

Creación de imágenes de alto rango dinámico por interpolación de fuentes

Damián Benassi, Fernando Nellmaddin y Mariano Peyregne

Trabajo práctico final de “Procesamiento digital de imágenes”, II-FICH-UNL.

Resumen—Las imágenes de alto rango dinámico son imágenes en las cuales se cuenta con zonas con mucha iluminación y otras zonas con muy poca. La fotografía tradicional se basa en tomar una imagen a un nivel de exposición determinado, lo que no permite tener detalle simultáneo en las partes claras y oscuras. El método que se utiliza en este trabajo, se basa en tomar un conjunto de fotografías a diferente exposición y utilizar técnicas para combinarlas, de manera de obtener detalles en todas las regiones de la imagen. Se presentan dos métodos para la obtención de una imagen de alto rango dinámico. El primer método consiste en un promediado de las imágenes, mientras que el segundo realiza una interpolación cúbica en vecindarios definidos a partir de todas las fuentes.

Palabras clave—HDR, valor de exposición, interpolación cúbica

I. INTRODUCCIÓN

La fotografía tradicional consiste en capturar una escena y almacenar en un archivo digital la cantidad de luz recibida por los sensores de la cámara. Las cámaras actuales, al tener un software incorporado de fábrica, permiten la configuración de ciertos parámetros en la captura de las imágenes. Uno de los parámetros de interés para el desarrollo de este trabajo, es el valor de exposición (Exposition Value, EV en inglés). Este parámetro resume dos cantidades de las que depende la exposición: el tiempo de exposición y la apertura. El tiempo de exposición se refiere al período de tiempo que se abre el obturador de la cámara. Por otra parte, la apertura define el grado en que se abre el obturador y por lo tanto, el ángulo de visión del dispositivo de captura. Al variar el valor de exposición, el resultado neto es que se capture más o menos luz, dando lugar a imágenes más o menos claras. Un valor de exposición alto dará lugar a imágenes más claras, donde se verán los detalles de zonas oscuras, mientras que un EV bajo obtendrá imágenes oscuras en donde se apreciarán los detalles de zonas claras o muy iluminadas. El valor $EV = 0$ define la imagen capturada sin modificación de los parámetros de fábrica de la cámara.

El rango dinámico de una imagen digitalizada se define como la razón entre el mayor y menor valor de gris que contiene la imagen. Sean I_{min} y I_{max} los valores mínimo y máximo que se detectaron en los sensores del dispositivo de captura, entonces el rango dinámico medido en decibeles, se puede definir como:

$$DR = 20 \log \frac{I_{max}}{I_{min}} \quad (1)$$

De este modo, la cantidad de grises que representa una imagen es una medida del rango dinámico de la misma. En un dispositivo estándar de captura, se tienen 8 bits para representar todos los grises, cuyo mayor valor es 255, y el mínimo se toma como igual a 1. Por lo tanto, su rango dinámico DR es $DR = 20 \log 255 = 48.13$ dB.

En las imágenes de alto rango dinámico, se busca combinar la información contenida en varias imágenes capturadas a distinta exposición, de manera de tener información sobre las zonas claras y oscuras. Al combinar las fuentes, la información que se pierde en una imagen oscura (de EV negativo), se ve compensada con la información contenida en una imagen clara, resultado de un valor de exposición alto (positivo); y la misma situación se presenta en el caso inverso. El rango dinámico en este caso, se define [1] como:

$$HDR = 20 \log \frac{I_{max} e_{max}}{I_{min} e_{min}} \quad (2)$$

donde e_{max} denota el tiempo de exposición máximo y e_{min} el mínimo.

II. MÉTODO

Imágenes fuente

Para la obtención de una imagen de alto rango dinámico, se parte de varias imágenes con diferentes niveles de exposición. Para el presente trabajo, se limitó a la utilización de hasta 4 imágenes fuente, con la condición de que sean capturas de la misma escena (mismos objetos, misma iluminación y misma posición de la cámara) pero con distinta exposición. Por ahora se supondrá que se tienen 4 imágenes fuente con distinto nivel de exposición. Más tarde se considerará el comportamiento del método en los casos donde hay imágenes faltantes.

