

# Implementación del Detector de Bordes de Canny sobre Redes Neuronales Celulares

Miguel A. Jaramillo\*, J. Álvaro Fernández\*, Enrique Martínez de Salazar\*\*

\*Dpto. de Electrónica e Ingeniería Electromecánica. \*\*Dpto. de Expresión Gráfica.

Escuela de Ingenierías Industriales. Universidad de Extremadura. Avda. de Elvas s/n. 06071. Badajoz (España)

Tel.: +34 924 289628. E-mail: [miguel@unex.es](mailto:miguel@unex.es), [jalvarof@unex.es](mailto:jalvarof@unex.es), [dsalazar@unex.es](mailto:dsalazar@unex.es).

## Resumen

La detección de bordes mediante la localización de discontinuidades en el nivel de gris de una imagen digital es una de las aproximaciones más utilizadas en el área del Procesamiento Digital de Imágenes (PDI). El presente trabajo revisa los conceptos básicos de los métodos de detección de bordes mediante filtrado espacial, y en particular el método de Canny, para proporcionar una técnica basada en una estructura de procesamiento paralelo: la Red Neuronal Celular (CNN). Las ventajas de utilizar esta estructura sobre el procesamiento secuencial clásico son básicamente dos: posibilidad de aunar convolución local de filtros espaciales, realimentación, control de luminosidad y umbralización en una sola capa, y aumento drástico de la velocidad de cálculo del algoritmo secuencial.

## 1. Introducción

El problema de la detección de bordes en imágenes digitales se ha abordado en la literatura desde dos puntos de vista diferentes: técnicas orientadas al pixel, y técnicas orientadas a la región. Las técnicas de detección de bordes orientadas al pixel se basan en la detección de discontinuidades (transiciones fuertes) en la información de los píxeles de una imagen, mientras que las técnicas orientadas a la región determinan el proceso de detección mediante una medida cuantitativa de ciertas propiedades de un grupo conexo de píxeles (p. e. continuidad, energía, etc.) [1].

Las aproximaciones orientadas al pixel se clasifican en función del dominio donde se realiza la detección de discontinuidades: espacio (máscaras espaciales), frecuencia (filtros de Fourier) y espacio/frecuencia (filtros de subbanda). La transformada discreta de Fourier (DFT) determina de manera biunívoca la correspondencia entre cualquier señal digital 2D definida en el dominio del espacio (p. e. una imagen) y su espectro frecuencial [2]. Las transformadas de subbanda (p. e. Gabor, DCT, o Wavelet) tienen la propiedad de poseer una buena localización conjunta en ambos dominios: espacio y frecuencia [3].

El presente trabajo se engloba dentro de las técnicas de filtrado espacial. El documento se organiza como sigue: en la sección 2 se define el concepto de detector de borde óptimo, para especificar en la sección 3 el filtro óptimo de Canny y su algoritmo secuencial para la detección de bordes. En la sección 4 se define una estructura de

procesamiento paralelo mediante una Red Neuronal Celular (CNN) multicapa, demostrando su equivalencia con el algoritmo original.

## 2. Definición de Detector de Borde Óptimo

En Procesamiento Digital de Imágenes (PDI), un borde es la frontera entre un objeto y el fondo. Identificando este borde, se puede localizar todo el objeto, así como medir sus propiedades básicas. La utilización de la información de borde simplifica en gran medida el análisis de imágenes [1].

Los objetivos principales de un detector de borde son dos: una baja tasa de error y una buena localización del borde. El primer objetivo se refiere a la capacidad del detector para clasificar píxeles de la imagen como bordes, sin incluir elementos espurios como ruidos o manchas. Por su parte, la propia forma del filtro puede dar lugar a una detección de borde ligeramente trasladada, o incluso duplicada [1].

