

Reconocimiento de patrones aplicado a la exposición fotográfica

M. en C. Enrique Alejandro Navarrete Paredes (enavarre@campus.cem.itesm.mx)

Dr. Eduardo de Jesús García García (egarcia@campus.cem.itesm.mx)

Dr. Raúl Monroy Borja (raulm@campus.cem.itesm.mx)

I.T.E.S.M.-C.E.M.

Carretera Lago de Gpe. Km 3.5, Atizapán de Zaragoza, 52926 Edo. de Mex., México.

Tel. (52) 5864-5657 FAX (52) 5864-5651

Abril de 2000

Abstract: La exposición fotográfica es un problema complejo que aún las cámaras fotográficas más sofisticadas (sistemas de zonas con forma pre-definida) no resuelven satisfactoriamente en todos los casos. En esta tesis se desarrolla un sistema capaz de calcular la exposición fotográfica para escenas de tipo "paisaje" extrayendo las características de la imagen para alimentar una red neuronal. La extracción de los datos característicos de la imagen se hace mediante una segmentación de la misma, para obtener regiones con la forma de los elementos que componen la escena. Se hacen conjuntos de patrones (validación y entrenamiento) en función de la entropía de cada imagen. Estos conjuntos se usan para el proceso de entrenamiento y validación de la red neuronal. Se emplea el método de momentum y eliminación de punto plano".

Palabras Clave: Fotografía, Exposición Fotográfica, Reconocimiento de Patrones, Segmentación,

1.0 Introducción

Los avances tecnológicos han permitido al hombre simplificar diversas tareas de su actividad cotidiana. Tal es el caso de la fotografía, donde el proceso de plasmar imágenes en papel se ha ido mejorando y facilitando cada vez más. Las cámaras modernas dan más opciones de composición y apoyan al fotógrafo en la exposición de la película fotográfica. Sin embargo, la determinación automática de la exposición no es confiable en el 100% de los casos.

El problema de la exposición fotográfica consiste en encontrar la combinación de velocidad de obturación y apertura de diafragma para exponer adecuadamente la película fotográfica. La emulsión fotográfica de la película tiene un rango de sensibilidad a la cantidad de luz que incide sobre ella. Si recibe una cantidad de luz mayor al límite superior del rango de sensibilidad, se dice que la película queda "sobreexpuesta"; en el caso contrario, la película queda "subexpuesta". En ambas situaciones el evento no queda registrado adecuadamente.

Las cámaras fotográficas más modernas cuentan con sistemas de medición de luz de multizona, donde la escena se divide en zonas y dependiendo de la programación que hace el usuario se calcula la exposición fotográfica. La programación del usuario indica al sistema el tipo de escena que se va a tomar para que haga las correcciones de exposición pertinentes. Sin embargo, las pruebas hechas a estos sistemas indican que una cámara de este tipo no es capaz de identificar cuál es la mejor exposición. Los métodos empleados hasta el momento no aseguran el éxito de exposición en su totalidad. Por lo tanto es necesario seguir trabajando para la mejora de estos sistemas.

En el presente artículo se explica el desarrollo de un sistema de exposición que extrae la información de imágenes previamente digitalizadas, mediante técnicas de procesamiento digital de señales. Esta información alimenta una red neuronal que determina cual es la exposición que debe darse a una escena. Parte importante del procesamiento de la imagen es la segmentación, que permite dividirla en regiones acordes a la forma de los elementos que componen la escena.

2.0 Fotografía

La fotografía es una actividad que el hombre ha realizado hace más de cien años, buscando immortalizar eventos significativos dentro de su existencia o expresar sus sentimientos a través de este medio. Para dejar claro que es la fotografía se define enseguida: “La fotografía es el arte de fijar y reproducir, mediante reacciones químicas, en superficies convenientemente preparadas, las imágenes

2.1 Exposición fotográfica

La exposición fotográfica es el proceso de regulación del tiempo y cantidad de luz que debe recibir la película fotográfica. Los elementos que intervienen en este proceso son la velocidad de obturación, la apertura de diafragma, la sensibilidad de la película y la luminancia de la escena.

