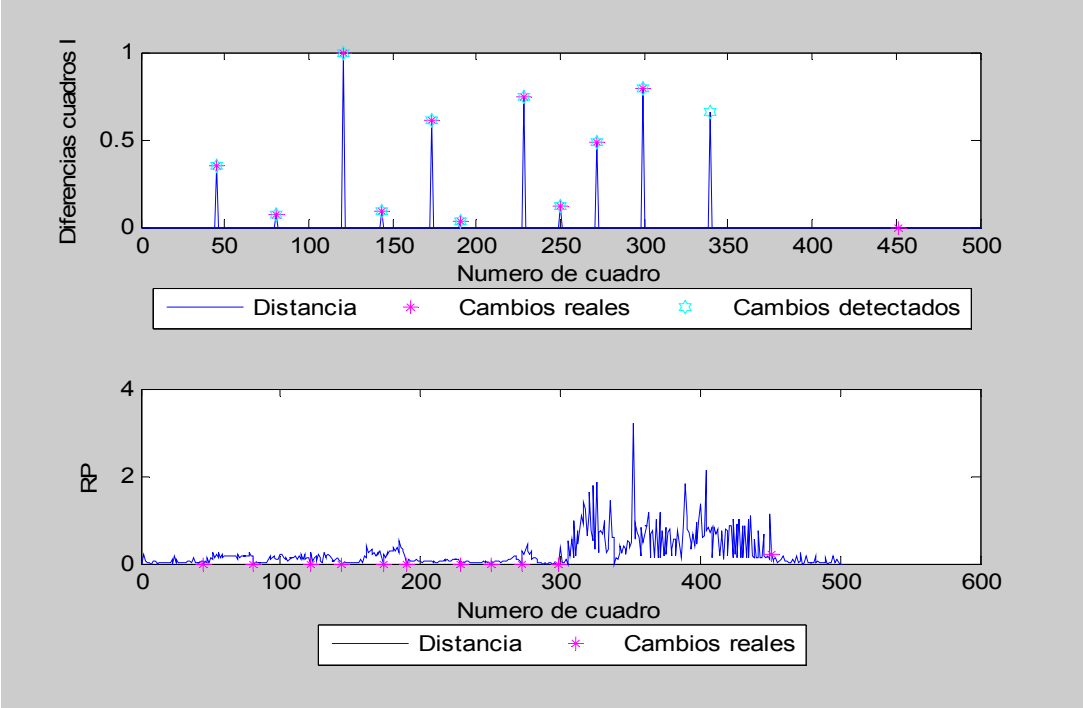

Ilustración 49: Detección por tipos de VOP en vídeo 1

4.4.2. Vídeo 2

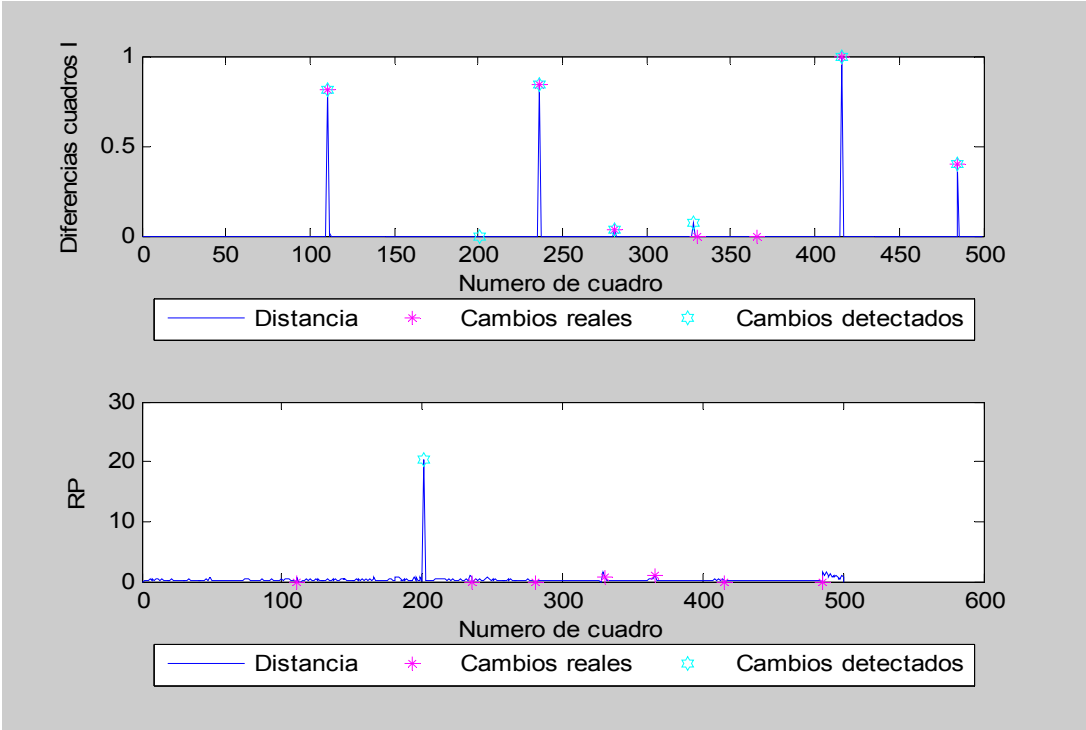

Ilustración 50: Detección por tipos de VOP en vídeo 2

4.4.3. Vídeo 3

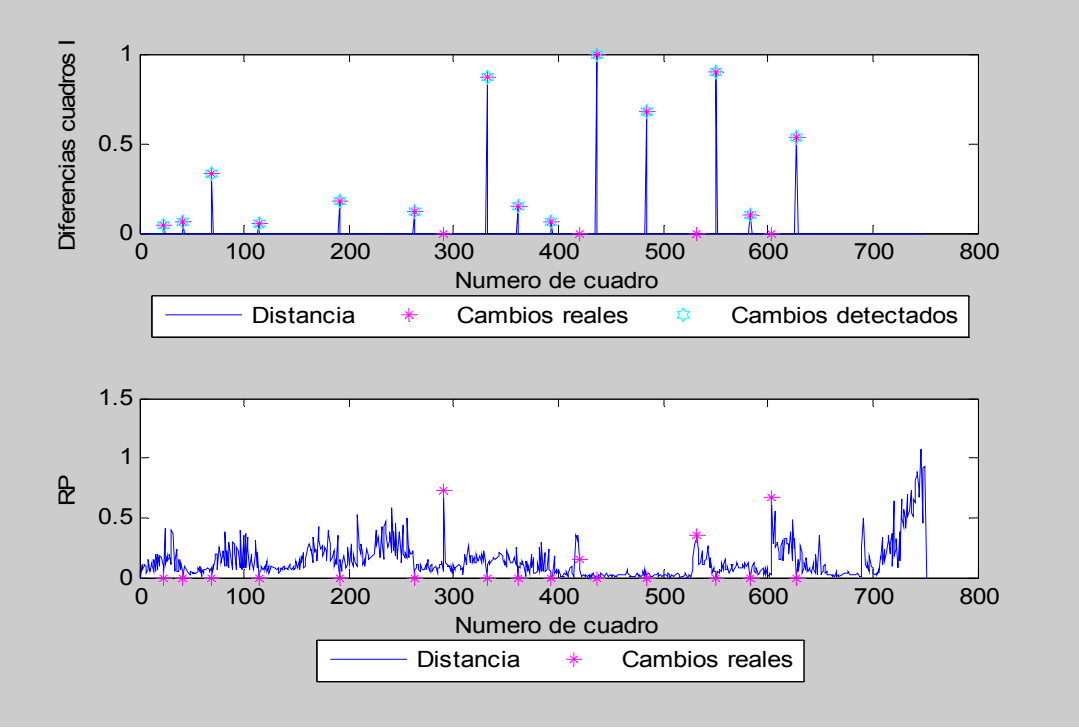

Ilustración 51: Detección por tipos de VOP en vídeo 3

4.5. Detección con la característica generalizada de la secuencia

Este método hace una combinación de las dos formas más comunes de detección de cambios de plano: la evaluación de la diferencia entre cuadros a nivel de píxel, usando la desviación típica, y de forma estadística mediante el uso de histogramas. Las condiciones para declarar un cambio de plano pueden darse de dos maneras:

1) La distancia es mayor que 0.3 y es el máximo en una ventana de 21 cuadros (en la ilustración 52 la figura de la izquierda).
2) La distancia es mayor que 0.1, es el máximo en la ventana de 21 cuadros y supera 3 veces el valor de la distancia media en la ventana (en la ilustración 52 la figura de la derecha).