A. Método de promediado

A partir de las imágenes fuente I_k , con

$$1 \leq k \leq M \leq 4$$

se realiza un promediado píxel por píxel entre todas las imágenes. Para cada píxel (i, j) de la imagen de salida O , su valor final se calcula como:

$$O(i, j) = \frac{1}{M} \sum_{k=1}^M I_k(i, j) \quad (3)$$

Este método es la primera aproximación a la obtención de imágenes de alto rango dinámico. Como se verá más tarde, a pesar de ser muy sencillo, logra un resultado aceptable.

B. Método de interpolación cúbica bidimensional

El segundo método que se propuso fue el de realizar una interpolación cúbica bidimensional [2],[3], con la información contenida en pequeños vecindarios del conjunto de imágenes. Para la aplicación de este método, primero se debe realizar una combinación de todas las imágenes fuente en una única imagen. Se ordenan estas 4 imágenes de manera de que la primer imagen I_1 corresponda a la fuente de exposición más negativa, e

I_4 corresponda a la imagen con valor de exposición más positivo. Se supone que todas las imágenes tienen el mismo tamaño $M \times N$. A continuación, se construye una nueva imagen combinada C de tamaño $2M \times 2N$ en la que se ubican los píxeles de las cuatro imágenes en un patrón como se muestra en la Fig. 1.

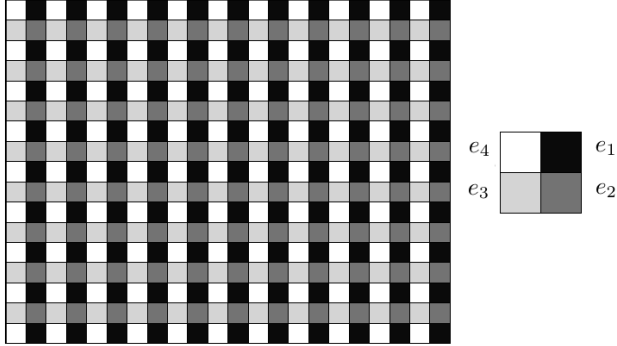


Fig. 1. Patrón de combinación de imágenes

En este patrón, un nivel de gris distinto identifica un valor de exposición distinto, siendo blanco el mayor valor de exposición (+2 por ejemplo) y negro el menor. El objetivo de esta combinación es la de tener la información de las cuatro imágenes en una sola. Se puede apreciar que, tomando de a grupos de 4 píxeles en un patrón de 2×2 , se tiene el mismo píxel de la escena pero a diferente exposición. La elección de este patrón es arbitraria y se eligió según el trabajo de [1], pero si se cambiara las posiciones, el resultado final sería distinto debido a que los pesos en las interpolaciones cambian.

El método propuesto, realiza una interpolación en pequeños vecindarios de 4×4 , tomando así 4 grupos de 4 píxeles cada uno. Se crea una imagen O de salida del tamaño original $M \times N$, donde el valor de píxel en cada posición se calcula como una interpolación cúbica bidimensional en el vecindario de 4×4 de la imagen C combinada.

En la Fig. 2, se muestra un vecindario de 4×4 de la imagen C . Se deja la explicación de los valores a_i para más adelante. La interpolación cúbica, calcula un sólo valor a partir de este vecindario, y ese valor será el que se copiará a la posición $O(\frac{i}{2}, \frac{j}{2})$ de la imagen de salida.

a_1	a_2	a_3	a_4
$I_4(i, j)$	$I_1(i, j)$	$I_4(i + 1, j)$	$I_1(i + 1, j)$
$I_3(i, j)$	$I_2(i, j)$	$I_3(i + 1, j)$	$I_2(i + 1, j)$
$I_4(i, j + 1)$	$I_1(i, j + 1)$	$I_4(i + 1, j + 1)$	$I_1(i + 1, j + 1)$
$I_3(i, j + 1)$	$I_2(i, j + 1)$	$I_3(i + 1, j + 1)$	$I_2(i + 1, j + 1)$

Fig. 2. Vecindario de 4×4

La interpolación cúbica bidimensional se realiza en dos pasos. Primero se interpola con un interpolador cúbico unidimensional en cada una de las 4 filas, obteniendo 4 valores distintos, uno para cada fila. En el segundo paso, se interpola estos 4 valores nuevamente con un interpolador cúbico unidimensional pero en la dirección vertical.