2.2 Clasificación de escenas y sistemas de exposición fotográfica

Las escenas comunes presentan patrones típicos en función del tema que se va a fotografiar. Una clasificación basada en la experiencia de fotógrafos[2] es: retrato, paisaje, periodismo, desnudo, niños, naturaleza, deportes y acción, arquitectura y naturaleza muerta.

se ha hecho en función del motivo de cada una de las escenas y no en función de la luminancia que se tiene.

En la composición de una fotografía se consideran el motivo de la escena y la luminancia de la misma. El motivo de la escena definirá si se requiere una profundidad de campo amplio o reducido y una velocidad de obturación alta o baja. Sin embargo, la luminancia es el principal factor para determinar la exposición. La luminancia es la variable independiente del proceso de exposición fotográfica.

Las cámaras fotográficas más sofisticadas poseen sistemas de medición de luz que automáticamente modifican la velocidad de obturación y la apertura del diafragma. Son sistemas de medición de luz multizona, donde la escena se divide en regiones y se considera el tipo de escena que se va a tomar; previamente el usuario indica el tipo de la escena. La forma de división de la escena y el número de secciones está en relación con la marca de la cámara. Existen sistemas que pueden dividir la escena en desde 8 zonas (Figura 1 y Figura 2, Pentax y Nikon respectivamente) hasta las que emplean 16 zonas (Figura 3, Minolta).

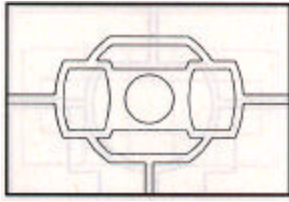


Figura 1. Pentax PZ-1

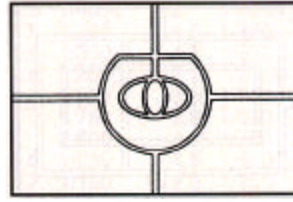


Figura 2. Nikon N90

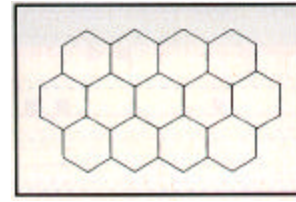


Figura 3. Minolta 9xi

Una prueba hecha por la revista Popular Photography [3], hace una comparación entre 17 cámaras SLR (Single Lens Reflex) de las marcas: Canon, Minolta, Nikon, Pentax y Sigma, mostrando los resultados de los sistemas de medición en 7 escenas diferentes. Los resultados que da el experimento de fotografiar la misma escena, con las mismas condiciones de luz, varían según la cámara que se emplee y el modo de operación escogido. Estas pruebas indican que una cámara moderna no es capaz de identificar cuál es la mejor exposición. Por lo tanto se puede afirmar que es necesario seguir trabajando para la mejora de los mismos.

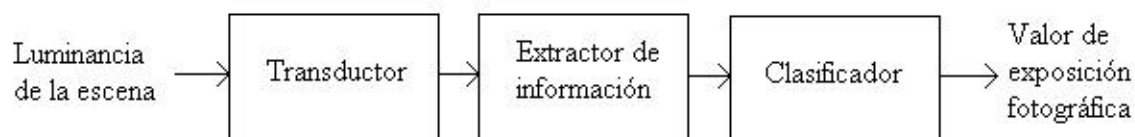
2.3 Guía para la exposición fotográfica

La norma American National Standard for Photography – Photographic Exposure Guide [4], tiene como propósito proveer información sobre la iluminación producida por el sol, la luna y fuentes artificiales, para encontrar la exposición adecuada de la película fotográfica. En otras palabras, la determinación de la velocidad de obturación y apertura de diafragma correspondientes a la situación particular.

El estándar indica paso a paso los valores de exposición y de corrección correspondientes. Se deben considerar la latitud a la que se encuentra la escena, el mes y la hora del día, el ángulo entre la cámara y el sol, las condiciones del cielo, el tono del sujeto, el tono del entorno y la bruma área.