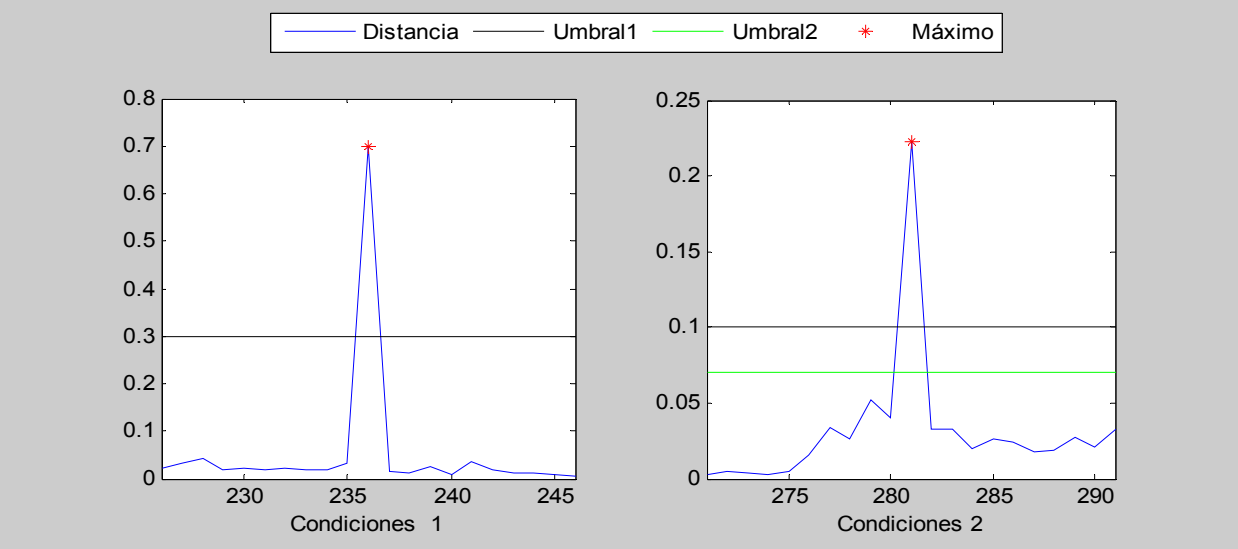

Ilustración 52: Condiciones para la detección

En las ilustraciones 53, 54 y 55 pueden verse los resultados de esta técnica para los vídeos 1, 2 y 3 respectivamente:

4.5.1. Vídeo 1

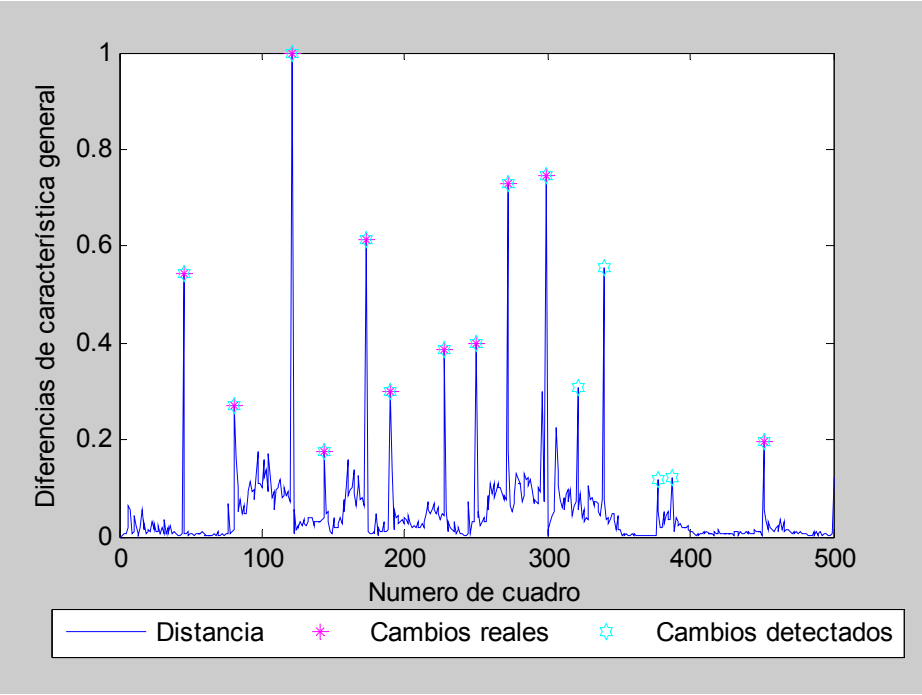

Ilustración 53: Detección por característica en vídeo 1

4.5.2. Vídeo 2

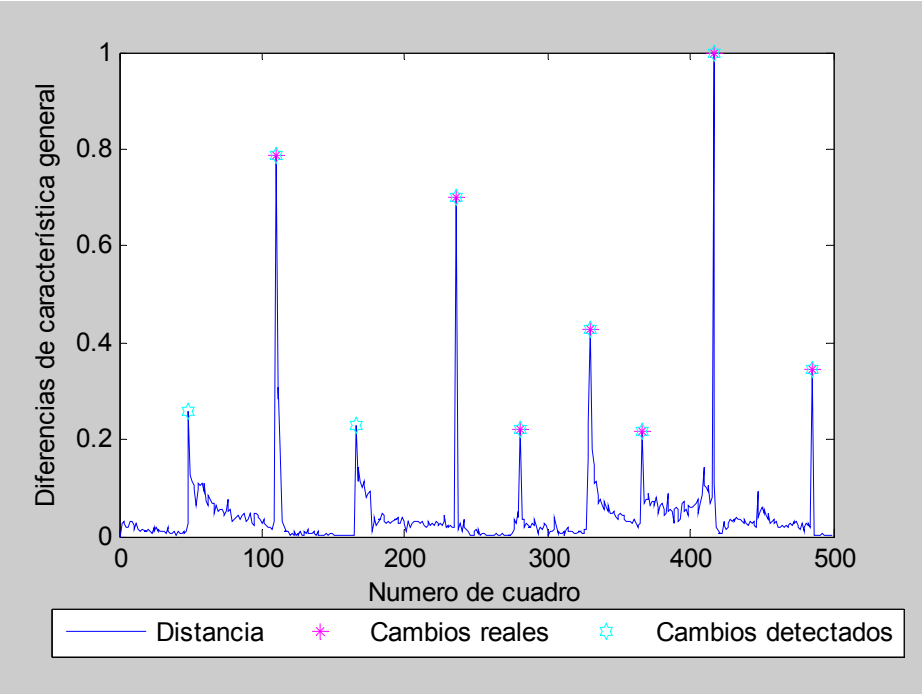

Ilustración 54: Detección por característica en vídeo 2

4.5.3. Vídeo 3

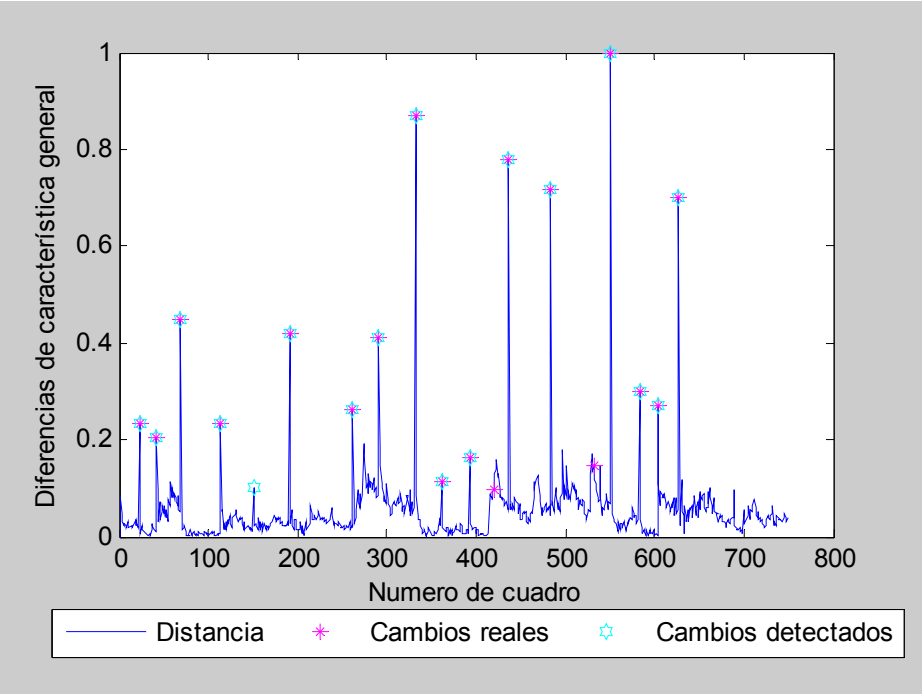

Ilustración 55: Detección mediante característica en vídeo 3

4.6. Detección LBDF

Este algoritmo utiliza la correlación entre histogramas para detectar los posibles cambios de plano. De forma análoga a los anteriores, se decide que ha habido un cambio de plano cuando se cumplen estas condiciones:

- La distancia es la máxima en una ventana de 15 cuadros.
- La distancia supera en dos veces el valor medio de las distancias de la ventana

La ilustración 56 muestra la detección de un plano al haberse cumplido las condiciones.