El interpolador cúbico unidimensional tiene un parámetro $0 < dx < 1$ que define el peso relativo que se le da a

los términos de la interpolación. La interpolación cúbica unidimensional para 4 valores a_i , con $1 \leq i \leq 4$, se define como:

$$a_{interp} = dx \{ dx [dx (a_4 - a_3 + a_2 - a_1) + (a_3 - a_4 - 2a_2 + 2a_1)] + (a_3 - a_1) \} + a_2 \quad (4)$$

Se emplea esta fórmula para calcular los valores interpolados para cada una de las 4 filas del vecindario. Se utilizan como los valores a_i los correspondientes, en cada fila, según la Fig. 2. Al finalizar estas primeras 4 interpolaciones, se obtienen los valores I_a , I_b , I_c y I_d de cada fila. A continuación, se realiza una interpolación final entre estos valores. Utilizando un paso dy vertical, la interpolación vertical tiene la siguiente forma:

$$O(\frac{i}{2}, \frac{j}{2}) = dy \{ dy [dy (I_d - I_c + I_b - I_a) + (I_c - I_d - 2I_b + 2I_a)] + (I_c - I_a) \} + I_b \quad (5)$$

Este valor de interpolación final se copia a la posición $O(\frac{i}{2}, \frac{j}{2})$ de la imagen resultado. Se repite este procedimiento para toda la imagen, pero siempre se toman vecindarios de 4×4 según la Fig. 2. Es decir que por ejemplo, no se consideran vecindarios donde los a_1 correspondan a posiciones pertenecientes a las imágenes I_1 o I_2 .

Al terminar, se obtiene una imagen del tamaño original $M \times N$ donde cada píxel es resultado de una interpolación cúbica bidimensional entre las 4 imágenes fuente. Como se puede apreciar, este procedimiento tiene mucha más complejidad que el método de promediado, donde sólo se consideraban vecindarios de 2×2 y se los promediaba.

Posproceso

Para finalizar, se propone un apartado de posproceso en el que se aplican unas correcciones a la imagen resultado. Se añaden aquí como un paso final de los métodos, pero se debería evaluar su aplicación según cada caso en particular.

1) *Aplicación de corrección gamma*: Como se notó que las imágenes quedaban por lo general más oscuras de lo deseable, se decidió aplicar una corrección gamma [4]. Sea r el valor de gris de entrada, y s el valor de salida, la corrección gamma se define como: $s = ar^\gamma$

Para los casos en que se probó el método, se utilizaron los valores $\gamma < 1$ y $a = 1$. La aplicación de esta operación, en cada uno de los canales RGB, tiene un doble efecto. El primero es el realce de las zonas oscuras de la imagen resultado. Además, y como una consecuencia de esto, se obtuvo una mejor definición de los bordes debido al contraste ganado en las zonas oscuras.

2) *Realce de bordes*: Debido a que el promediado y, en menor medida la interpolación, realizaban un difuminado de bordes, se optó por aplicar un filtro de realce de bordes, de manera de volver a recobrar aquellos que se habían perdido en el proceso principal. Para llevarlo a cabo, se aplicó a las imágenes fuente que mayor definición de bordes tenían, un filtro de Sobel de realce de bordes [4]. La imagen de bordes se sumó a la imagen obtenida en la interpolación,

