

Extracción Automática de Fotogramas Informativos en Videos de Colonoscopia

Juan Arcila, Heber Mejía, Victor Tuesta, Maria Trujillo

Escuela de Ingeniería de Sistemas, Universidad Señor de Sipán – Perú
Escuela de Ingeniería de Sistemas y Computación, Universidad del Valle - Colombia
{diarcilaju, hmejia, vtuesta}@crece.uss.edu.pe, maria.trujillo@correounivalle.edu.co

Keywords: Estimación de movimiento, fotogramas borrosos.

Resumen

Los videos de colonoscopia contienen secuencias de fotogramas no-informativos, tales como fotogramas borrosos debido a movimientos rápidos del endoscopio; siendo necesario excluir dichos fotogramas para que el médico experto realice su trabajo en un menor tiempo. En este artículo se proponen dos métodos de visión artificial para la extracción automática de los fotogramas informativos basándose en las características que posee cada fotograma. El primer método permite la clasificación de fotogramas entre informativos y no informativos, basándose en la cantidad de contornos que posee cada fotograma. El segundo método utiliza flujo óptico denso para obtener el porcentaje de movimiento que tiene cada fotograma y agruparlos en informativos y no-informativos utilizando agrupamiento con K-Means. Ambos métodos procesan el video de colonoscopia y muestran como resultado un nuevo video que incluye solo los fotogramas informativos.

I. Introducción

Durante un procedimiento endoscópico, una pequeña cámara incluida en la punta del endoscopio genera una señal de video del interior de un órgano humano, dicha señal se almacena en un vídeo para el análisis por parte del médico. La colonoscopia que es un tipo de endoscopia, se emplea como una gran herramienta para la detección del cáncer colorrectal (enfermedad considerada como una de las principales causas de muerte en varios países) [1]. Los inconvenientes presentados en el análisis de videos de colonoscopia es que muchas partes del video resultante están fuera de foco debido a la rapidez con que avanza la cámara, la cámara choca con la mucosa del colon, entonces se crean fotogramas borrosos – llamados “fotogramas no-informativos” – que no muestran información significativa [2].

La aplicación de técnicas de visión artificial para la extracción automática de segmentos informativos en video de diagnóstico médico de colonoscopias es el propósito principal de este trabajo. Las propuestas para mejorar la visualización de los exámenes endoscópicos [3] [4] [5] así como la identificación de fotogramas informativos es un tema de investigación reciente [6], Hwang *et al* [7] proponen un método que se basan en que los fotogramas informativos y no informativos muestran patrones diferentes utilizando

Transformada Discreta del Coseno (DCT), como método de detección de contornos, permitiendo de esta manera clasificar los fotogramas en informativos y no informativos.

En este artículo se propone dos métodos de clasificación de fotogramas utilizando técnicas de detección de contornos – extrayendo como característica el total de contornos de cada fotograma –, detección de movimiento con flujo óptico denso [8] – para extraer como característica el porcentaje de movimiento que posee la transición de un fotograma a otro –, y usando el algoritmo K-Means con las características obtenidas para clasificar los fotogramas que muestran información importante (fotogramas informativos) y los fotogramas que no muestran información significativa (fotogramas no-informativos). Los resultados experimentales muestran los métodos tienen un desempeño similar, el método basado en la cantidad de contornos obtuvo una precisión promedio de 76,7%, y el método utilizando flujo óptico denso obtuvo una precisión promedio de 74,7%.

II. Trabajos Relacionados

En investigaciones previas se utilizaron técnicas para poder detectar los fotogramas informativos en una secuencia de video de diagnóstico médico, se encontró que en [9] fue implementado un método para excluir los fotogramas no-informativos, y asegurar la inclusión de los fotogramas que son importantes para documentar zonas endobronquiales o que revelan lesiones patológicas durante un examen broncoscópico. Para la detección de los fotogramas no-informativos fue utilizado el algoritmo DCT, ya que la propuesta se basa en el hecho de que los fotogramas no-informativos muestran espectros considerablemente menores en el dominio de la frecuencia en comparación con los fotogramas que sí poseen numerosas aristas, esta característica de DCT permite que sea usado como un simple pero eficaz método para la detección de bordes. Al final de la propuesta fue desarrollada una aplicación donde el usuario puede seleccionar la secuencia de vídeo a resumir; el paso siguiente se considera la intervención de la persona que supervisa y selecciona los parámetros de inclusión y exclusión de fotogramas; si el usuario no está satisfecho con el resumen resultante, este puede volver a ajustar los parámetros.

