

Un estudio empírico de los fotomosaicos

Héctor Benítez Pérez¹, Manuel López Michelone²

¹ Universidad Nacional Autónoma de México,
México

² Universidad Nacional Autónoma de México,
Instituto de Investigación en Matemáticas Aplicadas y Sistemas,
México

hector.benitez@iimas.unam.mx, morsa@la-morsa.com

Resumen. Un fotomosaico es una imagen compuesta por muchas otras pequeñas imágenes dispuestas de tal manera que, a cierta distancia, forman una imagen más grande y reconocible. Cada pequeña fotografía es seleccionada cuidadosamente para que tenga un color y una textura similares a la parte de la imagen original que representa. Más allá de su valor en los campos de entretenimiento y arte, la creación de los fotomosaicos involucra una metodología para crearlos, así como requisitos mínimos necesarios para un resultado visual aceptable. En este artículo proponemos describir cómo se crea un fotomosaico así como plantear algunos criterios como repetición de imágenes, entropía y blending (que serán representados por índices numéricos), para intentar valorar cuando un fotomosaico es mejor que otro. Los resultados obtenidos pueden considerarse una primera aproximación a estos criterios.

Palabras clave: Fotomosaicos, entropía, algoritmos, métricas.

An Empirical Study of Photomosaics

Abstract. A photomosaic is an image composed of many small images arranged in such a way that, at a certain distance, they form a larger, more recognizable image. Each small photograph is carefully selected so that it has a similar color and texture to the part of the original image it represents. Beyond their value in the fields of entertainment and art, the creation of photomosaics involves a methodology to create them, as well as minimum requirements necessary for a result visually acceptable. In this article we propose to describe how a photomosaic is created as well as to propose some criteria such as repetition of images, entropy and blending (which will be represented by numerical indices), to try to assess when a photomosaic is better than another. The results obtained can be considered a first approach to these criteria.

Keywords: Photomosaics, algorithms, metrics, entropy.

1. Introducción

Un fotomosaico es una imagen compuesta por muchas fotografías pequeñas que se combinan para crear una imagen más grande. La técnica de crear un fotomosaico implica dividir la imagen original en pequeñas secciones y luego reemplazar cada sección con una fotografía diferente que tenga colores y tonos similares. El resultado final es una imagen digital compuesta por cientos o incluso miles de fotografías individuales que se combinan para crear una imagen única y coherente [9].

El crédito por la idea de los fotomosaicos probablemente sea para Robert Silvers, quien desarrolló esta técnica cuando era estudiante en el MIT [4]. Silvers finalmente patentó su proceso y fundó Runaway, una empresa dedicada a generar fotomosaicos [7]. Hay algunos elementos que nos permiten mostrar cuando un fotomosaico está bien hecho:

- Las imágenes que se utilizan para crear el fotomosaico deben ser de alta calidad y tener una resolución adecuada para que se vean bien cuando se reduzcan de tamaño.
- La imagen resultante debe ser clara y nítida, y no debería haber distorsiones ni errores visibles (artefactos), en las imágenes que la componen.
- El fotomosaico debe ser coherente y tener un aspecto uniforme, lo que significa que las imágenes que lo componen deben tener un tamaño y una orientación similares.
- El fotomosaico debe ser reconocible como la imagen original a la que se ha aplicado el efecto. Es decir, la imagen resultante debe tener una forma y un aspecto general similares al de la imagen original.
- La elección de las imágenes que componen el fotomosaico debe ser coherente con el tema o la temática de la imagen original.

La calidad visual de un fotomosaico es algo que no se ha analizado formalmente. Y aunque la idea ya tiene algunos años, parece evidente que la calidad de un fotomosaico depende en parte de las propiedades del ojo humano para ver colores e imágenes.

Aun así, hay muchos puntos de interés sobre los criterios que se deben utilizar para seleccionar las imágenes específicas como parte del proceso de construcción del fotomosaico. Algunas de las preguntas más relevantes serían:

- ¿Cómo se puede definir algo similar a una distancia de color entre una parte de la imagen original y la colección de imágenes que se utilizarán en el fotomosaico?
- ¿Cuántas fotografías debe tener una biblioteca para construir un fotomosaico?
- ¿Cuántas repeticiones de imágenes puedes tener antes de que el fotomosaico parezca monótono?