3.0 Sistema de reconocimiento de patrones para determinar la exposición

El objetivo del presente trabajo es desarrollar un sistema de cómputo capaz de determinar la exposición fotográfica correcta de una escena, conjuntando técnicas de procesamiento de imágenes, reconocimiento de patrones y la experiencia de fotógrafos. Un sistema que puede hacer la tarea descrita se muestra en el diagrama a bloques siguiente:



3.1 Extracción de la información

La función de la segunda es extraer la información relevante de la imagen como son los segmentos que conforman la escena, el área de cada uno de ellos y su luminancia asociada. Los pasos que se siguen en el proceso son: normalización, filtrado y segmentación y conteo de píxeles y promedio de luminancia [6] [7] [8] [9].

3.2 Normalización

La normalización fija la luminancia de todas las imágenes analizadas a un nivel estándar. Así ya no es necesario ajustar las etapas siguientes del procesamiento una vez que están sintonizadas. Este proceso consiste en escalar cada píxel de la imagen al máximo nivel posible en función del píxel de mayor luminancia. Cada píxel se multiplica por el factor siguiente:

$$\text{escala} = \text{valor máximo de luminancia} / \text{píxel de mayor luminancia} \quad (\text{Ec. 1}).$$



Figura 5. Imagen original



Figura 6. Imagen normalizada

3.3 Filtrado

Este proceso permite suavizar los componentes de la imagen, de manera que se enfatizan los cambios de contraste más relevantes. En nuestro caso se emplea la mediana por ser el filtro que preserva mejor los contornos de la imagen. La mediana de un conjunto de valores es aquel número que se encuentra a la mitad de éstos cuando el conjunto está ordenado. En otras palabras, arriba de la mediana se tendrá el 50% de los números del conjunto, y abajo de ella el otro 50%.



Se observó que para este trabajo, el empleo de una mediana de 3x3 resultaba más conveniente. Primero, porque el proceso de ordenar los píxeles es costoso en tiempo; para un recuadro de 3x3 se necesitan ordenar 9 números; para uno de 5x5, 25 valores. Segundo, al aplicar una mediana de un orden mayor a 3x3 se percibe la pérdida de

Figura 7. Mediana 3x3

para detectar los cambios bruscos de luminancia. Se define de la forma siguiente:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (\text{Ec. 2})$$

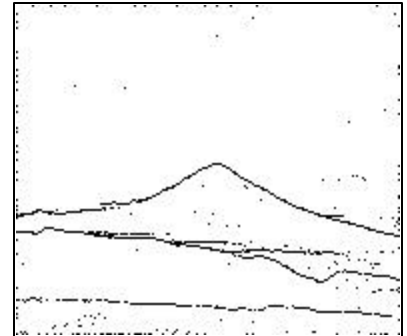
3.3.2 Laplaciano discreto

Para efectos de aplicar el filtro a una imagen digitalizada, se requiere de la versión discreta del Laplaciano[11]. Empleando diferencias finitas se puede obtener el gradiente de la función $f(x,y)$. Si se buscan los puntos donde hay un cambio de contraste, el gradiente debe ser igual a cero, por lo que la máscara digital queda como:

$$\nabla^2 f(x,y) = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (\text{Ec. 3})$$

Para aplicar esta función discreta a una imagen, se requiere de la convolución. Al aplicar el Laplaciano a las imágenes se tiene como resultado un conjunto de puntos que indican donde se produce un punto de inflexión, es decir, donde se ubica un cambio brusco de contraste. Un ejemplo de este proceso se muestra la Figura 8.

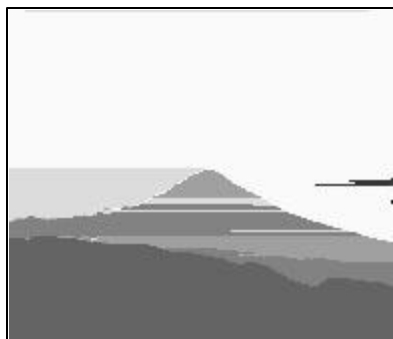
Figura 8.
Puntos de
inflexión



3.3.3 Segmentación

La fase final emplea la matriz de puntos de inflexión como referencia para construir cada uno de los segmentos que constituye la escena. Este método es una adaptación del propuesto por Castleman para almacenar información de los segmentos, "Codificación de segmento de línea"[6]. El algoritmo de segmentación usa cada uno de los puntos de inflexión como indicación del término de un segmento y el inicio de uno nuevo. El análisis se hace línea por línea y conforme se encuentra un punto de inflexión se genera una nueva estructura dinámica donde se guarda la coordenada inicial y final de la fracción de línea. Debido a la combinación de los elementos contenidos en las imágenes, no es posible obtener todos los puntos de inflexión requeridos para encontrar los contornos cerrados. Por tanto la segmentación basada únicamente en los puntos de inflexión no es completa.