En la investigación en [7], fue desarrollado un método para evaluar la calidad de una colonoscopia y para excluir los fotogramas no-informativos su método desarrollado se basa en la Discrete Fourier Transform (DFT), donde los espectros de frecuencia de los fotogramas informativos y no-informativos

muestran diferentes patrones; por ejemplo, los fotogramas no-informativos no tienen información de objeto claro, excepto los cuatro bordes fuertes en las esquinas de una imagen corriendo aproximadamente a ± 45 direcciones. Mientras que los fotogramas informativos tienen una gran cantidad de información clara de los bordes para su espectro y no muestran componentes prominentes a lo largo de los ± 45 direcciones. Además obtienen de cada fotograma siete características de textura: entropía, contraste, correlación, homogeneidad, la desemejanza, el segundo momento angular y energía. Utilizaron después el algoritmo K-means en dos etapas para clasificar fotogramas en dos grupos (es decir, el grupo de fotogramas informativos y el grupo de fotogramas no-informativos), usando en las siete características de textura.

III. Metodos de Clasificación

Con el objetivo de la extracción automática de fotogramas informativos, utilizando técnicas de visión artificial, se proponen dos métodos adaptados para la extracción de fotogramas informativos en video de colonoscopia. El primer método se basa en los contornos teniendo como hipótesis que los fotogramas no-informativos poseen pocos contornos en comparación con los fotogramas informativos. El segundo método se basa en la detección de movimiento usando flujo óptico denso, para agrupar los fotogramas, en informativos y no-informativos, utilizando el algoritmo K-Means.

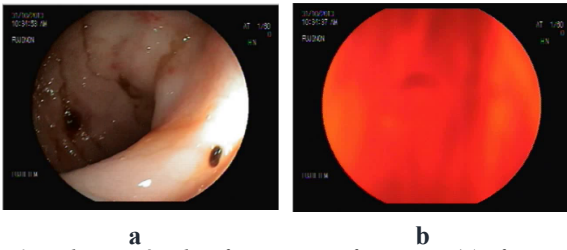


Fig. 1 – Ilustración de fotograma informativo (a), fotograma no-informativo (b).

3.1. Extracción de fotogramas informativos basándose en los contornos

El método propuesto se basa en la investigación de [9] donde se considera la cantidad de contornos que posee cada fotograma para clasificarlos en informativos y no informativos. El método usa un algoritmo distinto para la detección de contornos [10] [11] y determina de esta manera si un fotograma es informativo o no-informativo dependiendo de los contornos que posee, si tiene pocos contornos es un fotograma que no posee información significativa ya que puede ser un fotograma desenfocado por la rapidez con la que avanza la cámara o puede ser un fotograma capturado cuando la cámara ha chocado con la mucosa del colon.

a. Procesamiento inicial

Después de la entrada del video, se procesan inicialmente dos primeros fotogramas consecutivos (en Fig. 2), iterando hasta concluir con toda la secuencia de video.

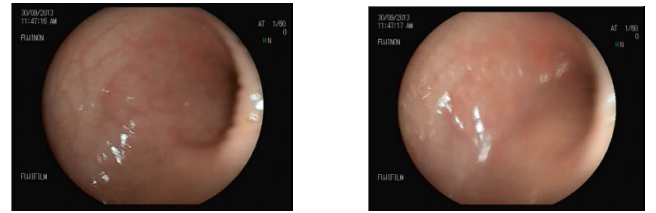


Fig. 2 – Ilustración de fotogramas consecutivos.

Se convierten los fotogramas de RGB (modelo de color que muestra la intensidad de los colores primarios de la luz Rojo, Verde, Azul) a escala de grises utilizando la siguiente ecuación:

$$y = 0.299R + 0.587G + 0.114B. \quad (1)$$

La Fig. 3 ilustra el resultado de la conversión a escala de grises de fotogramas consecutivos.