La clave para comprender estos problemas se analiza en el artículo El reconocimiento de rostros, de Leon Harmon [3] La primera imagen del artículo mencionado es el rostro de Abraham Lincoln, que se construyó a partir de una colección de mosaicos grises sólidos.

Harmon, un investigador de Bell Labs en la década de 1970, que se consideraba un ciberartista, escribió el programa Magna Dott, que probablemente fue la primera aplicación de un filtro mosaico a una foto digital. Lo que importa aquí, sin embargo, es el estudio de la redundancia de la imagen, que toma la esencia de la imagen con un filtro de mosaico simple.

De hecho, Salvador Dalí tomó la imagen de Lincoln y la usó en su pintura Gala mirando al mar. Si se observa la imagen, digamos desde 30 metros de distancia, se puede ver el rostro de Lincoln, pero al acercarse a la pintura, se observa la imagen de la esposa de Dalí. La imagen se conoce popularmente como Lincoln en Dalivisión.

2. Construyendo un fotomosaico

Estas dos imágenes –la de Harmon y la de Dalí– son la clave del secreto de los fotomosaicos. En lugar de escribir un filtro de mosaico que cambie un área de la imagen original con un color sólido, es posible escribir un filtro de mosaico mejorado que cambie cada área de la imagen original con alguna imagen que contenga un color similar. Así, los elementos básicos de un programa que construye un fotomosaico son:

1. Tómesese una foto escaneada (imagen fuente) para procesar.
2. Defina una malla de cuadrícula en la imagen fuente. Cada región rectangular en dicha cuadrícula define el tamaño de una de las fotos que se colocará en esa posición de la imagen.
3. Calcúlese el color promedio de cada cuadrado de la malla definida.
4. Encuéntrese la imagen más cercana, con el color promedio más similar al cuadro de la cuadrícula, y sustitúyala en esa celda.
5. Repítase el proceso para toda la cuadrícula.

Un sistema para construir fotomosaicos no es un único programa. Se requieren diferentes programas para realizar una variedad de tareas. Un primer programa calcula el color promedio de cada imagen de la biblioteca a usarse en el fotomosaico que va a ser creado. Una vez teniendo esta información, se procede a crear el fotomosaico con un programa que sigue los pasos ya mencionados. El resultado es un archivo que describe qué foto va en cada región del fotomosaico. Un tercer programa ensambla finalmente la imagen final.

3. Colección de imágenes y color promedio

Aunque el enfoque elegido es muy sencillo, está claro que se requieren decenas de miles de imágenes digitales de alta calidad (fotografías) para construir el fotomosaico. (Silvers indica que tiene una biblioteca con más de 100 mil imágenes escaneadas para este fin [6].) Nuestro programa original apenas usaba 6 mil imágenes fotográficas de alta calidad (color de 24 bits), aunque hoy contamos con más de 120 mil fotografías.

Para entender cómo encontrar el color promedio de una imagen, debemos decir que cada fotografía digital está formada por píxeles que contienen tres componentes de color –que van de 0 a 255 en valor– (Rojo, Verde y Azul), que hacen el esquema RGB. La combinación de valores en cada componente RGB nos da diferentes colores específicos. Para encontrar el color promedio de la foto, tomamos la suma de cada píxel de la foto y dividimos el resultado por la cantidad de píxeles utilizados:

$$\text{Color Promedio} = \sum_{1}^n \sum_{1}^m (R_{nm}G_{nm}B_{nm})/nm, \quad (1)$$

donde n y m es el ancho y alto respectivamente, de cada foto procesada. El programa que procesa todas las imágenes genera un archivo que contiene el color promedio para cada una de las imágenes de la biblioteca, separada en sus componentes RGB.