Ahora se recurre al análisis de los niveles promedio de luminancia de los segmentos. Para cada segmento se encuentra el valor medio de luminancia y la desviación estándar. A continuación se comparan los segmentos y se fusionan aquellos segmentos que tienen una luminancia en un rango de 2.5 desviaciones estándar. La comparación se hace de cada segmento con el resto de los segmentos.



La imagen de la Figura 9 es el resultado del proceso ya descrito. Con el objeto de mostrar con mayor claridad esta segmentación, se han puesto niveles de grises falsos para distinguir las diferentes regiones. Con esta nueva segmentación se recalcula la luminancia del segmento promediando los niveles de luminancia de los pixeles que lo componen. También se hace un conteo de los pixeles que tiene el segmento para encontrar su área.

Figura 9. Segmentación

3.4 Clasificador

Partiendo del hecho de que se cuenta con el estándar “Guía para la exposición fotográfica”[4], que es la reunión del conocimiento y experiencia de un equipo de fotógrafos-investigadores, se puede determinar la exposición fotográfica para cada escena. Este valor de exposición (Ev) fotográfica y la luminancia-área de los segmentos obtenida en los pasos anteriores es la información que se debe relacionar a través de un sistema de reconocimiento de patrones.

3.5 Luminancia-Área de los segmentos

En este trabajo definimos que el patrón característico de una imagen está compuesto por cinco valores, cada uno de ellos corresponde a la probabilidad de que se presente un segmento con una área y una

El número de valores del patrón se fijó considerando dos aspectos, el grado de complejidad que se genera al tener un vector de mayor dimensión y el tipo de escenas que se analizan (tipo paisaje) que presentan grandes regiones con una misma luminancia, por lo que se decide fijar el número a cinco. La relación empleada para el cálculo del valor del segmento es:

$$\text{Valor} = (\text{Luminancia media del segmento} * \text{Área del segmento}) / (\text{Ancho de la imagen} * \text{Alto de la imagen})$$

(Ec. 4)

Es importante mencionar que se emplea el "Valor de luminancia máximo de la imagen" en el cociente de la relación (Ec. 5) para uniformizar los grados de luminancia que tiene cada imagen. De esta forma

3.6 Red neuronal con entrenamiento Propagación hacia atrás con momentum y

Uno de los sistemas capaces de clasificar un patrón entre varias categorías, es la red neuronal. Una red neuronal tiene la posibilidad de generalizar su comportamiento, dando buenos resultados con patrones con los que no fue entrenada. De tal forma que la red se puede entrenar con un grupo finito de patrones, para poder discernir entre un número infinito de combinaciones de segmentos y niveles de luminancia que se presentan en la naturaleza.

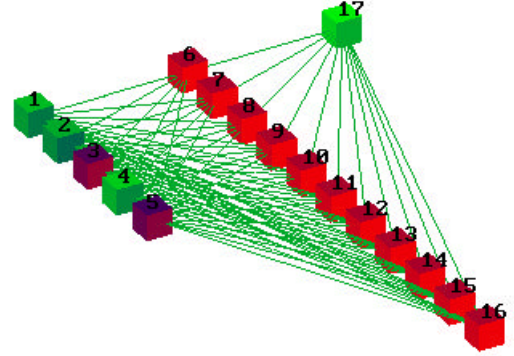


Figura 10. Estructura de la red neuronal

La red neuronal propuesta posee una capa de entrada constituida por 5 neuronas -correspondientes a los cinco valores que forman un patrón -, una capa oculta de 11 neuronas y una capa de salida compuesta de 1 neurona. Hetch-Nielsen[12] comenta que una red neuronal de tres capas, con un entrenamiento de propagación hacia atrás, puede ser capaz de aproximar cualquier función; considerando que la capa de entrada tenga n neuronas, la capa oculta $2n+1$ neuronas y la capa de salida m neuronas.