Los resultados de la técnica LBDF pueden verse en las ilustraciones 57, 58 y 59.

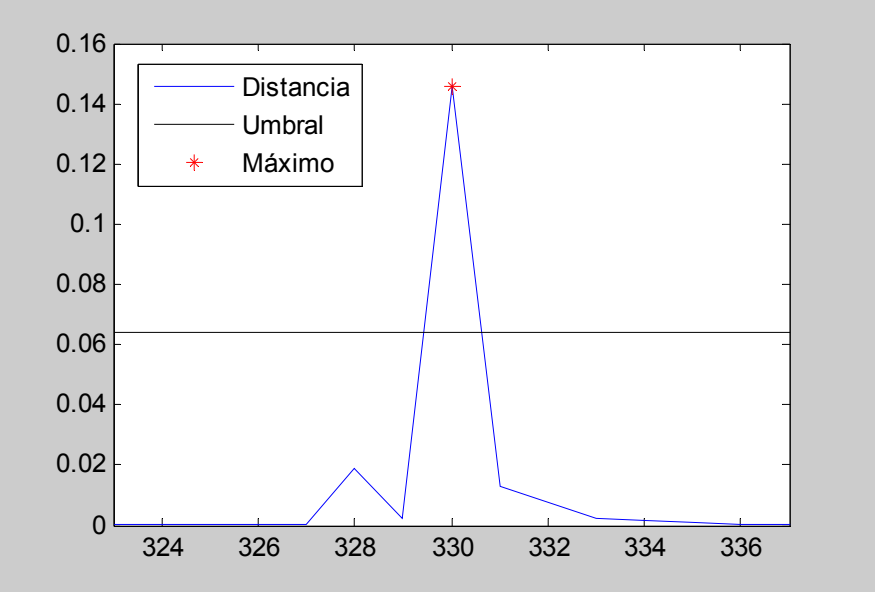

Ilustración 56: Condiciones para la detección

4.6.1. Vídeo 1

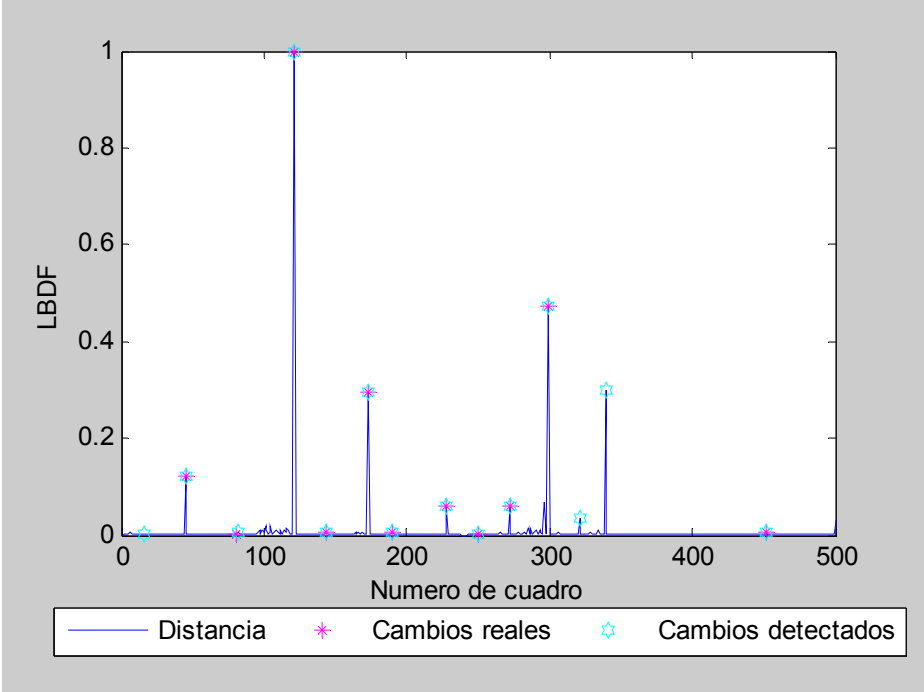

Ilustración 57: Detección LBDF en vídeo 1

4.6.2. Vídeo 2

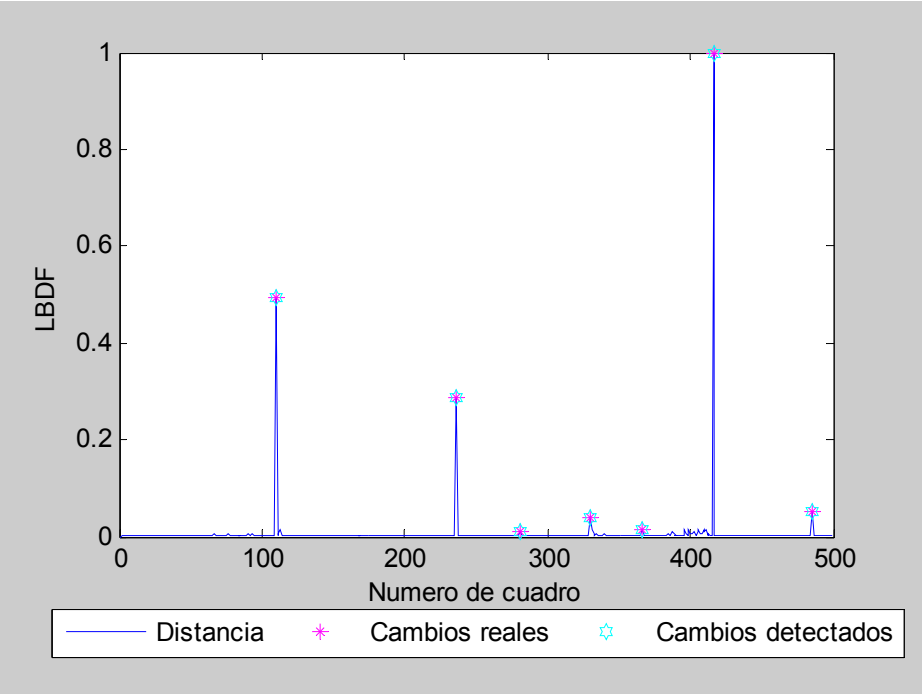

Ilustración 58: Detección LBDF para vídeo 2

4.6.3. Vídeo 3

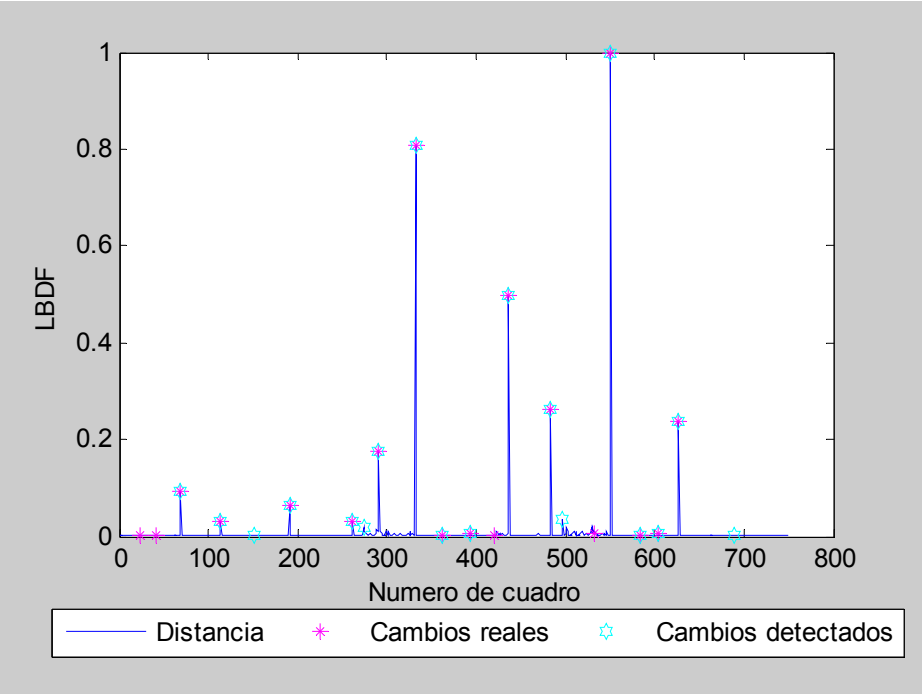

Ilustración 59: Detección LBDF para vídeo 3

4.7. Detección mediante análisis χ^2 de los histogramas de color

Este método aplica, en una ventana de diferencias χ^2 de los histogramas de color, un filtro temporal para declarar un cambio de plano. La condición que debe darse es que el cociente entre la diferencia entre el cuadro actual y los pasados sea mucho mayor que las diferencias entre el resto de cuadros vecinos dentro de la ventana. La ilustración 60 muestra esos cocientes en comparación con la distancia χ^2 y los umbrales que se deben cumplir:

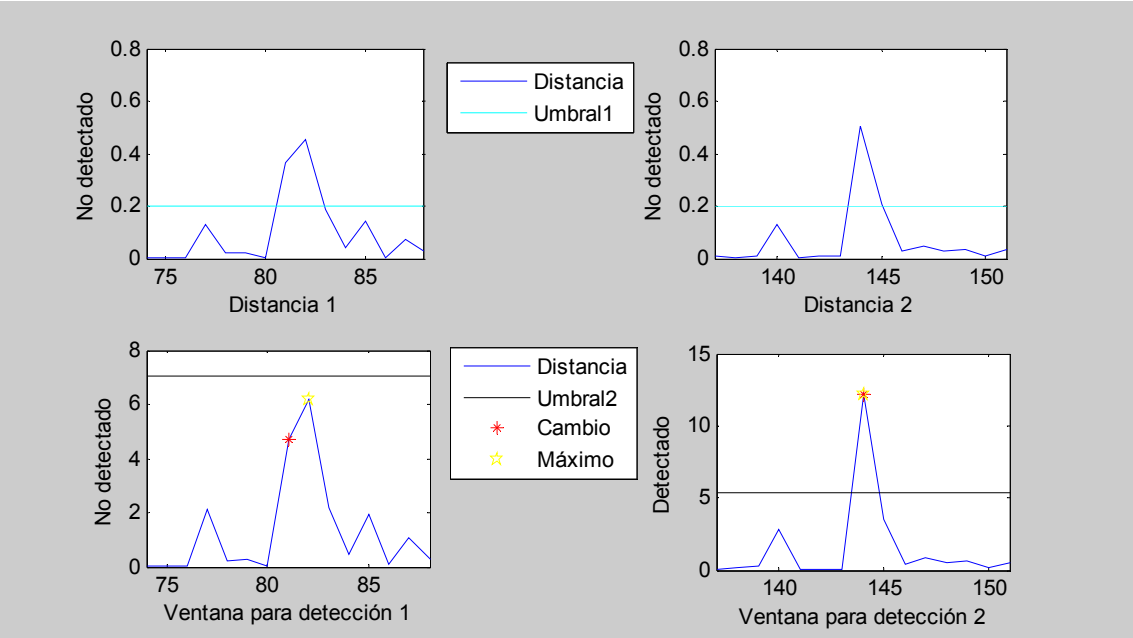

Ilustración 60: Detección en el método χ^2

En las ilustraciones 61, 62 y 63 se ven los resultados del método χ^2:

4.7.1. Vídeo 1

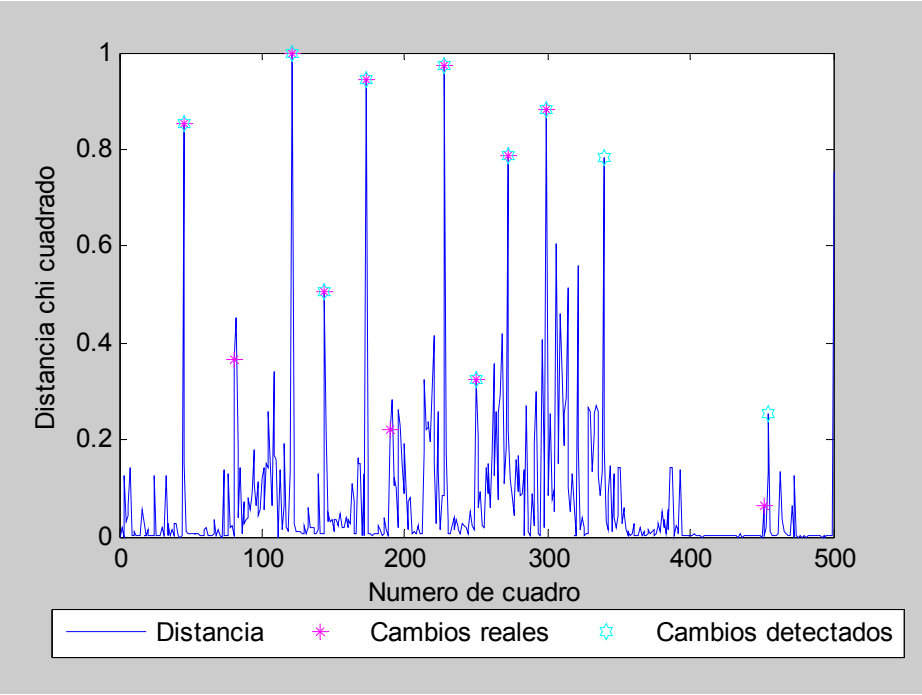

Ilustración 61: Distancia chi cuadrado en vídeo 1

4.7.2. Vídeo 2

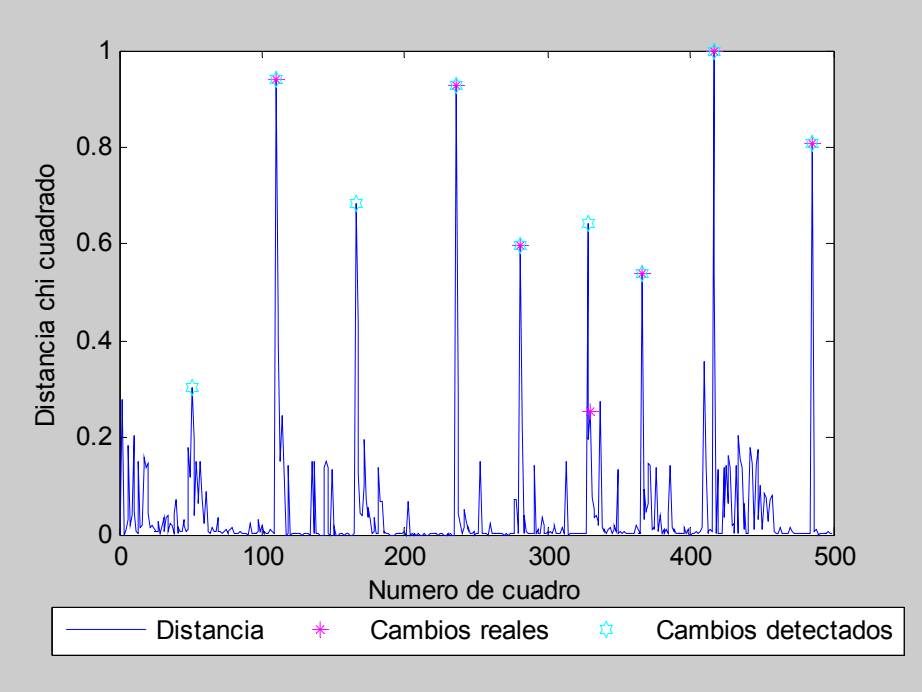

Ilustración 62: Distancia chi cuadrado para vídeo 2

4.7.3. Vídeo 3

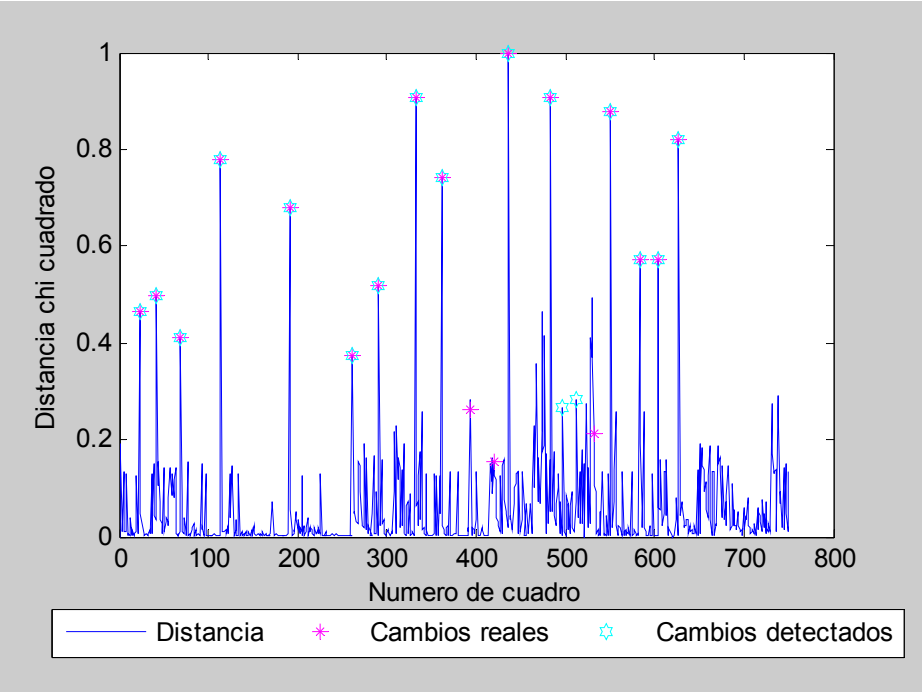

Ilustración 63: Distancia chi cuadrado para vídeo 3

4.8. Evaluación de los resultados

La bondad de los resultados en la detección de cambios de plano se mide con dos parámetros: *precisión* y *recall*. Estos parámetros se definen de la siguiente manera:

$$precisión = \frac{N_C}{N_C + N_{FA}} \qquad recall = \frac{N_C}{N_C + N_{MD}}$$

