

Utilización de FPGAs para el procesamiento digital de video

J. Javier Martínez, F. Javier Toledo, F. Javier Garrigós, J. Manuel Ferrández
Dpto. Electrónica, Tecnología de Computadoras y Proyectos. Universidad Politécnica de Cartagena
Campus Muralla de Mar. Edificio Antiguo Hospital de Marina
30202 Cartagena
Teléfono: 968326467 Fax: 968 326400
E-mail: javier.toledo@upct.es

Resumen. Este trabajo muestra alguna de las técnicas de procesamiento digital de video desarrolladas por el grupo de Hardware Reconfigurable del Dpto. de Electrónica, Tecnología de Computadoras y Proyectos de la UPCT. En él se resume la implementación de diversos algoritmos de procesamiento, los cuales plantean diferentes alternativas para la obtención de bordes de imágenes de video en tiempo real. Los algoritmos que se proponen son: extracción de bordes mediante el método del gradiente, algoritmo de CANNY y un método basado en redes neuronales celulares discretas (DT-CNN). Todas las implementaciones realizadas poseen como elemento principal la utilización de dispositivos lógicos reconfigurables (FPGAs), cuya aportación principal al desarrollo es conseguir la aceleración hardware necesaria para procesar las imágenes de video a alta velocidad y en el mínimo espacio posible.

1 Introducción

Las aplicaciones de procesamiento de imagen requieren, cada día más, algoritmos de cálculos complejos que demandan unidades de procesamiento rápidas y de gran capacidad. Esto es aún más patente, en aplicaciones en las que las imágenes tienen que ser procesadas en tiempo real o en las que su entrada es una señal de video generada por un CCD o por una video-cámara. Para estas aplicaciones, encontrar una solución adecuada que justifique su implementación suele ser una tarea complicada, ya que su alto coste y dimensiones físicas limitan su desarrollo de forma considerable. De este modo, potentes soluciones basadas en arquitecturas de cómputo paralelas, tales como redes de computadores o cluster de microprocesadores, tienen su uso limitado a determinadas aplicaciones en grandes industrias o centros de investigación debido a su alto coste y excesivas dimensiones.

En este sentido, en nuestro grupo estamos desarrollando y evaluando nuevas alternativas, basadas en dispositivos reconfigurables FPGAs, con el propósito de satisfacer las exigencias de las aplicaciones de procesamiento de video mejorando su velocidad, portabilidad y costes de desarrollo. Esto es posible, gracias a que las FPGAs ofrecen un interesante compromiso entre los circuitos integrados de aplicación específica (ASICs) y los procesadores de propósito general (GPPs). Por un lado las FPGAs son capaces de implementar las capacidades de cómputo y paralelismo inherentes de los pequeños ASICs, mientras que por otro lado, proporcionan la reconfigurabilidad, economía, rapidez y flexibilidad de los GPPs. Todas estas características hacen que las FPGAs deban ser tenidas en cuenta, a la hora de abordar la implementación de aplicaciones de procesamiento de imagen y video en tiempo real.

2 Aplicaciones

A lo largo de los últimos años hemos venido desarrollando diferentes técnicas de procesamiento de video que suponen importantes avances en cuanto a la economía y aceleración hardware de determinados algoritmos de procesamiento. En las secciones siguientes se describirán, a modo de ejemplo algunos de los trabajos desarrollados. Todos ellos son alternativas para obtener el contorno de imágenes de video en tiempo real y presentan importantes mejoras en velocidad frente a su correspondiente implementación software.

2.1 Algoritmo basado en gradiente

Una de las alternativas más sencillas para extraer el contorno de una imagen es mediante su derivada o gradiente espacial. El gradiente espacial es una herramienta que proporciona información de la variación de los niveles de gris de la imagen y por tanto de las fronteras de los objetos (los bordes). Una manera de obtener una aproximación discreta del gradiente es la siguiente:

$$G = \sqrt{G_x + G_y} \approx |G_x| + |G_y|.$$

Donde G_x y G_y son respectivamente los gradientes direccionales de la imagen en direcciones x e y , obtenidos tras las convolución de la imagen con las mascararas P_x y P_y ,

$$P_x = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}; \quad P_y = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix},$$

las cuales, representan una aproximación a los operadores derivada en las direcciones x e y respectivamente.