4. Métricas de color

En este punto tenemos el color promedio de cada una de las casi 6 mil imágenes de la colección. Ahora lo que necesitamos es un criterio que establezca la distancia mínima entre el color promedio de cada celda de la cuadrícula en la imagen de origen y las imágenes de la colección. La idea más simple fue usar la distancia euclidiana (de la geometría analítica), en términos del color promedio RGB³ [8, 10]:

$$\Delta = \sqrt{(R_2 - R_1)^2 + (G_2 - G_1)^2 + (B_2 - B_1)^2}, \quad (2)$$

donde R_1 , G_1 y B_1 es el color promedio de los componentes RGB de la celda en la cuadrícula de la imagen de origen y R_2 , G_2 , y B_2 es el color medio de una de las imágenes de la colección. La imagen mejor seleccionada será la que tenga la menor distancia después de aplicar esta última ecuación a todas las imágenes de la colección contra el color de la celda de la cuadrícula que se está analizando. A esto lo llamamos métrica de color.

Esta métrica en particular –a la que llamamos métrica euclidiana– parece suficiente para construir un fotomosaico de forma adecuada. Sin embargo, hay otra métrica interesante, basada en un estudio de Thiadmer Riemersma [4], quien analiza cómo el ojo humano percibe los colores. Implementando las ideas de Riemersma (a la que llamamos métrica Riemersma), hallamos que la distancia entre dos colores puede definirse de la siguiente manera:

$$\Delta = \sqrt{(2 + \bar{r}/256)(\Delta R)^2 + 4(\Delta G)^2 + (2 + (255 - \bar{r})/256)(\Delta B)^2}, \quad (3)$$

donde:

$$\bar{r} = (R_1 + R_2)/2, \quad (4)$$

$$\Delta R = R_1 - R_2, \Delta G = G_1 - G_2, \Delta B = B_1 - B_2. \quad (5)$$

³ La distancia entre dos colores se refiere a la diferencia numérica entre los valores RGB (rojo, verde, azul) de dos colores. Existen varias formas de calcular este valor, como la distancia euclidiana o la distancia de Manhattan, que pueden dar resultados diferentes según los pesos dados a cada canal de color. El valor de la distancia se puede utilizar para comparar qué tan similares o diferentes son dos colores entre sí.

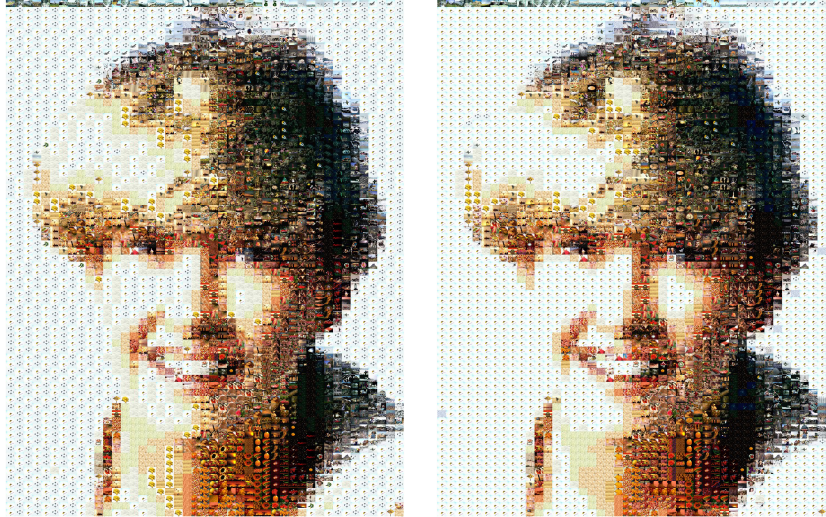


Fig. 1. Métrica euclidiana (izquierda) contra métrica Riemersma (derecha).

Riemersma tiene en cuenta la forma en que el ojo humano percibe el rojo, el verde y el azul. Con todos estos elementos, definimos es el algoritmo fundamental para producir un fotomosaico:

1. Se obtiene cada celda de la cuadrícula de la imagen de origen.
2. Se calcula el color promedio de esa celda en tiempo real.
3. Se encuentra la distancia mínima entre este número en comparación con el color promedio de cada imagen en la colección (usando alguna de las dos métricas definidas, métrica Euclidiana o Riemersma).
4. Se reemplaza la celda con la imagen elegida.
5. Se repite el proceso para cada celda de la cuadrícula de la imagen de origen.