Donde N_C es el número de detecciones correctas, N_{FA} es el número de falsas alarmas y N_{MD} es el número de cambios no detectados. Las tablas 6, 7, 8, 9, 10, 11 y 12 muestran los resultados para cada uno de los algoritmos implementados:

Por último, la tabla 13 hace una comparación entre las diferentes técnicas.

Tabla 6: Detección por histogramas de color

Vídeo	Precisión	Recall
01_01	90,91	90,91
01_02	90,00	81,82
01_03	100,00	100,00
02_01	85,71	85,71
02_02	100,00	100,00
02_03	75,00	85,71
03_01	88,24	83,33
03_02	93,33	77,78
03_03	82,35	77,78
Total	**89,32**	**85,19**

Tabla 7: Detección con imágenes DC

Vídeo	Precisión	Recall
01_01	85,71	54,55
01_02	37,50	27,27
01_03	75,00	54,55
02_01	87,50	100,00
02_02	87,50	100,00
02_03	85,71	85,71
03_01	70,59	66,67
03_02	100,00	61,11
03_03	68,75	61,11
Total	**76,67**	**63,89**

Tabla 8: Detección con método TWAM

Vídeo	Precisión	Recall
01_01	78,57	100,00
01_02	83,33	90,91
01_03	100,00	100,00
02_01	66,67	85,71
02_02	66,67	85,71
02_03	75,00	85,71
03_01	94,12	88,89
03_02	100,00	88,89
03_03	88,89	88,89
Total	**85,96**	**90,74**

Tabla 9: Detección por tipos de VOP

Vídeo	Precisión	Recall
01_01	90,91	90,91
01_02	83,33	90,91
01_03	91,67	100,00
02_01	71,43	71,43
02_02	75,00	85,71
02_03	85,71	85,71
03_01	100,00	77,78
03_02	100,00	83,33
03_03	100,00	88,89
Total	**91,18**	**86,11**

Tabla 10: Detección por característica de secuencia

Vídeo	Precisión	Recall
01_01	73,33	100,00
01_02	75,00	81,82
01_03	84,62	100,00
02_01	77,78	100,00
02_02	100,00	100,00
02_03	77,78	100,00
03_01	94,12	88,89
03_02	100,00	83,33
03_03	88,89	88,89
Total	**86,09**	**91,67**

Tabla 11: Detección con LBDF

Vídeo	Precisión	Recall
01_01	71,43	90,91
01_02	75,00	81,82
01_03	66,67	90,91
02_01	100,00	100,00
02_02	100,00	100,00
02_03	100,00	100,00
03_01	77,78	77,78
03_02	100,00	77,78
03_03	82,35	77,78
Total	**82,88**	**85,19**

Tabla 12: Detección análisis X^2

Vídeo	Precisión	Recall
01_01	80,00	72,73
01_02	87,50	63,64
01_03	54,55	54,55
02_01	66,67	85,71
02_02	85,71	85,71
02_03	54,55	85,71
03_01	88,24	83,33
03_02	88,24	83,33
03_03	88,24	83,33
Total	**78,50**	**77,78**

Tabla 13: Comparación de los resultados

Método	Precisión	Recall
Histogramas color	89,32	85,19
Imágenes DC	76,67	63,89
TWAM	85,96	90,74
Tipos VOP	91,18	86,11
Característica	86,09	91,67
LBDF	82,88	85,19
Análisis X^2	78,50	77,78

En las ilustraciones 63 y 64 pueden verse las comparaciones entre los resultados de los distintos métodos en cuanto a *precisión* y *recall* respectivamente. En la ilustración 65 se representan los dos parámetros de forma conjunta para todas las técnicas.