El esquema de la Fig.1 representa la arquitectura del algoritmo de extracción de bordes implementado sobre la FPGA. Para realizar este diseño se ha utilizado una herramienta de alto nivel, System Generator™ de Xilinx, la cual permite facilitar considerablemente el proceso de desarrollo. La arquitectura del algoritmo comienza con un bloque de convolución, que realiza un suavizado (mediante un filtrado gaussiano) eliminando parte de ruido existente en la entrada de video. Posteriormente aparecen los bloques de convolución P_x , P_y , para obtener las derivadas direccionales, y por último los bloques de valor absoluto y de suma con los que se obtiene el resultado definitivo. La implementación completa del circuito ha sido realizada sobre una FPGA tipo XC4VLX25 utilizando bloques de convolución de altas prestaciones desarrollados por nosotros [1]. El total de los recursos consumidos por la aplicación ha sido de un 3% de la capacidad de la FPGA.

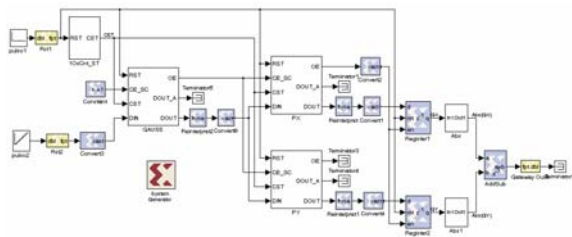


Figura 1. Implementación de etapa basada en gradientes.

Tabla 1. Velocidad de procesamiento de la etapa.

El circuito ha sido diseñado para conectarse a una fuente de video con calidad DVD (tamaño 720x576 píxeles y 25 frames por segundo). En la Fig. 2 se muestra la salida de video obtenida tras el procesamiento.



Fig. 2. Respuesta del procesamiento basado en gradiente.

La Tabla 1 muestra la relación existente entre el tamaño de la imagen y el número de imágenes procesadas por el circuito.

Tam. imagen (Píxeles.)	Imágenes/s
208x160	733.47
320x240	317.84
640x480	79.46
720x576	58.86
800x600	50.85
1024x768	31.04
1024x1024	23.28

Tabla 1. Velocidad de procesamiento de la etapa basada en el gradiente.

2.2 Método de CANNY

El algoritmo de Canny [2] está considerado como uno de los mejores métodos de detección de contornos existentes en la actualidad. Este método obtiene muy buenos resultados con el uso de tres criterios: localización, detección y respuesta única ante un borde. Básicamente el algoritmo consiste en las siguientes cinco fases:

- Suavizado de la imagen mediante una Gaussiana para reducir los efectos del ruido.
- Convolución de la imagen por operadores de gradiente direccionales.
- Supresión de los gradientes que no son máximos en la dirección del contorno.
- Aplicación de un umbral con histéresis para distinguir los contornos reales del ruido.
- Integración de los contornos detectados con diferentes umbrales en un solo mapa de contornos consistentes.

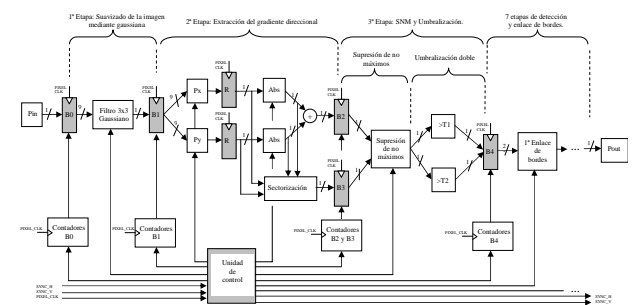


Fig. 3. Arquitectura hardware utilizada para implementar el algoritmo de extracción de bordes Canny.

La implementación de la arquitectura de la Fig. 3 ha sido realizada sobre una FPGA XC2v1000-6 utilizando el lenguaje de descripción hardware VHDL. El circuito es capaz de procesar más de 50 millones de píxeles/s por etapa, o de manera equivalente, imágenes de vídeo de 1024x768 píxeles a 60 hz, consiguiendo una capacidad de cálculo superior a 3 GOPS. La Fig. 4 muestra un ejemplo de

las distintas etapas de procesamiento realizadas por el algoritmo

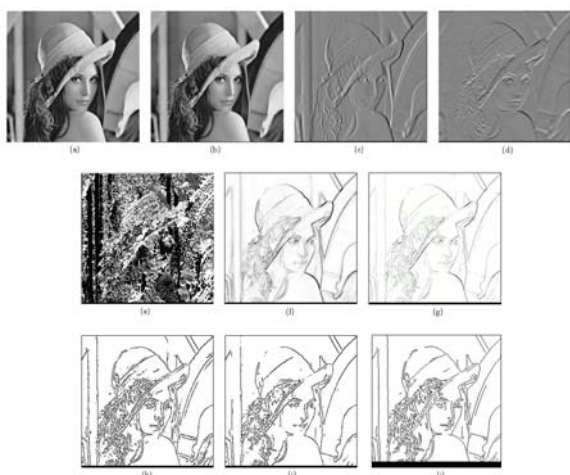


Fig. 4. Distintas etapas del procesamiento de una imagen: (a) Imagen original. (b) Imagen suavizada. (c) Gradiente en dirección x . (d) Gradiente en dirección y . (e) Sectorización del ángulo del módulo. (f) Módulo del gradiente. (g) Eliminación de no máximos. (h) Umbralización menos exigente. (i) Umbralización más exigente. (j) Última etapa de enlace.