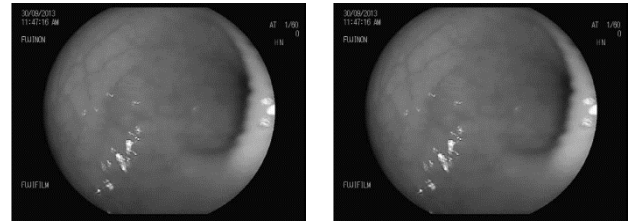


Fig. 3 – Fotogramas consecutivos en escala de grises.

b. Diferencia absoluta

Se crea un nuevo fotograma de la diferencia absoluta entre dos fotogramas. Dado los fotogramas $I_i(x, y)$ y $I_{i+1}(x, y)$, se obtiene el fotograma de diferencia $I_{diff}(x, y)$, y se crea una imagen binaria tal que:

$$I_{diff}(x, y) = \begin{cases} |I_i(x, y) - I_{i+1}(x, y)|, & \text{si } |I_i(x, y) - I_{i+1}(x, y)| \leq 255 \\ 255, & \text{si } |I_i(x, y) - I_{i+1}(x, y)| > 255. \end{cases} \quad (2)$$

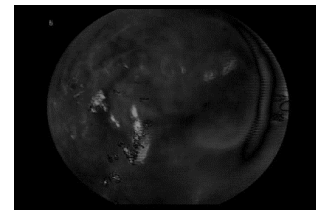


Fig. 4 – Fotograma resultante de la diferencia absoluta entre los fotogramas consecutivos.

c. Filtración de fotogramas no-informativos

Para determinar si un fotograma es informativo o no-informativo, se calcula la cantidad de contornos en cada fotograma. Después utilizando una estructura condicional se clasifican los fotogramas en informativos (Fig. 5) y no-informativos (Fig. 6) según el número de contornos que se presentan en ellos.

Si hay un número elevado de contornos se supone que el fotograma es informativo.

Si hay un número relativamente pequeño de contornos se supone que el fotograma es no-informativo.

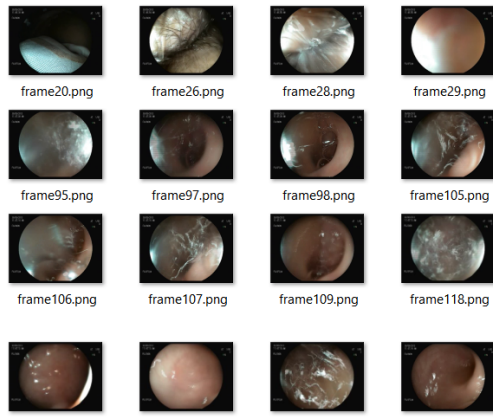


Fig. 5 – Fotogramas clasificados como informativos.

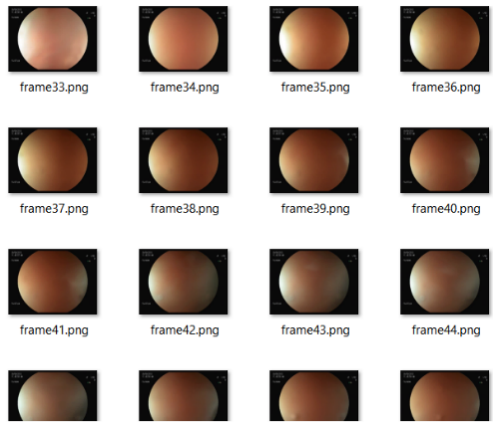


Fig. 6 – Fotogramas clasificados como no-informativos.

3.2. Extracción de fotogramas Informativos utilizando flujo óptico denso

Este método está basado en el movimiento que presenta la transición entre fotogramas, para la detección de movimiento se utiliza flujo óptico denso [8] [12]. Si los fotogramas consecutivos no presentan o presentan poco movimiento se establecerá que la cámara ha chocado con la mucosa del colon o con algún residuo intestinal, por el contrario si presentan movimiento regular se establecerá que son fotogramas informativos y si el movimiento es rápido se establecerá que las tomas están desenfocadas.

a. Procesamiento inicial

Al igual que el método anterior, los fotogramas se convierten de RGB a escala de grises.