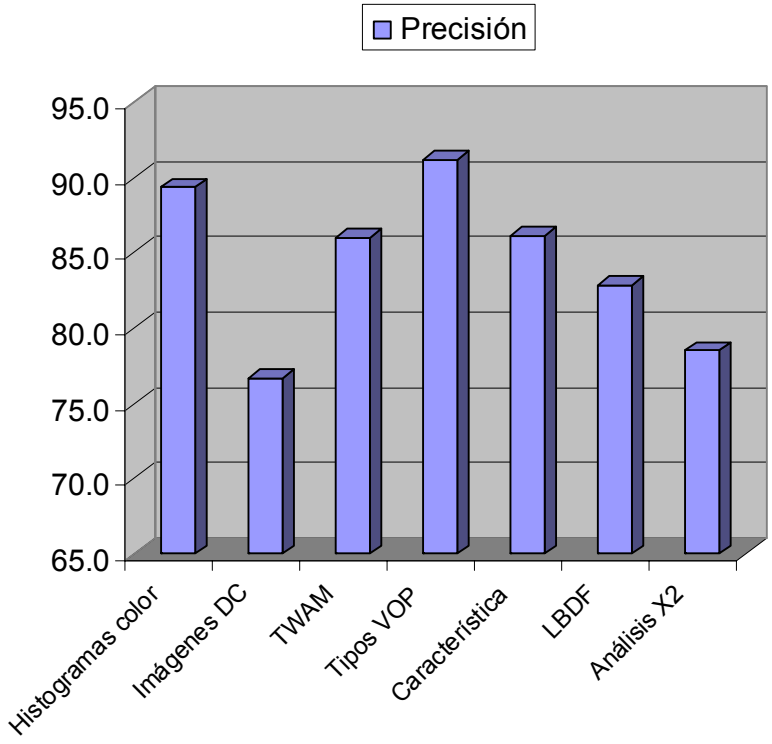

Ilustración 64: Comparación de la *precisión* de las diferentes técnicas

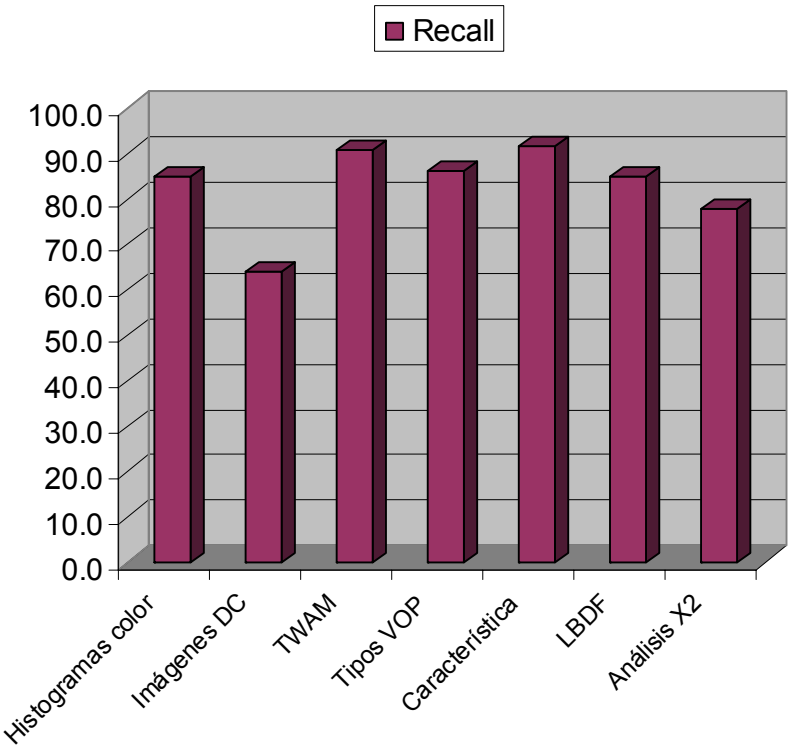

Ilustración 65: Comparación del *recall* de las diferentes técnicas

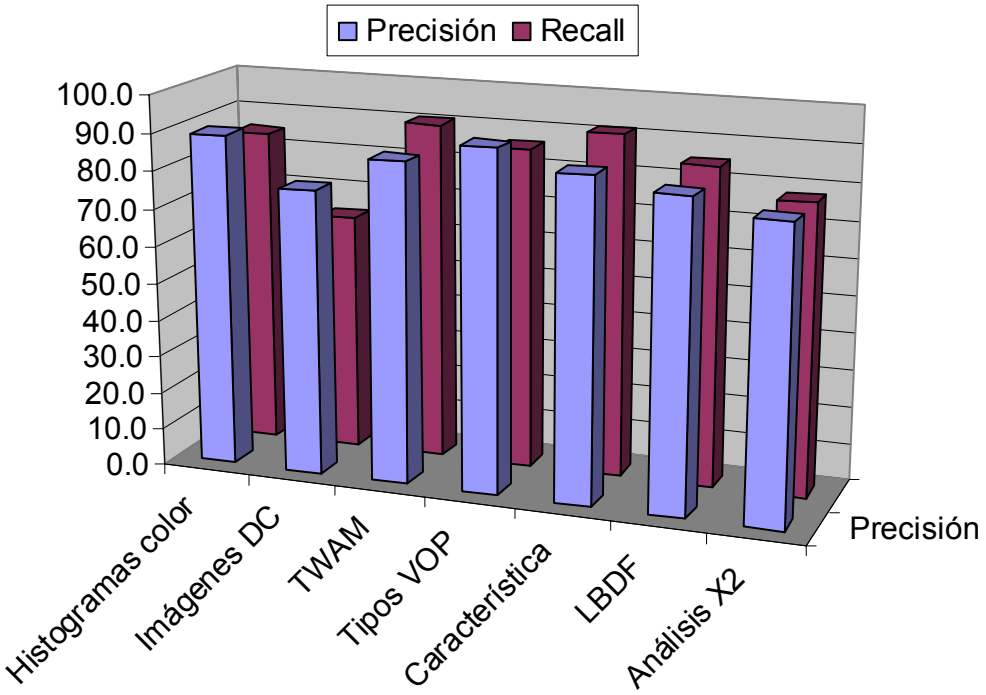

Ilustración 66: Comparación global de los resultados de las diferentes técnicas

De estos métodos, el que ha dado mejores resultados es el que realiza la detección distinguiendo por tipos de VOP, con una *precisión* del 91,18% y un *recall* de 86,11%. Es fácil ver el porqué de su buen funcionamiento. El codificador, al detectar un cambio grande de un plano a otro, prefiere codificar el cuadro como I en lugar de enviar la compleja información diferencial obtenida al compararla con un cuadro de referencia. Así, es muy probable que los nuevos planos sean de tipo I. De forma análoga, en un cuadro tipo P, habrá una gran cantidad de macrobloques tipo *intra* en un nuevo plano, lo que resulta también fácil de detectar.

La técnica que detecta los nuevos planos por los histogramas de color ha tenido también unos buenos resultados. Esta técnica aprovecha que, al haber un cambio de plano, es difícil que la información de color sea la misma en ambas imágenes. Evidentemente, esta técnica falla cuando, aún siendo distintos, los planos se parecen mucho en los colores o cuando cambia el color de un plano como, por ejemplo, por un movimiento de cámara.

El método TWAM ha dado también unos buenos resultados. El uso de una ventana doble para comparar distancias anteriores y posteriores funciona bien, ya que puede haber cambios de plano que generen distancias absolutas insignificantes y, sin embargo, ser significativamente mayores al resto de las distancias de su ventana.

Las medidas de las diferencias de forma estadística, como hacen el análisis X^2 y el LBDF, que parte de la correlación entre planos, funcionan algo peor que los anteriores. El problema del análisis X^2 es que la distancia fluctúa mucho entre planos, lo que dificulta la detección. Con la medida de la correlación ocurre algo parecido las diferencias en nuevos planos son tan dispares que elegir un umbral adecuado resulta difícil.

De todas las técnicas estudiadas, la que utiliza directamente las diferencias de imágenes DC es la que más falla. Al restar imágenes DC, un pequeño movimiento de cámara o movimiento de los objetos hará que la diferencia entre dos planos crezca, lo que se traducirá en una falsa alarma. Además, también habrá cambios que no se detecten cuando, por ejemplo, dos planos tengan el mismo fondo pero distinto contenido.

El último método combina dos métricas. La primera, la diferencia de histogramas de luminancia, que no funciona tan bien como la de histogramas de color. La segunda, la desviación típica, que mide la variación de intensidad de los bloques de la imagen. La combinación de ambas hace que se complementen de cierta manera, consiguiendo que el número de cambios no detectados sea pequeño y, por tanto, se consiga un *recall* alto. Sin embargo, el uso de ambas características hace que el número de falsas alarmas sea mayor, haciendo así la *precisión* más pequeña.

4.9. Pruebas de robustez

Una característica deseable en las técnicas de detección de cambios de plano es su robustez frente a cambios en la calidad o las características de los vídeos. Por este motivo resulta interesante comprobar cómo se comportan los algoritmos frente a algunos de estos cambios. A continuación se presentan las dos pruebas de robustez realizadas y los resultados obtenidos.

4.9.1. Cambios frente al ruido

Para probar cómo funcionan los algoritmos frente a variaciones en las imágenes se ha introducido ruido de varias potencias en las secuencias de imágenes DC de los vídeos.

Tabla 14: Precisión (%) frente al ruido

SNR (dB)	5,0	10,0	15,0	20,0	25,0	30,0	35,0	40,0	45,0	50,0	55,0	60,0
Histograma color	89,8	89,4	88,3	87,9	87,3	86,0	85,7	85,5	85,1	85,5	85,2	85,4
Imágenes DC	77,7	77,9	76,4	76,4	76,4	76,4	76,4	76,7	76,7	76,7	76,7	76,7
TWAM	85,2	84,5	83,7	83,4	83,6	83,9	84,5	85,0	85,5	85,7	85,8	85,8
Tipos VOP	52,7	63,2	73,4	81,3	86,6	90,4	91,7	92,3	92,3	92,6	92,4	92,4
Característica	88,6	87,7	86,5	86,1	85,7	85,4	85,0	85,2	85,0	85,0	85,0	85,0
LBDF	84,6	84,5	83,3	83,4	83,0	83,5	83,4	83,1	83,1	83,1	83,1	82,9
X^2	80,4	79,6	80,1	80,7	80,7	80,5	80,0	79,5	79,6	79,9	78,8	78,9

Tabla 15: Recall (%) frente al ruido

SNR (dB)	5,0	10,0	15,0	20,0	25,0	30,0	35,0	40,0	45,0	50,0	55,0	60,0
Histograma color	81,2	83,4	86,0	85,7	85,3	84,5	84,2	83,9	83,6	83,8	83,6	83,8
Imágenes DC	64,6	63,7	63,0	63,0	63,0	63,0	63,0	63,9	63,9	63,9	63,9	63,9
TWAM	89,7	90,1	90,6	90,8	90,4	90,2	90,2	90,0	90,2	89,8	90,1	89,8
Tipos VOP	82,0	82,8	83,2	83,1	83,5	83,7	83,3	83,7	83,8	84,0	83,9	83,9
Característica	90,2	90,4	89,7	89,7	89,4	89,1	88,9	89,3	88,9	88,9	89,0	89,0
LBDF	86,0	86,1	85,7	85,8	85,7	85,4	85,3	85,4	85,4	85,6	85,6	85,4
X^2	77,4	78,0	79,0	79,2	79,1	79,4	79,2	79,3	78,7	78,8	78,6	78,7