2.3 Algoritmo basado en DT-CNN

La capacidad de aprendizaje de las redes neuronales artificiales (ANNs) hace posible que su respuesta pueda ser ajustada a un valor deseado en función de una serie de coeficientes o pesos. La versatilidad de las ANNs hace que estas sean una herramienta muy utilizada en diversos campos de la ciencia e industria y que con ellas se obtengan resultados excelentes en muchas aplicaciones. De entre todas las ANNs, existe un tipo especial, las Redes Neuronales Celulares (CNN), que ofrecen características ideales para ser utilizadas en aplicaciones de procesamiento de imagen y video. Una de las características que diferencian a estas redes de otras es su conectividad local. La conectividad de las CNNs es similar a las encontradas en las redes neuronales biológicas y permiten, de forma sencilla, obtener diferentes procesamientos de una imagen simplemente modificando los pesos de la red (sus templates). Las CNN fueron propuestas por Chua y Yang en 1988 y son la base de las redes neuronales discretas (DT-CNN). En la Fig. 5 se muestra una imagen sintética de este tipo de red.

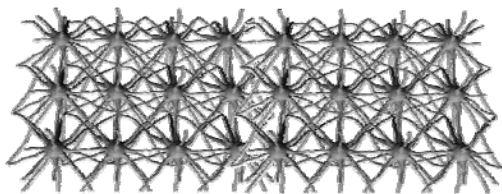


Fig. 5. Simulación 3D de una CNN de 24 neuronas con conectividad local de radio 1.

Las ecuaciones dinámicas que rigen la conducta de las DT-CNNs son las siguientes.

$$X_{ij}[k] = \sum_{k,l \in Nr(ij)} A_{kl} Y_{kl}[k-1] + \sum_{k,l \in Nr(ij)} B_{kl} Y_{kl}[k-1] + I_{ij}$$

$$Y_{ij}[k] = f(Y_{ij}[k]) = \frac{1}{2} (|X_{ij}[k] + 1| - |X_{ij}[k] - 1|)$$

Este modelo ha sido implementado sobre una FPGA, empleando una arquitectura altamente eficiente desarrollada por nosotros [3]. La arquitectura propuesta proporciona importantes ventajas en velocidad de procesamiento y consumo de área, posibilitando que el sistema sea capaz de emular el comportamiento de más de 300.000 neuronas procesando video en tiempo real.

De las diferentes y potentes posibilidades que ofrecen las DT-CNN, en la Fig. 6 se muestra un ejemplo obtenido tras el entrenamiento de la red para obtener los bordes de una imagen.



Fig. 6. Extracción de bordes utilizando DT-CNN.

Agradecimientos

Estos trabajos se realizan en el marco de los proyectos de investigación TIC 2003-09557-C02-02 y TIC 2003-09400-C04-02 del MCYT.

Referencias

- [1] J. J. Martínez, F. J. Toledo, F. J. Garrigós, J. M. Ferrández. "Convolución 3x3 de alta eficiencia para el diseño de aplicaciones de procesamiento de video a alto nivel". Proc. Jornadas Computación Reconfigurable y Aplicaciones, JCRA06. Cáceres, Sept. 2006.
- [2] J. J. Martínez, F. J. Toledo, F. J. Garrigós, J. M. Ferrández. "Implementación sobre Hardware Reconfigurable del Algoritmo de CANNY para la Extracción de Contornos en Imágenes de Vídeo de Gran Tamaño en Tiempo Real". Proc. Jornadas Computación Reconfigurable y Aplicaciones, JCRA03. Madrid, Sept. 2003.
- [2] J. J. Martínez, F. J. Toledo, F. J. Garrigós, J. M. Ferrández. "Implementation of a Discrete Cellular Neuron Model (DT-CNN) Architecture on FPGA", Proc. 2° SPIE Int. Conf. on bioengineered and bioinspired systems, vol. 5389, 389-397, Sevilla, 2005.