5. Resultados

Escribimos todos los programas de computadora en Delphi 7 (Embarcadero), usando una computadora HP EliteDesk 705 (procesador AMD Ryzen 5 PRO de 6 núcleos), con 16 GB de memoria, con Windows 10, los cuales procesaron el color promedio de imágenes, la construcción del fotomosaico, el software que hace blending y el ensamble de la imagen final del mosaico en formato JPG. Esta es la descripción de los programas escritos:

- **Programa de pre-procesamiento de la colección de imágenes.** Este programa encuentra el color promedio de cada imagen. La salida de este software es un archivo de texto, que llamamos archivo indexado, que contiene en cada línea el color

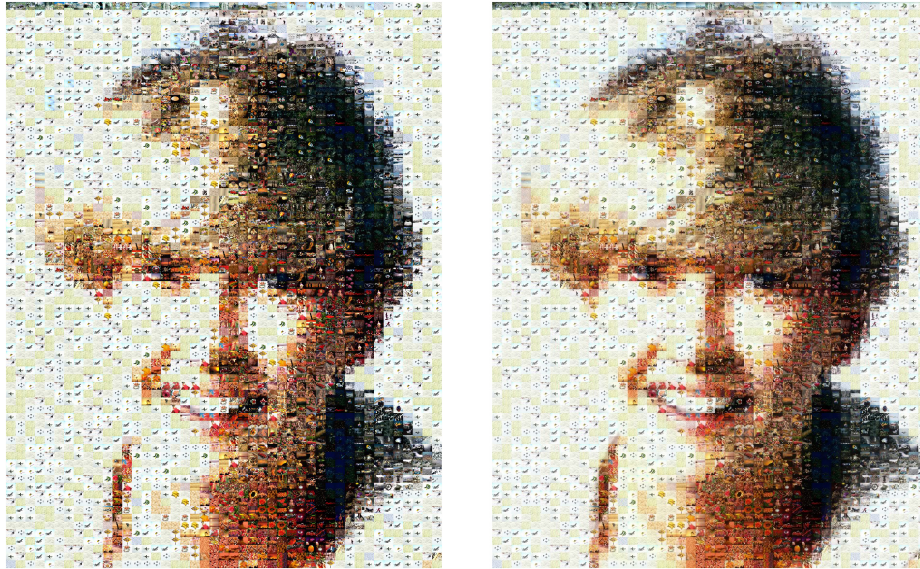


Fig. 2. Fotomosaico sin repetición con un 25 % de fusión (derecha).

promedio de cada imagen de la colección. Este conjunto de imágenes se procesa solo una vez, ya que el software de construcción de fotomosaicos lee este archivo para encontrar la mejor imagen para colocar en cada celda de la cuadrícula de imágenes (de acuerdo a algún criterio en particular).

- **Generador de fotomosaicos.** Una vez que se generado el índice de las fotografías a usar en un archivo, el programa que construye el fotomosaico. Se requiere la imagen fuente, la que se va a procesar y el nombre de un archivo que generará la descripción del fotomosaico a construir y el archivo de índice que contiene la descripción del color promedio de todas las imágenes de la biblioteca.
- **Programa para ensamblar el fotomosaico final.** El sistema genera un archivo de texto que contiene la descripción en qué región se coloca cada pequeña imagen. Normalmente, el software de creación de fotomosaicos necesita en promedio de dos a cuatro minutos para crear la descripción del fotomosaico que se va a construir.
- **Generador de imágenes.** Toma el archivo de texto de salida creado por el generador de fotomosaicos y ensambla la imagen final, que se puede guardar en formato JPEG. El archivo de salida es un conjunto de líneas que indican qué imagen colocar en cada posición de la cuadrícula. Esta se genera línea por línea.
- **Programa que analiza los índices de belleza del fotomosaico.** Este programa cuenta las imágenes usadas y permite fusionar la imagen original con el fotomosaico generado para ver si mejora visualmente.