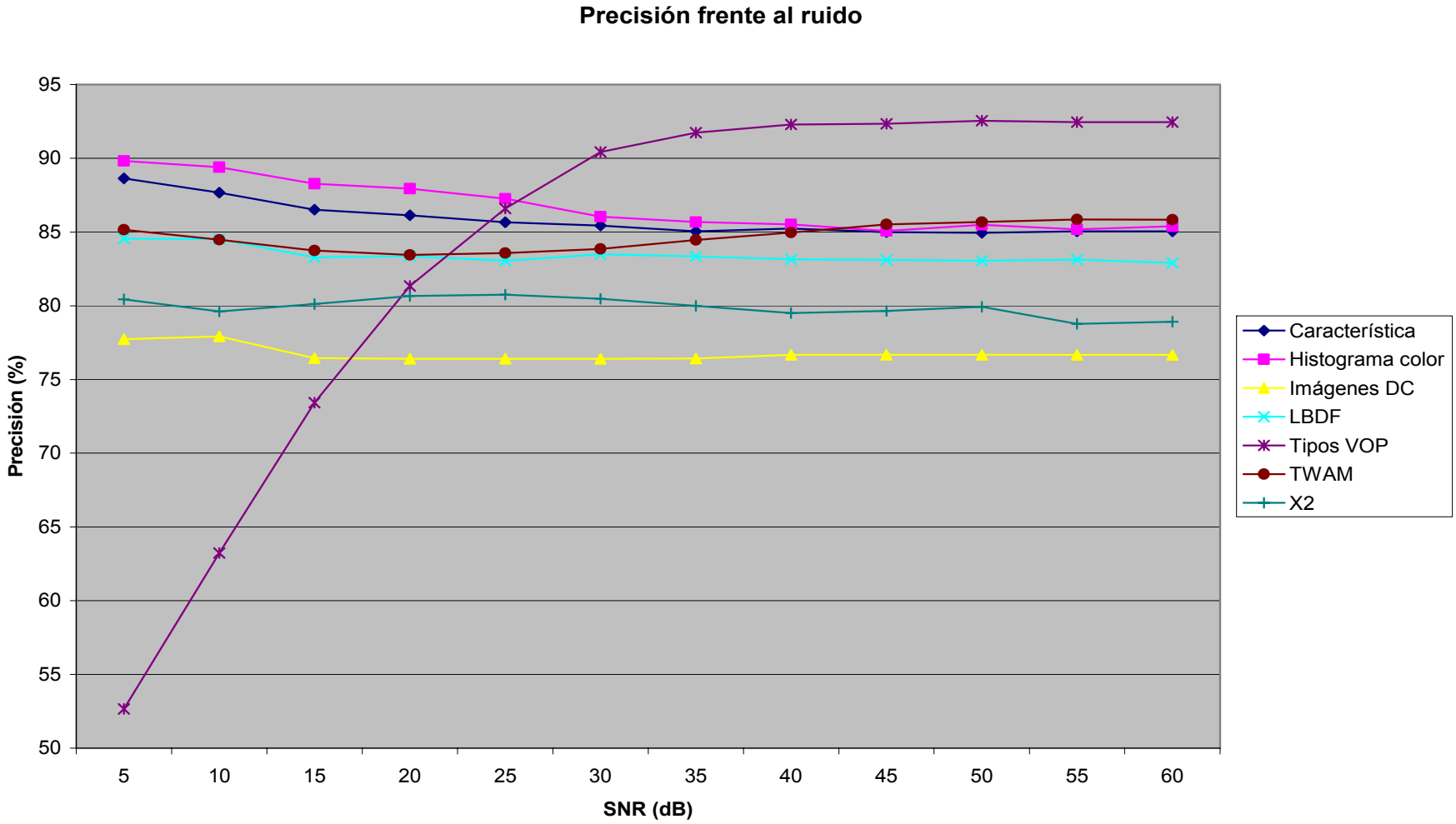

Ilustración 67: Precisión frente al ruido

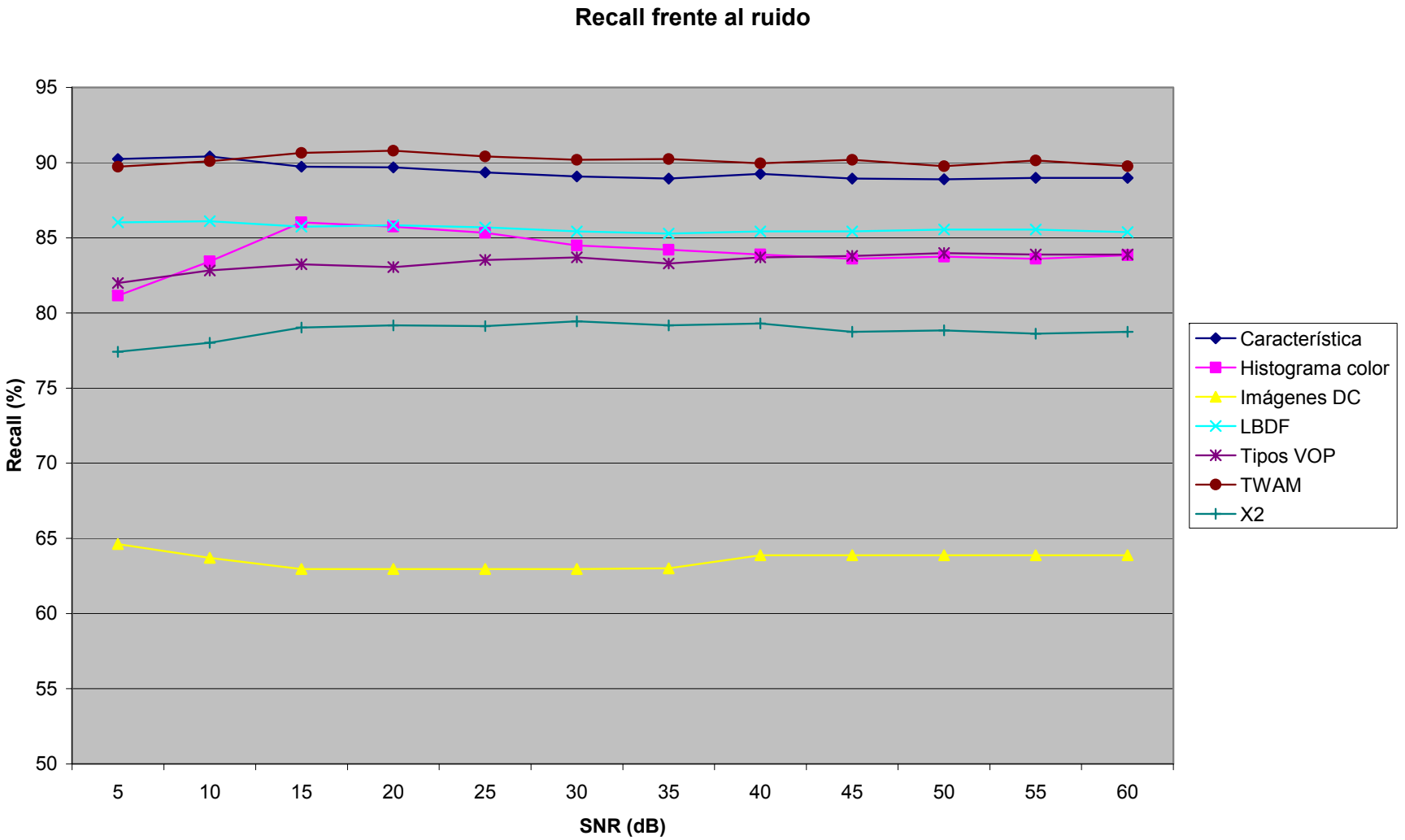

Ilustración 68: Recall frente al ruido

En las ilustraciones anteriores, puede apreciarse que los métodos varían ligeramente respecto a los resultados obtenidos anteriormente, y que permanecen estables con varias potencias de ruido. Sólo en el método de tipos de VOP se ven cambios producidos por el ruido introducido. Esos cambios sólo se aprecian en la precisión, ya que lo que ha hecho el ruido es aumentar el número de falsas detecciones. La mayoría de estas detecciones han ocurrido porque, al introducir ruido, el parámetro RP (que mide la proporción de macrobloques intra e inter en un VOP), que varía mucho de un cuadro a otro, se ve muy afectado y dispara las falsas alarmas.

4.9.2. Cambios en la tasa de bits

En esta prueba se han codificado los vídeos en varias tasas de bits para ver las variaciones que este parámetro introduce en los resultados.