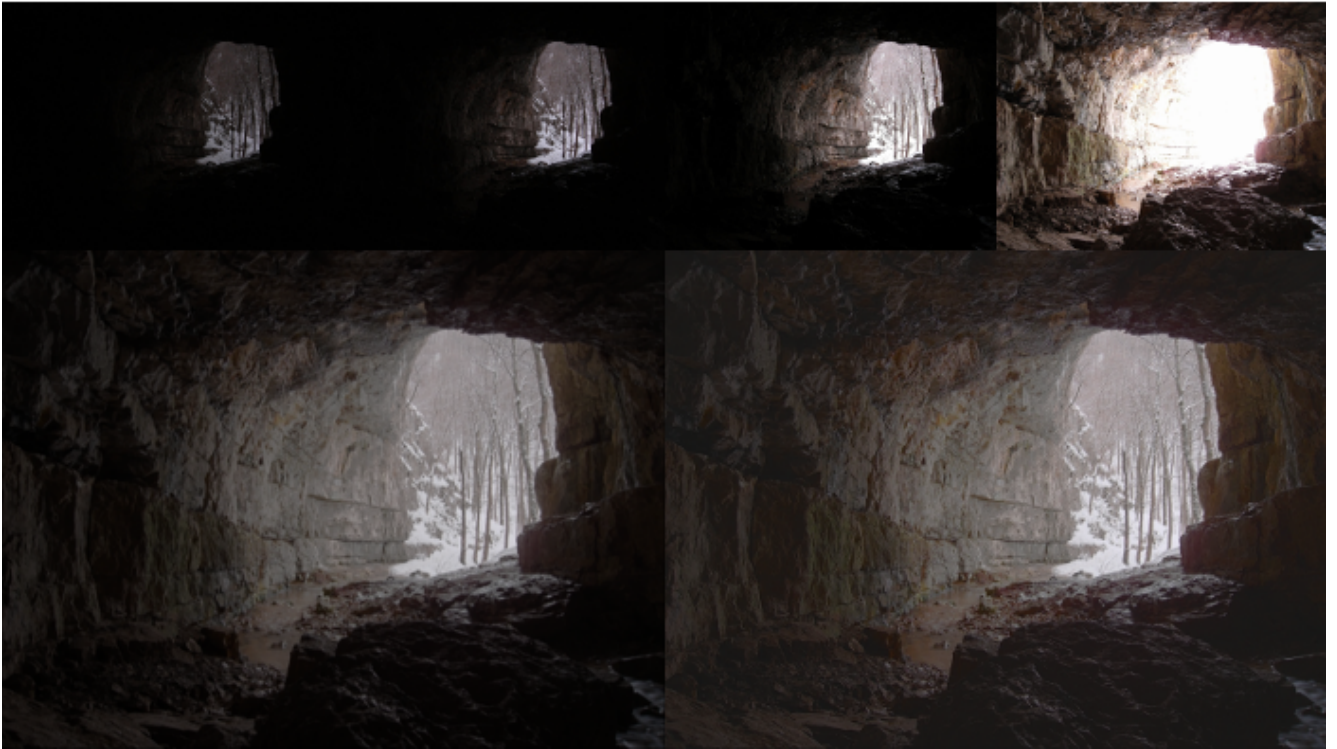


Fig. 3. Arriba: Imágenes fuente con exposición -3, -2, +0 y +4. Abajo: Resultado de promediado y de interpolación (sin posproceso)



Fig. 4. Resultado final luego del posproceso

dando como resultado una imagen de alto rango dinámico y con los bordes originales presentes y acentuados.

III. RESULTADOS

Para comprobar la efectividad del método, se utilizaron grupos de 4 imágenes con diferente exposición. Los valores de exposición que se utilizaron, según cada caso de prueba, se ubicaron entre: -4 y $+4$. Es de notar que las apreciaciones aquí mencionadas son exclusivamente subjetivas y basadas en el juicio de los autores; esto es debido a que no es simple obtener una medida objetiva representativa de qué es una buena imagen de alto rango dinámico.

El método de promediado, a pesar de ser simple, da resultados satisfactorios. Sin embargo, genera un difuminado mínimo que no es deseable y tiene un impacto negativo en el resultado final. Como primera aproximación trivial, fue de utilidad para compararlo con el resultado del método de interpolación.

Por su parte, el método de interpolación cúbica tiene excelentes resultados. A pesar de ser computacionalmente más costoso, el producto final es superior al del método de promediado. Razón de esto es que la interpolación no tiene un gran efecto de difuminar bordes, si no que incluso los realza.

Sin embargo, la aplicación del método de promediado o de interpolación no resultan suficientes, si no que se necesita aplicar la corrección gamma en cada uno de los canales RGB, de manera de obtener una imagen resultado más definida y con colores más vivos. Asimismo, al apreciar que se pueden perder bordes en el proceso, se aplica un filtro de Sobel para realzar los bordes y obtener el producto final.