El éxito de los resultados depende de la métrica ya que es en definitiva el criterio de elección de cada imagen que formará el fotomosaico. Las siguientes imágenes muestran los resultados con las dos métricas aplicadas, que no parecen ser radicalmente diferentes.

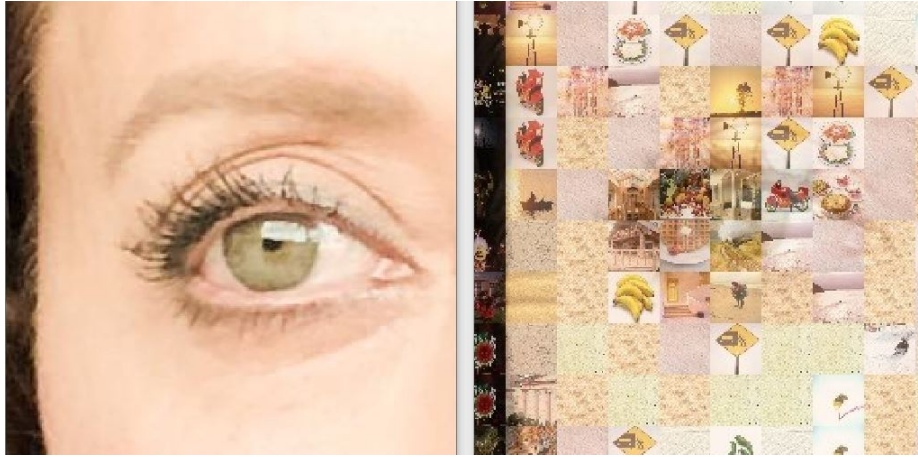


Fig. 3. Comparación en un fragmento de un fotomosaico contra la imagen original (blending de la imagen original al 35 %).

Un problema evidente es la repetición de las mismas imágenes, lo que hace que el resultado final sea monótono y poco deseable. Sin embargo, en este proceso observamos puntos importantes en la tecnología de creación de fotomosaicos:

- **Calidad de la imagen de origen.** No todas las imágenes son lo suficientemente buenas para ser procesadas por un software de fotomosaico. Las imágenes con más contraste son generalmente mejores. Obviamente, cuantos más colores tenga una imagen, mejor. Un equilibrio adecuado entre el contraste y el brillo es el mejor enfoque. Afortunadamente existen muchas aplicaciones que te permiten manipular imágenes en estos parámetros (Photoshop, ACDSee, Retriever, etcétera).
- **Colección de imágenes.** Para obtener mejores resultados, el software de fotomosaico debe tener acceso a una gran colección de imágenes de alta calidad (color de 24 bits), con una resolución promedio de al menos 800 x 600 píxeles. Las imágenes más pequeñas no dan buenos resultados en general.
- **Imágenes con muchos colores.** Es deseable que la colección de imágenes utilizadas sea numerosa y, además, que contenga imágenes con muchos colores. Por ejemplo, utilizamos la colección de 10 mil imágenes de Global Star Software (<http://www.globalstarsoftware.com/>) y descubrimos que, en comparación con las 6 mil imágenes originales utilizadas, las 10 mil contenidas tenían menos distribución de colores que nuestra colección original. ¿La moraleja de la historia? Una colección de imágenes con más balance de color es mejor que una gran cantidad de imágenes (demasiado oscuras o demasiado claras).
- **Las colecciones de imágenes.** Una gran colección de imágenes puede dar mejores resultados que una colección limitada. Se pueden encontrar colecciones de entre 30 mil y 50 mil imágenes de alta resolución –gratis– en archive.org⁴