Tabla 16: Precisión (%) frente a la variación de la tasa de bits

Bitrate (kbps)	150	286	420	572	700	900	1000
Histograma color	85,2	84,9	86,7	89,3	84,8	87,4	86,5
Imágenes DC	70,0	74,4	77,8	76,7	73,1	80,7	72,9
TWAM	88,2	88,0	87,7	86,0	86,5	88,1	85,5
Tipos VOP	93,7	92,9	92,0	91,2	91,3	89,6	84,8
Característica	83,0	86,4	85,7	86,1	88,5	86,1	90,9
LBDF	81,3	78,5	83,7	82,9	78,5	80,5	80,5
X^2	81,1	74,8	73,5	78,5	77,8	71,7	75,2

Tabla 17: Recall (%) frente a la variación de la tasa de bits

Bitrate(kbps)	150	286	420	572	700	900	1000
Histograma color	85,2	83,3	84,3	85,2	82,4	83,3	83,3
Imágenes DC	51,9	53,7	58,3	63,9	63,0	65,7	64,8
TWAM	83,3	88,0	86,1	90,7	88,9	88,9	87,0
Tipos VOP	82,4	85,2	85,2	86,1	87,0	88,0	88,0
Característica	76,9	88,0	88,9	91,7	92,6	91,7	92,6
LBDF	72,2	77,8	80,6	85,2	88,0	84,3	84,3
X^2	79,6	76,9	76,9	77,8	71,3	75,0	73,1

En las siguientes ilustraciones se pueden ver los resultados de forma gráfica. En ellas se puede apreciar que existe una cierta inestabilidad de las técnicas respecto a la tasa de bits. Los métodos que muestran esa inestabilidad de forma más evidente son el que emplea la diferencia de imágenes DC y el que realiza un análisis X^2 de los histogramas, los mismos que han dado peores resultados en el punto anterior.

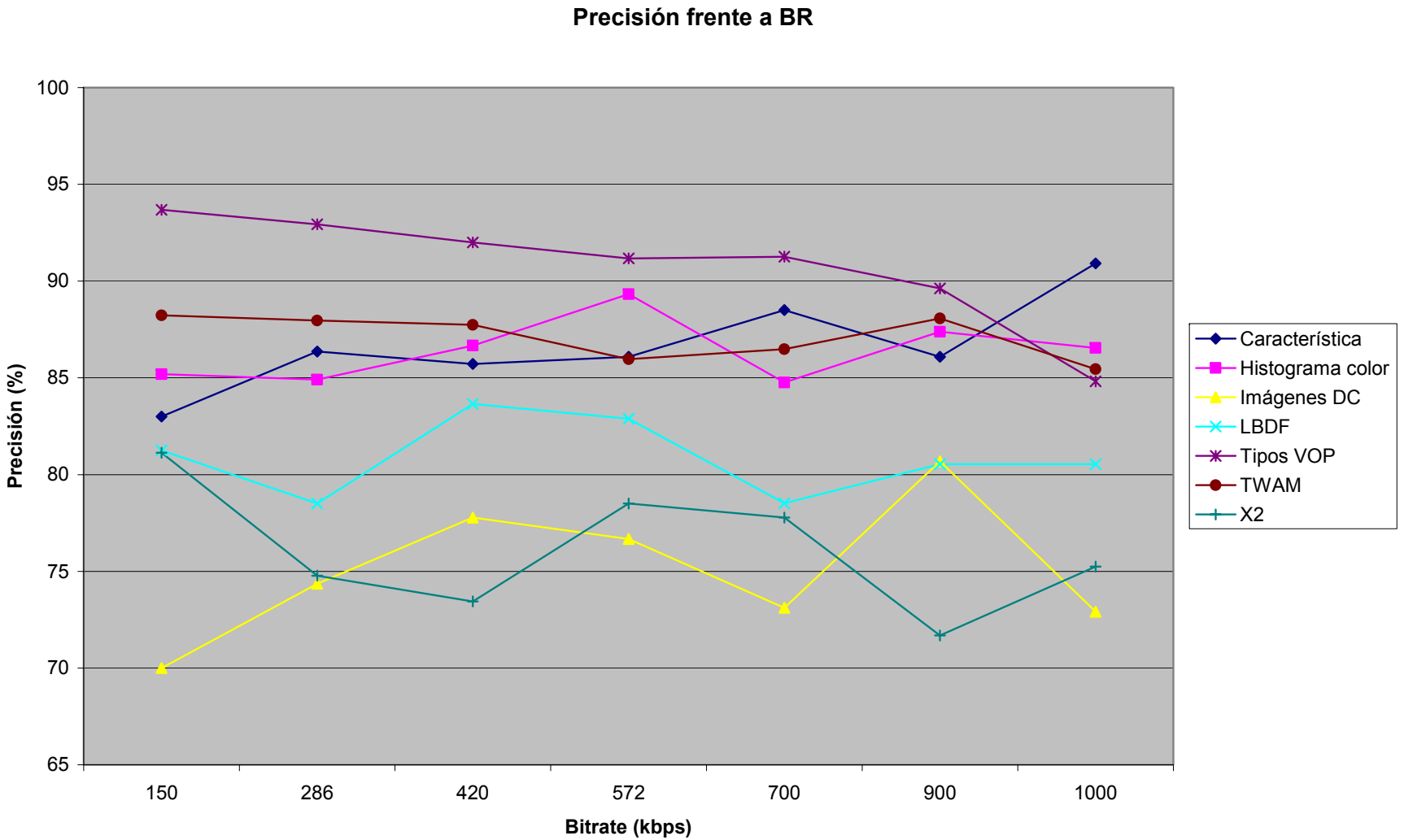

Ilustración 69: Precisión frente a la tasa de bits

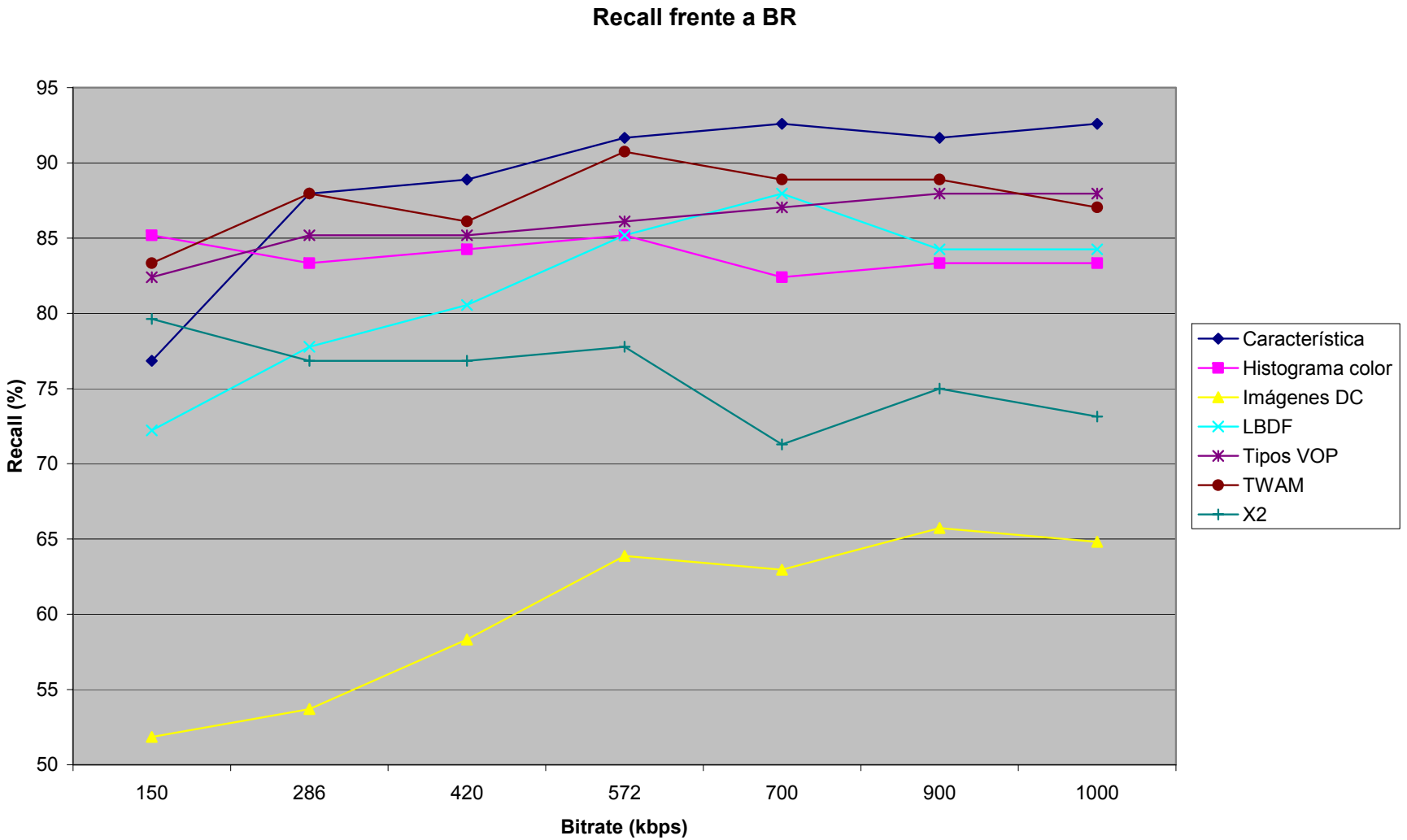

Ilustración 70: Recall frente a tasa de bits