b. Porcentaje de Movimiento de cada fotograma

Para obtener el movimiento, llamamos $I(x,y,t)$ a la intensidad del fotograma en el píxel (x,y) en el tiempo t y suponemos que existe una traslación en (v_x, v_y) tal que:

$$I(x + v_x, y + v_y, t + 1) = I(x, y, t). \quad (3)$$

Usando el desarrollo de primer orden de la serie de Taylor, se tiene que:

$$\frac{\partial I}{\partial t}(x, y, t) + \frac{\partial I}{\partial x}(x, y, t)v_x + \frac{\partial I}{\partial y}(x, y, t)v_y = 0. \quad (4)$$

Se obtienen dos matrices de flujo óptico denso, $A(x,y,t)$ donde se almacena si existe o no movimiento en la coordenada “x” y $B(x,y,t)$ para almacenar si existe o no movimiento en la coordenada “y”.

Se calcula $C(x,y,t)$, tal que $\{0,1\} \in C(x, y, t)$

$$C(x, y, t) = A(x, y, t) * B(x, y, t). \quad (5)$$

El porcentaje de pixeles (%P) que representan movimiento con valor 1 en $C(x,y,t)$, se obtiene:

$$P = \left(\frac{\sum_1^{x*y} C_{(x_i, y_i, t)}}{x * y} \right) * 100. \quad (6)$$

El porcentaje de movimiento (%M) que existe en la transición de un fotograma (F) a otro se muestra en la Tabla 1.

| F | %M | F | %M | F | %M | F | %M |
|----|-------|----|-------|----|-------|-----|-------|
| 1 | 2.52 | 32 | 53.17 | 63 | 0.52 | 94 | 45.63 |
| 2 | 0.00 | 33 | 35.04 | 64 | 0.00 | 95 | 52.38 |
| 3 | 4.73 | 34 | 4.13 | 65 | 2.13 | 96 | 50.13 |
| 4 | 0.00 | 35 | 2.21 | 66 | 1.60 | 97 | 55.48 |
| 5 | 4.42 | 36 | 0.52 | 67 | 0.50 | 98 | 67.54 |
| 6 | 6.38 | 37 | 0.31 | 68 | 1.79 | 99 | 44.04 |
| 7 | 3.48 | 38 | 0.00 | 69 | 0.94 | 100 | 34.96 |
| 8 | 7.77 | 39 | 2.35 | 70 | 0.88 | 101 | 53.48 |
| 9 | 37.58 | 40 | 8.94 | 71 | 0.00 | 102 | 43.48 |
| 10 | 6.31 | 41 | 30.73 | 72 | 2.25 | 103 | 38.71 |
| 11 | 13.52 | 42 | 41.17 | 73 | 2.27 | 104 | 62.67 |
| 12 | 20.92 | 43 | 48.42 | 74 | 6.77 | 105 | 59.06 |
| 13 | 41.81 | 44 | 8.29 | 75 | 12.42 | 106 | 67.75 |
| 14 | 11.08 | 45 | 18.71 | 76 | 20.42 | 107 | 58.85 |
| 15 | 7.31 | 46 | 21.08 | 77 | 36.52 | 108 | 64.21 |
| 16 | 14.81 | 47 | 18.94 | 78 | 0.08 | 109 | 62.50 |
| 17 | 17.58 | 48 | 16.56 | 79 | 3.98 | 110 | 65.31 |
| 18 | 5.27 | 49 | 34.33 | 80 | 21.96 | 111 | 79.04 |
| 19 | 30.94 | 50 | 60.96 | 81 | 6.67 | 112 | 68.54 |
| 20 | 42.52 | 51 | 71.81 | 82 | 12.10 | 113 | 62.13 |
| 21 | 37.31 | 52 | 13.81 | 83 | 6.71 | 114 | 33.42 |
| 22 | 50.19 | 53 | 19.33 | 84 | 6.75 | 115 | 19.63 |
| 23 | 61.44 | 54 | 11.81 | 85 | 3.00 | 116 | 22.65 |
| 24 | 70.48 | 55 | 5.02 | 86 | 6.79 | 117 | 14.71 |
| 25 | 69.58 | 56 | 9.35 | 87 | 9.71 | 118 | 17.90 |
| 26 | 55.98 | 57 | 3.88 | 88 | 29.13 | 119 | 21.42 |
| 27 | 68.77 | 58 | 3.27 | 89 | 48.48 | 120 | 22.50 |
| 28 | 72.06 | 59 | 3.17 | 90 | 46.92 | 121 | 5.77 |
| 29 | 75.29 | 60 | 0.79 | 91 | 24.35 | 122 | 7.48 |
| 30 | 25.77 | 61 | 1.31 | 92 | 47.42 | 123 | 4.92 |
| 31 | 42.46 | 62 | 0.00 | 93 | 51.19 | 124 | 4.10 |