El entrenamiento de la red neuronal consiste en presentar a su entrada el patrón compuesto por los 5 valores de luminancia-área. A la salida la red arrojará un valor de exposición que será comparado con el valor real y se producirá un error. Éste último se usará para modificar las características de las neuronas que componen la red. Específicamente se empleó la metodología de aprendizaje denominada

3.7 Entropía

Parte importante del proceso de entrenamiento es determinar cuales son los patrones para hacer esta tarea. En nuestro caso se decide considerar el grado de entropía de cada imagen para detectar el nivel de incertidumbre de su estructura[10].

Matemáticamente la entropía se define como

$$U = \sum -p_i \log_2 p_i \quad (\text{Ec. 5})$$

Para nuestra aplicación la redefinimos como sigue

$$U = \sum -t_i \log_2 t_i \quad (\text{Ec. 6})$$

Donde $t_i = r_i l_i$, r_i es el área relativa del segmento i (el área total de la imagen es de 192x165 pixeles, 31680 pixeles), l_i es la luminancia relativa del segmento i (la luminancia máxima considerada es el mayor nivel encontrado en la imagen).

4.0 Resultados

En total se contó con un conjunto de 75 imágenes, de donde se obtuvieron la muestra de entrenamiento

Se observó que el mayor valor de exposición que se tiene en el grupo de 75 imágenes, es de 15. Al normalizar el valor de exposición se obtiene el "valor de exposición relativo".

4.1 Imágenes de entrenamiento

El conjunto de patrones de entrenamiento se decidió en función del grado de entropía que tuvo cada imagen. Se escogieron 10 imágenes que tuvieron una entropía en un rango de 0.4 a 1.8, de manera que se mantuviera una representación adecuada del universo de imágenes.

4.2 Imágenes de validación

De la misma forma que para las imágenes de entrenamiento, se tomaron las mismas consideraciones para escoger las imágenes de validación. Los valores de exposición de un conjunto de 10 imágenes cuya entropía se encuentra en un rango de 0.6 a 1.63.

4.3 Resultado del entrenamiento de la red neuronal y validación

Para construir y entrenar la red neuronal se empleó el Simulador de Redes Neuronales de la Universidad de Stuttgart[13]. El empleo del método de entrenamiento de "propagación hacia atrás mejorado" (en el simulador: "Backpropagation with momentum term and flat spot elimination") mostró un mejor resultado que empleando el método de "propagación hacia atrás estándar".