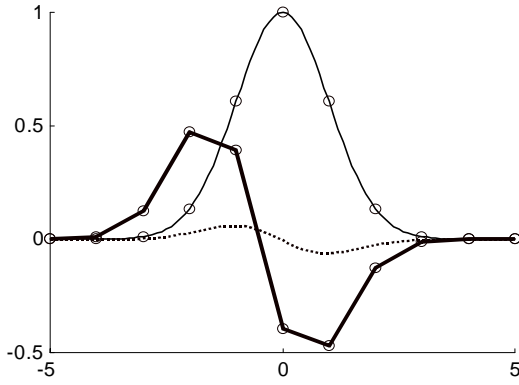
Asumiendo el caso unidimensional, un borde diferencial sujeto a ruido blanco gaussiano, y un filtro espacial  $f$  que suavice el ruido y localice correctamente los bordes, se pueden definir los siguientes criterios matemáticos:

$$SNR = \frac{A \left| \int_0^w f(x) dx \right|}{n_o \sqrt{\int_{-w}^w f^2(x) dx}} \quad (1)$$

$$L = \frac{A |f(0)|}{n_o \sqrt{\int_{-w}^w f^2(x) dx}} \quad (2)$$

$$X_r = \prod \left( \frac{\int_{-w}^w f^2(x) dx}{\int_{-w}^w (f'(x))^2 dx} \right)^{1/2} \quad (3)$$

donde  $A$  y  $n_o$  son constantes,  $SNR$  representa el cociente señal ruido a la salida del filtro  $f$ ,  $L$  es la localización del borde, inversa a la distancia entre el borde detectado y el borde real, y  $X_r$  es una restricción que asegura que la respuesta del filtro no tendrá una respuesta múltiple elevada a un mismo borde en un entorno  $W$  pequeño [4]. Un detector de borde óptimo es aquel que maximiza los criterios (1) y (2) sujetos a una restricción de respuesta múltiple (3).



**Fig. 1.** Filtro Óptimo Detector de Borde de Canny:  $g(x)$  normalizada continua y discreta para  $\sigma = 1$  (línea fina).  $dg(x)$  (línea punteada) y  $dg[x]$  (línea gruesa) a partir de  $g[x]$ .

### 3. El Método de Canny

Uno de los métodos más importantes para realizar una detección global de bordes sobre una imagen es el conocido como método de Canny [4]. Esta técnica, que se caracteriza por estar optimizada para la detección de bordes diferenciales, consta de tres etapas principales: filtrado, decisión inicial, e histéresis.

La primera etapa consiste en un filtrado de convolución de la derivada primera de una función gaussiana normalizada discreta sobre la imagen, realizada en dos direcciones: horizontal y vertical. La función gaussiana posee dos parámetros fundamentales, valor medio  $m$ , y desviación típica  $\sigma$ . En este caso, el valor medio es nulo, siendo la ecuación que define el filtro gaussiano de la forma:

$$g(x) = k \cdot \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (4)$$

donde  $k$  se elige de manera que el máximo de  $g(x)$  sea la unidad en su versión discreta. Para la realización del filtro utilizado en la primera etapa, se deriva (4), obteniéndose:

$$g'(x) = \frac{d}{dx} g(x) = -\frac{k}{\sigma^2} x \cdot \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (5)$$

La versión discreta del filtro se define como  $dg[x] = g'(x)$ , donde  $x$  posee valores desde  $-5\sigma$  hasta  $5\sigma$ , con diferencias de 1 pixel entre muestras consecutivas. Por razones de cómputo, es preferible derivar directamente la versión discreta de (4), con el fin de obtener el valor  $k$ . En la fig. 1 se muestran todas estas funciones para  $\sigma=1$ .

Canny demuestra que este filtro es óptimo en el sentido de que maximiza el producto de los criterios (1) y (2), sujetos a la restricción (3), definidos en la sección anterior.

El paso siguiente consiste en realizar la convolución de  $dg[x]$  sobre la imagen  $U$  de tamaño  $M \times N$ , con  $U(i,j) \in [0,1]$ , obteniéndose dos nuevas matrices de la forma:

$$dx = \text{abs}(dg * U), \quad dy = \text{abs}(dg^T * U) \quad (6)$$

donde el superíndice  $T$  indica una operación de trasposición, y  $*$  es el operador de convolución discreta. Las matrices  $dx$

y  $dy$  son imágenes  $M \times N$  donde el valor de nivel de gris de un pixel determina la variación o gradiente, en valor absoluto, con los pixeles colindantes, en las direcciones horizontal y vertical, respectivamente. Estas matrices son el punto de partida para la etapa siguiente.