Tabla. 1– Primeros 124 fotogramas de un video de muestra y su correspondiente porcentaje de movimiento.

c. Agrupación de Fotogramas

El porcentaje de movimiento obtenido es usado como variable de agrupamiento con el algoritmo K-means de una etapa para clasificar los fotogramas, se establece inicialmente el número de grupos igual a 3 ($\mathbf{m}_1^{(1)}, \mathbf{m}_2^{(1)}, \mathbf{m}_3^{(1)}$): bastante movimiento, movimiento medio y poco movimiento

Se continua iterando dos pasos:

Paso Asignación: Se realiza primero la asignación de cada valor (x_p) al grupo $(S_i^{(t)})$.

$$S_i^{(t)} = \{x_p: |x_p - m_i^{(t)}| \leq |x_p - m_j^{(t)}| \forall 1 \leq j \leq k\} \quad (7)$$

Paso de actualización: Se calculan los nuevos centroides, como el centroide de las observaciones en el grupo.

$$\frac{\partial I}{\partial t}(x, y, t) + \frac{\partial I}{\partial x}(x, y, t)v_x + \frac{\partial I}{\partial y}(x, y, t)v_y = 0. \quad (8)$$

Se repiten estos pasos hasta que las asignaciones de los valores a los grupos no cambian.

Finalizando de esta manera con la agrupación de los fotogramas en tres grupos (G): 0 indica fotogramas con mucho movimiento, 1 indica grupo de los fotogramas con movimiento medio y 2 indica grupo de los fotogramas con poco movimiento. La tabla 2 ilustra la clasificación de los primeros 124 fotogramas y el grupo correspondiente al que pertenece.

| F | G | F | G | F | G | F | G | F | G | F | G | F | G |
|----|---|----|---|----|---|----|---|-----|---|-----|---|---|---|
| 1 | 2 | 22 | 0 | 43 | 1 | 64 | 2 | 85 | 2 | 106 | 0 | | |
| 2 | 2 | 23 | 0 | 44 | 2 | 65 | 2 | 86 | 2 | 107 | 0 | | |
| 3 | 2 | 24 | 0 | 45 | 2 | 66 | 2 | 87 | 2 | 108 | 0 | | |
| 4 | 2 | 25 | 0 | 46 | 2 | 67 | 2 | 88 | 1 | 109 | 0 | | |
| 5 | 2 | 26 | 0 | 47 | 2 | 68 | 2 | 89 | 1 | 110 | 0 | | |
| 6 | 2 | 27 | 0 | 48 | 2 | 69 | 2 | 90 | 1 | 111 | 0 | | |
| 7 | 2 | 28 | 0 | 49 | 1 | 70 | 2 | 91 | 2 | 112 | 0 | | |
| 8 | 2 | 29 | 0 | 50 | 0 | 71 | 2 | 92 | 1 | 113 | 0 | | |
| 9 | 1 | 30 | 1 | 51 | 0 | 72 | 2 | 93 | 0 | 114 | 1 | | |
| 10 | 2 | 31 | 1 | 52 | 2 | 73 | 2 | 94 | 1 | 115 | 2 | | |
| 11 | 2 | 32 | 0 | 53 | 2 | 74 | 2 | 95 | 0 | 116 | 2 | | |
| 12 | 2 | 33 | 1 | 54 | 2 | 75 | 2 | 96 | 0 | 117 | 2 | | |
| 13 | 1 | 34 | 2 | 55 | 2 | 76 | 2 | 97 | 0 | 118 | 2 | | |
| 14 | 2 | 35 | 2 | 56 | 2 | 77 | 1 | 98 | 0 | 119 | 2 | | |
| 15 | 2 | 36 | 2 | 57 | 2 | 78 | 2 | 99 | 1 | 120 | 2 | | |
| 16 | 2 | 37 | 2 | 58 | 2 | 79 | 2 | 100 | 1 | 121 | 2 | | |
| 17 | 2 | 38 | 2 | 59 | 2 | 80 | 2 | 101 | 0 | 122 | 2 | | |
| 18 | 2 | 39 | 2 | 60 | 2 | 81 | 2 | 102 | 1 | 123 | 2 | | |
| 19 | 1 | 40 | 2 | 61 | 2 | 82 | 2 | 103 | 1 | 124 | 2 | | |
| 20 | 1 | 41 | 1 | 62 | 2 | 83 | 2 | 104 | 0 | | | | |
| 21 | 1 | 42 | 1 | 63 | 2 | 84 | 2 | 105 | 0 | | | | |