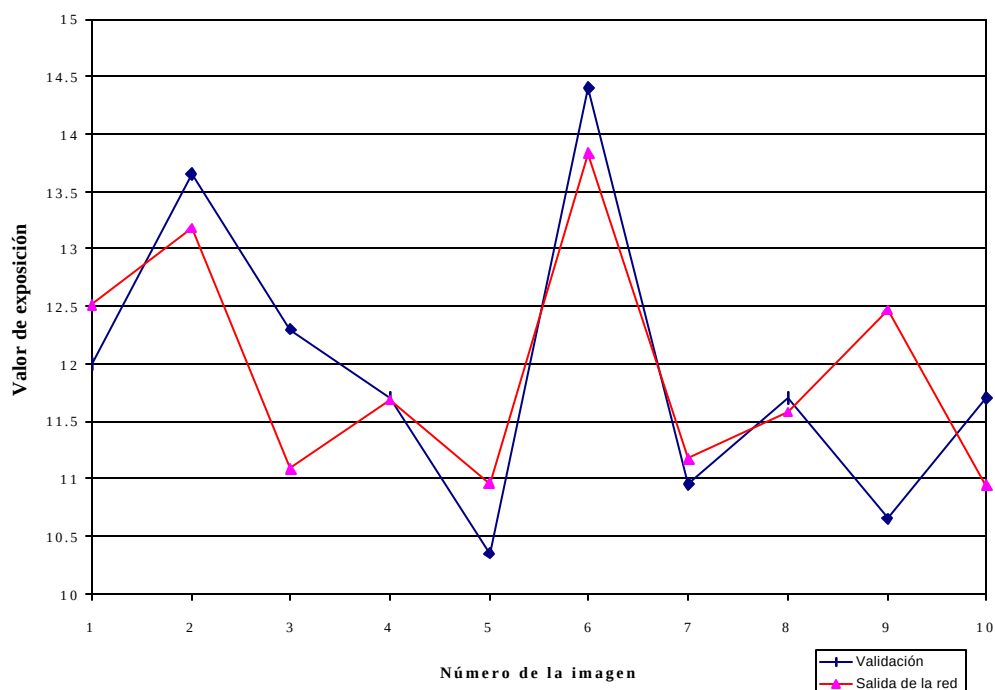


Figura 11. Gráfica de validación. Aprendizaje = 0.74, momentum = 0.87, $c = 0.22$, error = 0.0

Los resultados mostrados con la línea azul con marcas de rombo corresponden a la salida de la red neuronal, mientras que los valores de exposición unidos por la línea naranja y marcas de triángulo

4.4 Prueba del sistema

El funcionamiento del sistema se probó en una situación real empleando la cámara digital "Color QuickCam" de la firma Connectix para obtener la digitalización de la escena. Los parámetros fotográficos se hizo empleando la siguiente relación matemática [14].

$$T = \frac{A^2}{2^{Ev}} \quad (\text{Ec. 8})$$

Donde A es la apertura del diafragma, T la velocidad de obturación y Ev el valor de exposición. La escena de tipo paisaje emplea una gran profundidad de campo, por tanto se fijó el parámetro A con un valor de 16 y se sustituyó en la relación (Ec. 8) los valores de exposición obtenidos por el sistema. La operación da el resultado mostrado en la columna "Velocidad de obturación calculada". La "Velocidad de obturación empleada" se fijó al valor que la cámara fotográfica puede usar. Enseguida se muestra un ejemplo gráfico de las 6 pruebas realizadas, en este caso se obtuvo un $Ev = 10.4$.



Figura 12. Imagen obtenida con la cámara digital



Figura 13. Segmentación



Figura 14. Fotografía obtenida con los parámetros indicados por el sistema desarrollado

5.0 Conclusiones

5.1 Fotografía

Es imprescindible el empleo de estándares; en nuestro caso los estándares ANSI PH2.7-1986 "Guía para la exposición fotográfica" y ANSI PH3.49-1971 "Exposímetros fotográficos de propósito general" han sido fundamentales. Con el primero se puede determinar la exposición en función de las condiciones que rodean a la escena (posición del sol, latitud, altitud, hora, etc.). Con el segundo se tiene la ecuación matemática que relaciona las variables fotográficas, apertura de diafragma, velocidad de obturación, sensibilidad de la película y luminancia. Gracias a estos estándares fue posible determinar la exposición fotográfica de las 75 imágenes empleadas; el cálculo del rango de luminancia al cual una película fotográfica funciona; y el cálculo del rango de luminancia para las escenas de tipo paisaje. En especial los dos últimos cálculos dejan en claro que la película fotográfica tiene limitaciones con respecto al grado de luminancia que se puede presentar en una escena a plena luz del día.

5.2 Reconocimiento de patrones aplicado a la exposición fotográfica

Las conclusiones se pueden agrupar de acuerdo a los bloques que componen el sistema, extracción de

5.2.1 Extracción de información

Experimentalmente se comprueba que el empleo del filtro Mediana-Laplaciano se obtiene una abstracción más real y poderosa de la imagen, que contribuye positivamente para la determinación de la

La segmentación permite analizar la luminancia y tamaño de cada segmento de la imagen. La relación que se hace de estos datos es vital para formar el patrón que caracteriza a la imagen. Es claro que la probabilidad de que se presente un segmento con un área y una luminancia específicas es una relación adecuada para representar estas características en un solo dato.

5.2.2 Clasificador

Una red neuronal es capaz de determinar la exposición fotográfica, considerando los segmentos más importantes - aquellos con mayor área - de la imagen y la luminancia asociada a cada uno. En especial queda comprobado experimentalmente que la estructura propuesta para la red neuronal es capaz de aproximar el proceso heurístico que sigue el hombre para determinar la exposición fotográfica.