La segunda etapa realiza una primera toma de decisiones sobre los posibles bordes de la imagen, mediante el cálculo de los picos de las imágenes obtenidas en la primera etapa,  $dx$  y  $dy$ , realizando el siguiente grupo de cálculos:

- Un pixel de  $U$ ,  $U(i,j)$ , se considera candidato a borde si  $dx(i,j) > dx(i-1,j)$  y  $dx(i,j) > dx(i+1,j)$ , o bien  $dy(i,j) > dy(i,j-1)$  y  $dy(i,j) > dy(i,j+1)$ , para  $i=1,...,M$  y  $j=1,...,N$ . Los pixeles pertenecientes al borde de la imagen no se contabilizan en esta decisión, y se clasifican como pixeles de fondo.
- Para los pixeles  $(i,j)$  que cumplan la condición anterior, se calcula el valor  $e(i,j) = (dx^2(i,j) + dy^2(i,j))^{1/2}$ . Para el resto de pixeles,  $e(i,j) = 0$ . Para eliminar efectos de borde, se eliminan los valores de las filas 2,  $M-2$  y  $M-1$ , así como los de las columnas 2,  $N-2$  y  $N-1$ . La matriz  $M \times N$  así obtenida, posee los candidatos de borde global (no orientado) de la imagen  $U$ .
- Para obtener una imagen con valores de gris en el rango  $[0,1]$ , se normaliza la imagen  $e$ .

La última etapa de procesamiento realiza una optimización de la decisión llevada a cabo en la etapa anterior, mediante la aplicación de una función de histéresis sobre la imagen  $e$ . Esta función se basa en la definición de dos umbrales,  $T_L$  y  $T_H$ , tales que  $T_L < T_H$ . Valores típicos para estos umbrales son 0.1 y 0.5, respectivamente, aunque se recomienda que  $T_H$  y  $T_L$  tengan una relación entre 2:1 y 3:1, dependiendo de la relación señal ruido (1), en el caso de que este valor sea conocido [4]. Esta etapa realiza los cálculos siguientes:

- Un pixel de  $e$ ,  $e(i,j)$ , se considera borde definitivo si  $e(i,j) \geq T_H$ .
- Un pixel de  $e$ ,  $e(i,j)$ , se considera fondo definitivo si  $e(i,j) < T_L$ .
- Todos los pixeles en un vecindario  $3 \times 3$  de los pixeles considerados como borde definitivo,  $e(k,l)$ , se consideran también borde definitivo si  $e(k,l) \geq T_L$ .

La información normalizada no binaria de salida del método de Canny indica la fuerza del borde a partir de un análisis del gradiente global entre pixeles vecinos. Este método es uno de los más robustos contra el ruido (filtro óptimo), en comparación con los métodos de Roberts, Sobel o Prewitt, por ejemplo [1].

### 4. Método de Canny sobre una Red Neuronal Celular Multicapa

El objetivo de esta sección es definir una estructura de procesamiento paralelo que realice el algoritmo de Canny. Generalmente, la velocidad de cálculo de cualquier algoritmo secuencial depende del tamaño de la imagen de entrada. Las estructuras de procesamiento paralelo son una buena alternativa para eliminar este tipo de dependencia a la hora de realizar tareas complejas de PDI.

Una de las estructuras de procesamiento paralelo más importantes es el modelo de Red Neuronal Celular (CNN)

[5][6], cuya principal característica es el conexionado local entre neuronas adyacentes. Esta característica ha sido aprovechada para la definición de estructuras capaces de realizar tareas típicas de PDI, como la extracción de bordes y líneas y detección de movimiento, entre otras [6][7][8]. Entre las diversas implementaciones reales de este tipo de redes cabe destacar las realizadas en [9][10].

El modelo utilizado en toda la estructura corresponde al definido en [8][11], y que puede expresarse de la forma:

$$X(t+1) = A * F[X(t)] + B * U + I \quad (7)$$

donde  $y=F(x)$  es una función de salida de la actividad neuronal,  $A$  es la plantilla de realimentación,  $B$  es la plantilla de control,  $U$  es la imagen de entrada e  $I$  es una constante de polarización y  $X$  la variable de estados, siendo  $*$  un operador de convolución, de la forma:

$$[A * X](i, j) = \sum_{C(k,l) \in N_r(i,j)} A(k, l) \cdot X(i, j) \quad (8)$$

$$N_r(i, j) = \{C(k, l) | \max\{|k - i|, |l - j|\} \leq r\} \quad (9)$$

donde  $N_r$  establece el vecindario o límite de conexiones laterales,  $r = 1, 2, \dots$ . En [12] se establece una nueva función sináptica para realizar la operación de convolución, la cual permite realizar operaciones de orden superior, además de funciones lógicas. La forma de este operador es:

$$\mathbf{x} \otimes_{\alpha} \mathbf{w} = \frac{1}{2^{\alpha}} \sum_{i=1}^N [\text{abs}_{\alpha}(x_i + w_i) - \text{abs}_{\alpha}(x_i - w_i)] \quad (10)$$

donde  $\mathbf{x}$ ,  $\mathbf{w}$  son dos vectores de longitud  $N$ ,  $|w_i|, |x_i| < 1$  y donde la función  $\text{abs}_{\alpha}$  se define según la relación:

$$\text{abs}_{\alpha}(x) = (x^2 + \varepsilon^2)^{\alpha/2} \quad (11)$$

donde  $\alpha = [1, \infty)$ , siendo  $\varepsilon$  un valor real positivo pequeño, menor que 1. Cuando  $N = 1$ , se tiene para  $\alpha = 1$  y  $\alpha = 2$ :

$$\begin{aligned} x \otimes_1 w &= \text{sgn}(x) \cdot \text{sgn}(w) \cdot \min[|x|, |w|] \\ &= \frac{1}{2} \cdot [|x + w| - |x - w|] \end{aligned} \quad (12)$$

$$x \otimes_2 w = x \cdot w \quad (13)$$

Puede demostrarse que ambos operadores son similares funcionalmente al operador (8), pudiendo aplicarse en su lugar para realizar operaciones de orden superior [12]. En [13] se define el operador  $\text{abs}(x)$  a partir de (10):

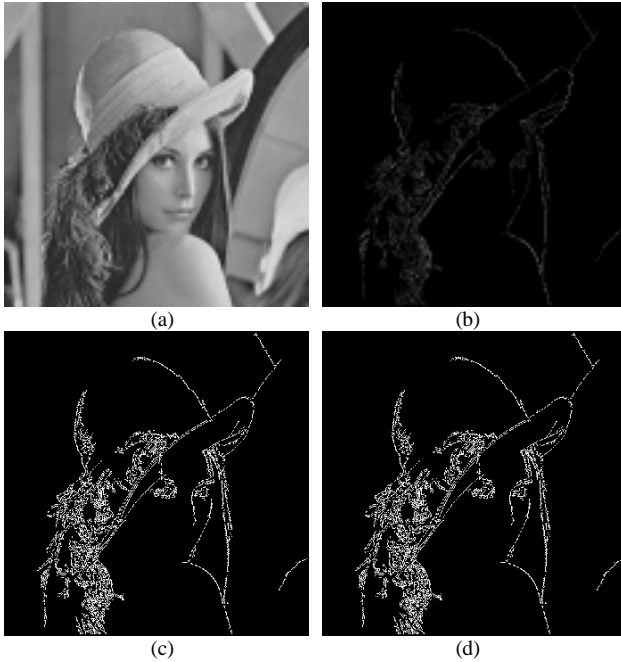
$$\mathbf{x} \otimes_{\alpha} = \sum_{i=1}^N [\text{abs}_{\alpha}(x_i)], \quad \text{abs}(x) = x \otimes_1 \quad (14)$$

Nótese además que (12) define la operación  $\text{AND}(x, w)$ . La red utilizada en la primera etapa incluye el filtro óptimo  $dg[x]$  definido en la sección anterior como valor de control,  $B$ . Según la orientación de este filtro, y colocando a la salida de cada neurona una función  $\text{abs}(x)$  (14) con  $N=1$ , se obtienen mediante (6) las matrices  $dx$  y  $dy$ .

La siguiente etapa se realiza, en su primera parte, mediante el uso de 4 CNN de comparación, con un filtro de control  $B=[-1, 1]$  y  $B=[1, -1]$  para  $dx$ , y sus correspondientes traspuestas para  $dy$ . La salida de esta red es una función escalón con límites  $[0, 1]$ , centrada en el origen. Las salidas de las dos redes de comparación para  $dx$  y las dos para  $dy$ , son las entradas de un operador lógico  $\text{AND}(x, y)$  (12). A la salida de las redes AND se introduce una red  $\text{OR}(x, y)$  ( $\text{AND}(1-x, 1-y)$ ). La salida de esta red es una máscara binaria de los píxeles candidatos a borde.