⁴ <https://archive.org/search.php?query=imagen+colección&page=3>

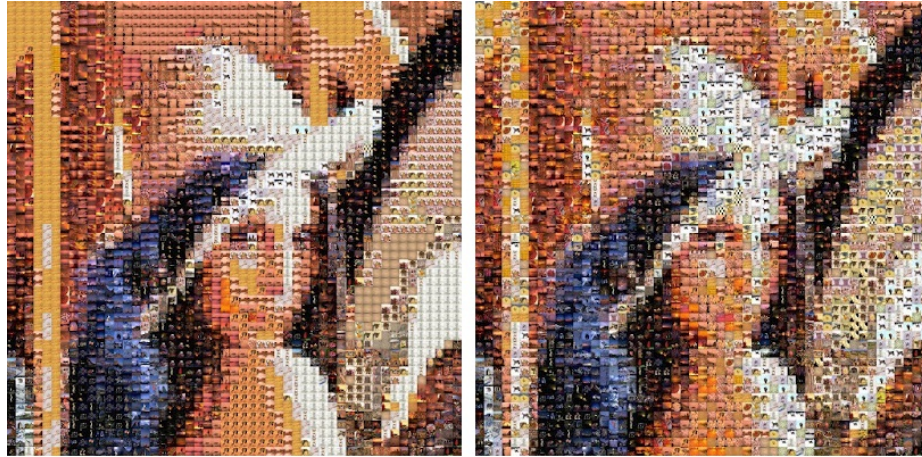


Fig. 4. Imagen de Lena con repetición (izquierda) y sin repetición de fotos.

5.1. Mejorando los fotomosaicos

Los fotomosaicos se pueden mejorar utilizando varias técnicas. Una de ellas es fusionar la imagen original encima de la imagen mosaico con un porcentaje bajo, lo que fusiona ambas imágenes. Esto se llama *blending*⁵ [11].

La combinación mejora los fotomosaicos porque la métrica solo tiene en cuenta el color de cada celda de la cuadrícula y muchas veces no encontrará una imagen que contenga un color similar al de la celda original y, además, si hay sombras o características pintadas en la región, la métrica no los tiene en cuenta. Por tanto, la fusión del fotomosaico con la imagen original (el *bending*) (aplicado ligeramente para que no se note demasiado) mejora el fotomosaico final.

Una idea utilizada por Silvers es no utilizar toda la imagen de la colección, sino un fragmento de la misma, buscando un mejor ajuste a la celda de la región analizada. Esta es una buena idea, pero se requiere mucho más procesamiento de imágenes. La repetición de las imágenes en el fotomosaico final genera monotonía.

Para evitar esto, se implementó la siguiente idea: se calcularon las 10 mejores imágenes para cada celda de la cuadrícula y se eligió una de ellas al azar. Por supuesto, esto no siempre le da a la imagen el color más cercano a la celda, pero elimina la monotonía y hace que el fotomosaico se vea mejor.

5.2. Entropía en la información de un fotomosaico

Para evaluar el grado de información transmitida en las regiones de fotografías que forman un fotomosaico podemos usar la entropía de información de acuerdo con Shannon [5], el cual evalúa el grado de información transmitida en las regiones del fotomosaico.

⁵ *Blending* es el proceso de combinar dos o más imágenes en una sola imagen, de manera que el resultado contenga la información visual de todas las imágenes originales.

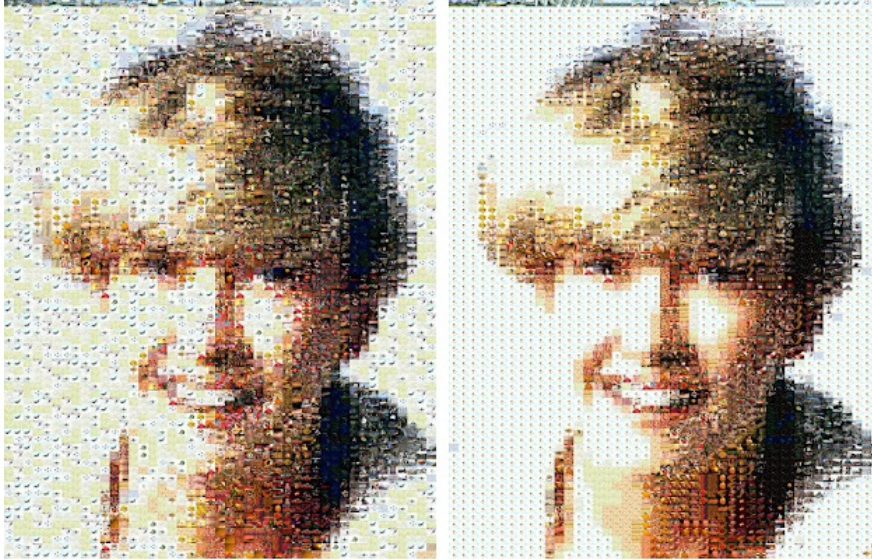


Fig. 5. Imagen sin repetición de fotos (izquierda) y sin repetición (derecha).