La entropía como medida del desorden que se puede tener en una imagen, fue un buen medio para seleccionar las imágenes que formaron el conjunto de entrenamiento y validación. Debe considerarse que si bien la estructura de la red teóricamente puede aproximar cualquier función real, el entrenamiento que se hace es condición fundamental para que se llegue a este punto de aproximación.

El Simulador de Redes Neuronales ha sido un instrumento invaluable. Sin este simulador no se hubiera podido llegar a los resultados que se presentan en este trabajo. Es digno de admirar el trabajo que han hecho los alemanes durante varios años para agrupar los diferentes programas que componen al simulador.

Analizando la gráfica de la Figura 11 se puede ver que el resultado obtenido de la red neuronal al usarse el conjunto de patrones de validación, se presenta una gran desviación en las imágenes 1 y 9. Estos datos nos indican que es probable que la exposición calculada empleando el estándar no se haya hecho de la manera adecuada. Esto se infiere del resto de los resultados de la gráfica, donde se puede ver claramente que el comportamiento de la red sigue de manera cercana la exposición real.

5.3 Perspectivas

El sistema de exposición fotográfica se debe extender a todos los tipos de escena. Se propone explorar las tres posibilidades siguientes. Primera, entrenar y validar la red propuesta en este trabajo con patrones de todos los tipos de escenas. Segunda, definir una red neuronal por tipo de escena y formar con estas redes una nueva red neuronal, la activación de la red neuronal correspondiente la haría el usuario. Tercera, aplicar el concepto de mapas auto-organizados para clasificar las escenas en función de su contenido y no en función de la percepción subjetiva del hombre; generar una red neuronal del tipo propuesto en la segunda posibilidad.

Para fines prácticos de traslado y consistencia en la información obtenida debe desarrollarse tanto el hardware para la adquisición de la información, como la implementación en circuitos integrados de la red neuronal. Este nuevo desarrollo puede considerar los niveles de color y el tipo de película empleada (luz de día, tungsteno, infrarrojo, etc.) para determinar si se requiere de un filtro en particular, para las compensaciones correspondientes.

Dado que el proceso de impresión de imágenes (del negativo al papel) es similar al de la exposición de la película fotográfica, este sistema se puede adaptar para el proceso de positivado.

6.0 Bibliografía

- [1] *Nuevo diccionario ilustrado Sopena de la lengua española*. Ramón Sopena S.A. 1972
- [2] Garret, John. *El arte de la fotografía en blanco y negro*. Herman Blume Ediciones TR164.G62218 1991
- [3] *Popular Photography*. Volume 58, No.5 May 1994. Pags. 40-51, 116 y 142
- [4] ANSI PH2.7-1986 *American National Standard for Photography-Photographic Exposure Guide*
- [5] *Area Array Image Sensor Products*. Data book. Texas Instruments.1994.

- [6] Castleman, Kenneth R. ***Digital image processing***. Prentice Hall. 1996.
- [7] Niblack, Wayne. ***An Introduction to Digital Image Processing***. Prentice/Hall International. 1986. TA1632.N53 1986
- [8] Marr, David. ***La visión***. Alianza Psicología. 1985. QP475.M2718 1985.
- [9] Pratt, Willian K. ***Digital Image Processing***. Wiley-Interscience. TA1632.P7 1991
- [10] Sedgewick, Robert. ***Algorithms in C***. Addison-Wesley. 1990. QA76.73.C15S43 1990
- [11] Wechsler, Harry. ***Computational vision***. Academic Press. 1990.
- [12] Hecht-Nielsen, Robert. ***Theory of the Backpropagation Neural Network***. HNC, Inc. 5501 Oberlin Drive. San Diego, CA 92121. 619-456-8877 and Departement of Electrical and Computer Engineering. University of California at San Diego. La Jolla, CA 92139
- [13] ***Stuttgart Neural Network Simulator. SNNS***. Versión 4.1. UNIVERSITY OF STUTTGART INSTITUTE FOR PARALLEL AND DISTRIBUTED HIGH PERFORMANCE SYSTEMS (IPVR). Applied Computer Science – Image Understanding
- [14] ANSI PH3.49-1971 ***American National Standard for general purpose photographic exposure meters***.