La matriz  $e(i, j)$  se construye aplicando la máscara binaria obtenida por OR sobre  $dx$  y  $dy$  por separado: esto se lleva a cabo con sendas CNN con  $B$  definida según la salida de OR y aplicada sobre  $dx$  y  $dy$ . Además, se realiza un operador cuadrático mediante la función (11), con  $\alpha = 1$  y  $\varepsilon = 0$ . Estas dos salidas son las entradas de una última red, que realiza el cálculo del valor cuadrático medio de las dos salidas anteriores. Esto puede realizarse alimentando la entrada  $U$  con la máscara cuadrática de  $dx$ , y la polarización  $I$  con la máscara cuadrática de  $dy$ . Sobre estos valores se aplica una función sináptica (11) con  $\alpha = 1/2$  y  $\varepsilon = 0$ , lo que proporciona la operación raíz cuadrada. Esta matriz puede normalizarse fácilmente calculando el máximo valor de la salida, y dividiendo cada valor por ese máximo.

Finalmente, el bloque de histéresis se implementa con dos CNN en cascada. La entrada de la primera CNN es la salida del bloque anterior,  $e(i, j)$ . El operador de control es unitario,  $B = 1$ , siendo  $A = 0$ , e  $I = 0$  en la primera iteración. La salida de esta red es una función escalón  $[0, 1]$ , donde el origen se establece igual a  $T_H$ . En la primera iteración, esta red proporciona los candidatos fijos a borde. La salida de esta red es la entrada de una segunda CNN, con  $B(i, j) = 1$ ,  $\forall i, j \in [1, 3]$ , excepto  $B(2, 2) = 0$ ;  $A = 0$  e  $I = 0$ , cuya función de salida es una función escalón  $[0, T_H - T_L]$  centrada en el origen. La salida de esta red es la nueva polarización de la primera red, calculando los píxeles de borde vecinos de los



**Fig. 2.** Detector de borde de Canny. (a) Imagen Lenna 512 x 512 píxeles. (b) Salida del detector (algoritmo).  $\sigma = 1$ ,  $T_H = 0.5$ ,  $T_L = 0.1$ . (c) Como (b), binarizada. (d) salida de la CNN multicapa de Canny (binarizada). Hay una correlación = 1 entre (c) y (d).

definitivos cuyo valor sea mayor que  $T_L$ . El proceso converge cuando ya no hay más píxeles a añadir. La salida binaria final de esta red puede multiplicarse sobre  $e(i,j)$ , para obtener la información de borde en escala de grises, estableciéndose una equivalencia total con el algoritmo secuencial, como puede comprobarse en la fig. 2, donde se realiza una prueba del algoritmo y la estructura con la imagen Lenna como entrada. En la fig. 3 se muestra un esquema de la estructura CNN descrita en esta sección.

## 5. Conclusiones

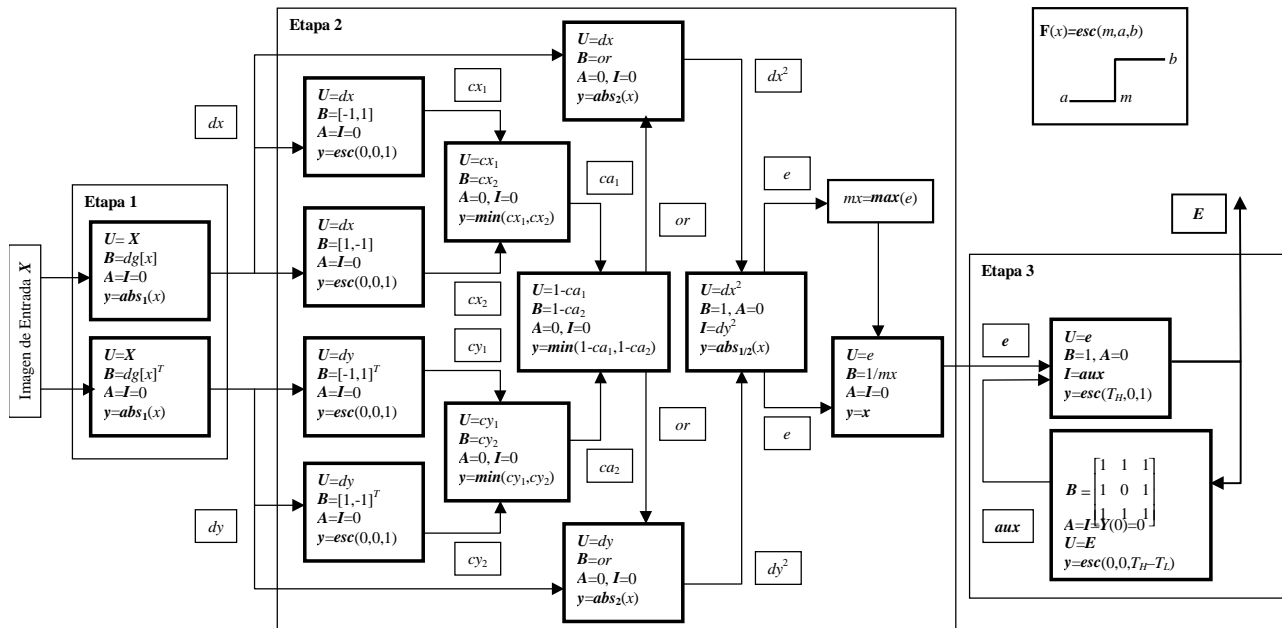
Se ha definido una estructura de procesamiento paralelo que realiza uno de los filtrados de detección de borde más robustos frente al ruido existentes en la literatura: el método de Canny. La aplicación de funciones sinápticas de orden superior para el modelo utilizado permite implementar el algoritmo completo sobre CNN sencillas, logrando una total uniformidad en la estructura, que facilita su implementación en circuito VLSI, y una réplica perfecta del algoritmo secuencial original.