Tabla 2– Primeros 124 fotogramas de un video de muestra etiquetados con su correspondiente grupo.

Finalmente, los fotogramas se clasifican dependiendo al grupo al que pertenezcan: el grupo 0 son fotogramas con mucho movimiento o no-informativos (Fig. 7), el grupo 1 son fotogramas con movimiento medio o informativos (Fig. 8) y el grupo 2 son fotogramas con poco movimiento o no-informativos (Fig. 9).

IV. EXPERIMENTOS

Para evaluar los métodos propuestos se utilizo una base de datos compuesta por 20 videos de colonoscopia, evaluando 309

fotogramas por cada video, con un total de 12360 fotogramas en cada prueba realizada.

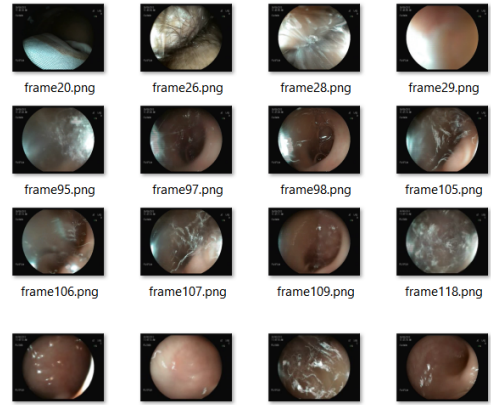


Fig. 7– Grupo 1, fotogramas informativos.

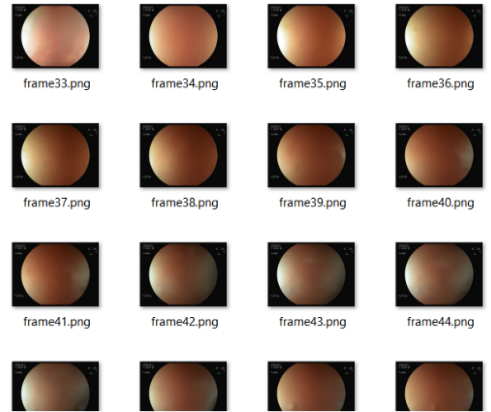


Fig. 8– Grupo 2, fotogramas no informativos.

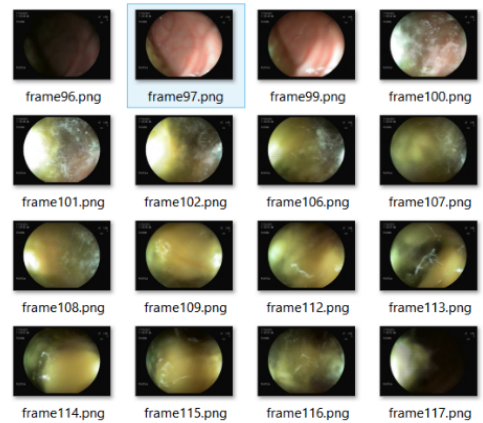


Fig. 9– Grupo 0, fotogramas no-informativos.