El concepto de entropía de la información describe cuanta incertidumbre (información) hay en una señal, en este caso una imagen. Por ejemplo, en una imagen en tonos de gris, con 256 posibles tonos, la entropía E puede calcularse como:

$$E = \sum_{i=0}^{255} p_i \log_2 p_i, \quad (6)$$

donde p_i es la probabilidad que un pixel al azar se elegido, teniendo una intensidad i [1]. En el contexto de una imagen en tonos de gris, la entropía calcula cuántos valores de pixel distintos hay en la imagen y cuánta variación hay en la distribución de esos valores. Si una imagen tiene una distribución uniforme de valores de pixel, entonces su entropía será alta, lo que significa que hay una gran cantidad de aleatoriedad presente.

Si una imagen tiene valores de pixel que se agrupan alrededor de un valor específico, entonces su entropía será baja, lo que significa que hay menos aleatoriedad presente. Por ejemplo, si tenemos una imagen I de color, podemos convertirla en tonos de gris y calcular su intensidad de la entropía.

Un valor de entropía mayor corresponde a una imagen o una región en el borde, mientras que un valor de entropía menor corresponde a un área de imagen suave, es decir, donde hay valores más homogéneos (de color), por ejemplo, cuando se aplica el blending [12, 2].

5.3. ¿Cuál es el mejor fotomosaico?

No hay ningún parámetro definitivo sobre cuando un fotomosaico es visualmente aceptable. Sin embargo, a pesar de que la belleza está en la mirada del observador, es posible definir una métrica en este sentido. Por supuesto que hay criterios variables.



Fig.6. Imagen con repetición de fotografías (izquierda) y sin repetición (derecha) (en escala de grises).

El hecho de que haya pocas repeticiones debería ser un buen criterio teniendo en cuenta que cuanto más se repiten algunas imágenes, más monótono o simple es el resultado final. Otro criterio a tener en cuenta es usar la fusión de imágenes (blending), es decir, mezclar la imagen original en un porcentaje contra la imagen del fotomosaico, para suavizar la imagen final y no se vea tan pixelada.

Cuando se usa blending en un fotomosaico, el resultado es que baja la entropía de la imagen final. Cuando analizamos cuántas imágenes se repiten en un fotomosaico se puede encontrar un índice muy sencillo que puede indicar claramente si un fotomosaico es mejor que otro a los ojos del observador bajo la premisa de que muchas imágenes repetidas hacen el fotomosaico una imagen monótona.

En este caso lo que se hizo es leer el archivo que describe el fotomosaico, el cual se utiliza para crear la imagen final. Luego, una rutina cuenta cuántas veces se repitieron las imágenes. Los resultados se graficaron para tener una idea de cómo se ven. Se grafica del menor al mayor número de repeticiones.

En un fotomosaico ideal probablemente se tendrá una línea con pendiente cero, es decir, donde todas las imágenes aparecen en la misma cantidad. Pero dado que este caso es el ideal y altamente improbable, tomamos esto simplemente como la referencia teórica. Con estos datos podemos contar cuántas imágenes diferentes aparecen en el mosaico.

Para encontrar el índice, tomamos el número total de imágenes que debe tener el fotomosaico (tomado del mismo archivo de descripción) dividido por el número de imágenes diferentes. En la Figura 4, tenemos que el índice del fotomosaico de Lena (izquierda), es 4.644. En la imagen derecha el índice del fotomosaico es 3.737.

Para esta imagen se usó una colección de 10,019 fotografías. En la siguiente imagen (Figura 5), se tiene un índice de 3.91 para la imagen izquierda y 4.70 para la imagen derecha (5560 imágenes usadas).