## Agradecimientos

Este trabajo ha sido financiado con el proyecto de investigación IPR98A016 concedido por la Junta de Extremadura.

## Referencias

- [1] R.C. González, R.E. Woods, *Digital Image Processing*, Addison-Wesley, 1993.
- [2] E.O. Brigham, *The Fast Fourier Transform*, Prentice Hall, 1974.
- [3] E.P. Simoncelli, E.H. Adelson, "Subband Transforms", *MIT Media Lab Vis. Mod. Tech. Rep.* 137, 1990, pp. 143-193.
- [4] J.F. Canny, "A Computational Approach to Edge Detection", *IEEE PAMI*, 8(6), 1986, pp. 679-698.
- [5] L.O. Chua, L. Yang, "Cellular Neural Networks: Theory", *IEEE Trans. Circ. Syst.*, 35(10), 1988, pp. 1257-1272.
- [6] L.O. Chua, L. Yang, "Cellular Neural Networks: Applications", *IEEE Trans. Circ. Syst.*, 35(10), 1988, pp. 1273-1290.
- [7] L.O. Chua, T. Roska, "The CNN Paradigm", *IEEE Trans. Circ. Syst. I Fund. Theory Appl.*, 40(3), 1993, pp. 147-156.
- [8] M.A. Jaramillo, F.J. López, M. Macías, M.I. Acevedo, "A CNN Model for Grey Scale Image Processing", *IWANN 95 Proc., Torremolinos (Málaga)*, LNCS 930, 1995, pp. 882-889.
- [9] H. Harrer, J.A. Nossek, R. Stelzl, "An Analog Implementation of Discrete-Time Cellular Neural Networks", *IEEE Trans. Neural Networks*, 3(3), 1992, pp. 466-476.
- [10] T. Roska, L.O. Chua, "The CNN Universal Machine: Analogic Array Computer", *IEEE Trans. Circ. Syst. I: Fund. Theory Appl.*, 40(3), 1993, pp. 163-173.
- [11] M.A. Jaramillo, J.A. Fernández, "Applications of Cellular Neural Networks (CNN) to Grey Scale Image Filtering", *ICANN 99 Proc., Edinburgh, IEE Conf. Pub.* 470(I), 1999, pp. 449-454.
- [12] R. Dogaru, K.R. Crounse, L.O. Chua, "An Extended Class of Synaptic Operators with Applications for Efficient VLSI Implementation of Cellular Neural Networks", *IEEE Trans. Circ. Syst.*, 45(7), 1998, pp. 745-755.
- [13] M.A. Jaramillo, J.A. Fernández, "Adaptive Adjustment of the CNN Output Function to Obtain Contrast Enhancement", *IWANN 99 Proc., Alicante*, LNCS 1607(II), 1999, pp. 412-421.



**Fig. 3.** Estructura multicapa para la implementación del detector de bordes de Canny en 3 etapas. Cada recuadro representa una CNN, en el que se definen los parámetros principales de cada subred ( $U$ : entrada,  $B$ : control,  $A$ : realimentación,  $I$ : polarización,  $y$ : función de salida). La salida es  $E$ , una matriz binaria del mismo tamaño que la entrada al sistema  $X$ , con los bordes seleccionados a 1.