Los métodos propuestos fueron puestos a prueba para poder evaluar las siguientes métricas:

- Tiempo que tarda cada método en procesar una secuencia de video

Esta medida es el resultado del tiempo que tarda cada método en procesar uno a uno los fotogramas de cada secuencia de vídeo para extraer los fotogramas informativos.

| Promedio de tiempo de procesamiento de cada método | |
|--|--------------|
| Métodos | Tiempo (Seg) |
| Método basado en contornos | 2,099 |
| Método basado en movimiento | 3,853 |

Tabla 3– Promedio de tiempo de procesamiento de cada método.

El método basándose en la cantidad de contornos procesa la secuencia de vídeo un 83,5% más rápido en comparación con el método basado en el movimiento de cada fotograma, utilizando flujo óptico denso.

b. Fotogramas informativos correctamente clasificados

El porcentaje de precisión de cada método para la clasificación de fotogramas informativos, está dado por:

$$P = \frac{|\#Taciertos \cap \#Tmuestras|}{|\#Tmuestras|} \quad (9)$$

En la Tabla 4 se observa que el método basándose en la cantidad de contornos obtiene mayor precisión, en promedio un 68,49% para la clasificación de fotogramas informativos; mientras que el método basado en el movimiento de cada fotograma obtuvo un porcentaje de 60,10 % de precisión para la clasificación de fotogramas informativos.

| Fotogramas informativos correctamente clasificados | | | | |
|--|--------------------|----------------------|-------------|---------|
| Métodos | Promedio Correctos | Promedio Incorrectos | % Precisión | % Error |
| Método basado en contornos | 97,40 | 43,85 | 68,49% | 31,51% |
| Método basado en movimiento | 60,25 | 36,20 | 60,10% | 39,90% |

Tabla 4– Promedio de precisión de cada método para la clasificación de fotogramas informativos.

c. Fotogramas no informativos correctamente clasificados

| Fotogramas No Informativos correctamente clasificados | | | | |
|---|--------------------|----------------------|-------------|---------|
| Métodos | Promedio Correctos | Promedio Incorrectos | % Precisión | % Error |
| Método basado en contornos | 139,55 | 28,20 | 82,85% | 17,15% |
| Método basado en movimiento | 170,70 | 41,85 | 79,58% | 20,42% |

Tabla 5– Promedio de precisión de cada método para la clasificación de fotogramas no-informativos.

De igual manera, utilizando la ecuación (9), en la Tabla 5 se observa que el método basado en la cantidad de contornos obtiene mayor precisión, en promedio un 82,85%, para la clasificación de fotogramas no informativos; mientras que el método basado en el movimiento obtuvo un porcentaje de 79,58 % de precisión para la clasificación de fotogramas no informativos.

d. Reducción de tiempo de vídeo resultante

El porcentaje de reducción que se obtiene en cada secuencia de vídeo usando la siguiente ecuación:

$$R = \frac{\text{TiempoFinalInformativos}}{\text{TiempoVideoInicial}} \times 100. \quad (10)$$

En la Tabla 6 se observa que el método basado en la cantidad de contornos obtiene un porcentaje 45,71% de reducción; mientras que el método basado en el movimiento obtuvo un porcentaje de 31,21 % de reducción en el tiempo de duración de los vídeos resultantes.

| Promedio de reducción de los vídeos resultantes | |
|---|-----------|
| Métodos | Reducción |
| Método basado en contornos | 45,71% |
| Método basado en movimiento | 31,21% |

Tabla 6– Promedio de reducción de los vídeos resultantes incluyendo solo los fotogramas informativos.

e. Precisión final

Para obtener la precisión final se ha considerado la precisión obtenida en la clasificación de los fotogramas informativos y la clasificación de los fotogramas no-informativos. Los resultados obtenidos de la precisión final de cada método utilizando la ecuación (9) se muestran en la Tabla 7. Comparando los métodos según el porcentaje de precisión promedio obtenido, se puede notar que el método con mayor precisión es el método basado en la cantidad de contornos que posee cada fotograma, con una precisión del 76.7 %, en comparación con el método basado en el porcentaje de movimiento que posee cada fotograma utilizando flujo óptico denso, que obtiene una precisión del 74.7%.

| Precisión final de cada método | |
|--------------------------------|-----------|
| Métodos | Precisión |
| Método basado en contornos | 76.7% |
| Método basado en movimiento | 74.7% |

Tabla 7– Precisión de cada método para la extracción de fotogramas informativos.