5.4. Resultados usando solamente imágenes en tonos de gris

El modelo RGB permite imágenes en tonos de gris, donde $R = G = B$, es decir, los tripletes de Rojo, Verde y Azul tienen el mismo valor, por lo que solo hay 256 tonos de gris... $(0, 0, 0), (1, 1, 1), \dots, (255, 255, 255)$. Está claro que la cantidad de imágenes disponibles puede cambiar el índice. Por lo tanto, tratamos de procesar una imagen (en tonos de gris), con una colección de 63.136 imágenes (con una repetición promedio de 246 imágenes con el mismo tono de gris en esa colección).

En el caso de permitir repetición, en la siguiente imagen (figura 6), el índice del fotomosaico fue 13.51. (imagen izquierda). Por otro lado, poniendo el criterio de elegir en cada imagen una de las 50 mejores, encontramos que el índice fue de 1.33. (imagen derecha). Un índice menor muestra menos repeticiones y por tanto, es menos monótono.

Con respecto a la entropía de las imágenes, para el caso de la figura 6, tenemos que el índice de entropía es de 7.9215 (izquierda) contra 7.9013 (derecha). De nuevo, la imagen con menos repeticiones presenta una entropía menor. Estos valores se obtuvieron usando el software MatLab, utilizando la biblioteca de procesamiento de imágenes.

6. Conclusiones

Algunos de los resultados más relevantes son:

- Una imagen sin permitir repeticiones hace un mejor fotomosaico, pero depende de cuantas imágenes consideremos para elegir la mejor. En las imágenes en tonos de gris, dado que solo podemos tener de 0 a 255 tonos diferentes, elegir uno entre los 50 mejores de entre más de 63 mil imágenes parece razonable. Cuanto menor sea el índice de repeticiones, mejor debe ser el fotomosaico.
- Las imágenes con repeticiones tienden a ser muy monótonas y tienden a tener índices más altos y visualmente son menos aceptables.
- El uso de blending suaviza el fotomosaico final y modifica favorablemente la entropía de la imagen (es decir, baja la entropía). El resultado final del fotomosaico es por ende más aceptable visualmente.
- El número total de imágenes de la colección es otro punto a tener en cuenta. Mientras más imágenes tiene una colección, la posibilidad de hacer un fotomosaico visualmente más aceptable, se incrementan.
- El uso de blending es una manera sencilla y económica para suavizar las imágenes procesadas y quitar en algún grado la pixelización del fotomosaico. Sin embargo, queda abierta la posibilidad de usar algunas técnicas de inteligencia artificial para hallar imágenes para las regiones del fotomosaico que contengan las sombras similares a las regiones en la foto a procesar.

Referencias

1. Farivar, R., Shokrollahi, A.: Adaptive photomosaic based on image entropy. *Journal of Real-Time Image Processing*, pp. 311–325 (2016)
2. González, R. C., Woods, R. E.: *Digital image processing*. Pearson Global Edition, pp. 545–549 (2017)
3. Harmon, L. D.: The recognition of faces. *Scientific American*, pp. 71–82 (1973)
4. Riemersma, T.: Colour metric (1999) www.compuphase.com/cmtric.htm
5. Shannon, C. E.: A mathematical theory of communication. *Bell System Technical Journal* (1948)
6. Silvers, R.: *Photomosaics*, Master Thesis. Massachusetts Institute of Technology (1996)
7. Silvers, R.: *Photomosaics*. Henry Holt and Co (1997)
8. Stricker, M. A., Orengo, M.: Similarity of color images (1995) doi: 10.1117/12.205308
9. Tran, N.: Generating photomosaics (1999) doi: 10.1145/298151.298213 <https://doi.org/10.1145/298151.298213>
10. Troebs, M.: Algorithms to create image montages (2000) pdfcoffee.com/algorithms-13-pdf-free.html
11. Wittenburg, T.: Alpha blending graphic images. *DDJ - Doctor Dobbs Journal* (1995)
12. Zhang, L., Ma, K. L., Yu, J.: Adaptively tiled image mosaics utilizing measures of color and region entropy. In: *Proceedings of the 9th International Symposium on Visual Information Communication and Interaction* (2016) doi: 10.1145/2968220.2968228