En la Fig. 3, se pueden ver las 4 imágenes fuente, y los resultados del promediado y la interpolación. En la Fig. 4 se ve el resultado final luego de la aplicación del posproceso. El rango dinámico logrado en este caso, es de 95.75 dB.

A. Comportamiento con menos imágenes

Durante todo el desarrollo, se limitó la cantidad de imágenes fuente a 4. Este número se eligió porque esta cantidad de tomas fue suficiente para la mayoría de los casos. Además la captura de más imágenes dificultaría la aplicación de este método, pues se necesita que la escena permanezca estática en cada una de las tomas y esto es algo difícil (y a veces imposible) de conseguir.

Si se dispone de menos de 4 imágenes, los métodos de promediado e interpolación se pueden aplicar sin modificaciones, con la única consideración de que los píxeles de las imágenes faltantes se tomen con color negro. De esta manera, su ausencia no añade información inexistente, y el proceso se puede realizar con sólo las imágenes que se disponen.

El resultado en estos casos es dependiente de qué información se cuenta, ya que si sólo se cuenta con imágenes de exposiciones negativas, el producto final no tendrá buena definición en las regiones oscuras. El mismo comportamiento se presenta en referencia a las regiones claras con exposiciones positivas. Sin embargo, a pesar de no tener información completa, se obtuvieron resultados aceptables.

B. Limitaciones

La principal limitación del método reside en su necesidad de disponer de muchas imágenes de la misma escena, y capturadas con diferente exposición. Esto requiere que se disponga de un soporte rígido y la escena no cambie entre la captura de todas las fotografías fuente. Si estos prerequisites no se cumplen, dos artefactos se generan en la imagen resultado. Por un lado, si la cámara se movió entre una imagen y la siguiente, el resultado final tendrá un efecto de desenfoque ocasionado por el movimiento relativo entre las imágenes fuente. El otro artefacto que se genera tiene relación a la aparición de “fantasmas” en el resultado, debido a la presencia de un objeto en una de las imágenes y su ausencia en alguna otra.

IV. CONCLUSIONES

El método de interpolación cúbica bidimensional, aplicado a imágenes que satisfacen los requisitos antes mencionados, resulta de gran utilidad para obtener imágenes de alto rango dinámico. A partir de un conjunto de imágenes capturadas a distinto valor de exposición, se logró obtener una imagen combinada que mantiene la información más importante tanto en las zonas claras como la oscuras. El resultado es superior al promediado y permite obtener una imagen de características profesionales a partir de imágenes capturadas con una cámara convencional. Las limitaciones del método son su principal debilidad, ya que en ocasiones se dificulta o es imposible obtener varias imágenes de exactamente la misma escena y en la misma posición. El método de promediado, como un enfoque trivial, fue de ayuda para evaluar el desempeño de la interpolación y tener un punto de comparación con el mismo.

A. Trabajo futuro

Existen dos frentes en los que se puede trabajar en el futuro. En el primer grupo, para atacar la limitación de que las imágenes se deban capturar sin movimiento entre sí, se pueden plantear métodos de corrección automática del movimiento. De esta manera, se podría permitir un mínimo movimiento entre las distintas fuentes, eliminando así una de las grandes limitaciones.

El segundo grupo en el que se puede trabajar a futuro, es en el de realizar una corrección gamma más acorde a las características propias de cada imagen. En el presente trabajo, la corrección gamma se realizó con el mismo operador en todos los canales, pero métodos más robustos realizan distintas correcciones gamma en cada canal. Sin embargo, en estos métodos y a diferencia del presentado en este trabajo, el proceso deja de ser automático y requiere la intervención del usuario.

REFERENCIAS

- [1] T. Nayar, S. y Mitsunaga, “High dynamic range imaging: Spatially varying pixel exposures,” en *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference*, vol. 1, pp. 472–479, 2000.
- [2] J. Richards y X. Jia, *Remote Sensing Digital Image Analysis, 4th Edition*. Springer, 2002.
- [3] R. Keys, “Cubic convolution interpolation for digital image processing,” en *IEEE Transactions On Acoustics, Speech and Signal Processing*, vol. 29, no. 6, pp. 1153–1160, 1981.
- [4] R. Gonzalez y R. Woods, *Digital Image Processing, 2nd Edition*. Prentice Hall, 2002.