V. Conclusiones

Después de las pruebas realizadas, se determinó qué de los dos métodos propuestos el método más eficiente – que consume menos recursos, y que tarda menos en procesar una secuencia de vídeo para extraer automáticamente los fotogramas informativos – es el método basado en la cantidad de contornos, el cual obtuvo una precisión de 76,7%, en comparación con el método que permite la extracción de fotogramas informativos utilizando flujo óptico denso, que también obtiene buenos resultados alcanzado un promedio de precisión de 74,7%.

En un futuro, se explorará el uso del algoritmo DCT para extraer las características de agrupamiento, asegurando de esta

manera que todos los fotogramas informativos, necesarios para el diagnóstico médico, estén incluidos en el video resultante.

Agradecimientos

Agradecimiento al Dr. Jeon HiunKing y Lic. Steven Hunt, quienes a pesar de la distancia nos brindaron apoyo.

Bibliografía

- [1] S. Kuwada, «Colorectal cancer 2000 - Education and screening are essential if outcomes are to improve.,» *Postgraduate Medicine*, pp. [2000, 107(5):96-8, 101-4, 107], 2000.
- [2] O. JungHwan, H. sae, L. JeongKyu, T. Wallapak, J. Wong y P. C. De Groen, «Informative frame classification for endoscopy video,» *Journal of Medical Image Analysis* 11(2):110-27. Epub 2007 Feb 27., Bridgeport USA, 2007 Apr.
- [3] V. Parot, L. Daryl, G. Germán, T. Giovanni, S. N. Norman, J. V. Benjamin y J. D. Nicholas, «Photometric stereo endoscopy.,» *Journal of Biomedical Optics* 18 (7): 076017. doi:10.1117/1.JBO.18.7.076017. <http://dx.doi.org/10.1117/1.JBO.18.7.076017>, 2013.
- [4] C. Li, A. B. Hamza, N. Bouguila, X. Wang, F. Ming y G. Xiao, «Online redundant image elimination and its application to wireless capsule endoscopy,» *Signal, Image and Video Processing* 10.1007/s11760-012-0384-3, vol. 8, n° 8, pp. 1497-1506, 2014.
- [5] Z. Mikrut y M. Duplaga, «Analysis of images of bronchial tree ramifications for summarization of Bronchoscopic Videos,» *Information Technologies in Biomedicine*, vol. 69, pp. 583-586, 2010.
- [6] S. Manivannan, R. Wang, E. Trucco y A. Hood, «Automatic Normal-Abnormal Video Frame Classification,» *CVIP Computer Vision and Image Processing group*, pp. 1-4, 2011.
- [7] S. Hwang, J. Oh, J. L. Lee, C. Cao, T. Wallapak, D. Liu, J. Wong y P. C. de Groen, «Automatic Measurement of Quality Metrics for Colonoscopy Videos,» *MULTIMEDIA '05 Proceedings of the 13th annual ACM international conference on Multimedia* doi: 10.1145/1101149.1101343, pp. 912-921, 2005.
- [8] S. Beauchemin y J. Barron, «The computation of optical flow,» *ACM Computing Surveys (CSUR)* doi: 10.1145/212094.212141, vol. 27, n° 3, pp. 433-466, 1995 .
- [9] M. I Leszczuk y D. Mariusz, *Algorithm for Video Summarization of Bronchoscopy Procedures*, Krakow-Polonia: BioMedical Engineering OnLine doi:10.1186/1475-925X-10-110, 2011.
- [10] S. Suzuki y K. Abe, «Topological Structural Analysis of Digitized Binary Images by Border Following,» *Computer Vision, Graphics, And Image processing* doi:10.1016/0734-189X(85)90016-7, vol. 30, n° 1, pp. 32-46, 2004.
- [11] E. Rostend y T. Drummond, «Machine Learning for High-Speed Corner Detection,» *Computer Vision – ECCV 2006* doi: 10.1007/11744023_34, vol. 3951, pp. 430-443, 2006.
- [12] J. Barron y N. Thacker, «Computing 2d and 3d optical flow,» *Imaging Science and Biomedical Engineering Division, Medical School, University of Manchester, Manchester*, 2005.