

INSTITUTO TECNOLÓGICO AUTÓNOMO DE MÉXICO



Foto-mosaicos: optimización en el arte

TESINA

QUE PARA OBTENER EL TÍTULO DE
LICENCIADO EN MATEMÁTICAS APLICADAS
PRESENTA

JULIO CÉSAR ESPINOSA LEÓN

ASESOR

DR. EDGAR POSSANI ESPINOSA

«Con fundamento en los artículos 21 y 27 de la Ley Federal del Derecho de Autor y como titular de los derechos moral y patrimonial de la obra titulada **“Foto-mosaicos: optimización en el arte”**, otorgo de manera gratuita y permanente al Instituto Tecnológico Autónomo de México y a la Biblioteca Raúl Baillères Jr., la autorización para que fijen la obra en cualquier medio, incluido el electrónico, y la divulguen entre sus usuarios, profesores, estudiantes o terceras personas, sin que pueda percibir por tal divulgación una contraprestación.»

JULIO CÉSAR ESPINOSA LEÓN

FECHA

FIRMA

*A mis padres,
por su apoyo, dedicación, paciencia y amor.*

Agradecimientos

Quiero agradecer a todas las personas que me acompañaron durante este camino

Al Dr. Edgar Possani, por sumarse al trabajo y apoyarme con tanta paciencia.

Índice general

1. Introducción	1
2. La programación lineal	3
2.1. Definición general	3
2.2. Programación entera	5
3. Fotomosaicos	10
3.1. Fotomosaicos: concepto y contexto histórico	10
3.2. Fotomosaicos en escalas de grises	12
3.3. Fotomosaicos en la triada RGB, o espacios de color . . .	17
3.3.1. Programa lineal	17
3.3.2. Idea del color promedio	20
4. El problema del agente viajero (TSP)	26
4.1. Descripción general del problema	26
4.2. Arte con el agente viajero, o arte con línea continua . .	32
5. Implementación y resultados	35
5.1. Partición de la matriz en rectángulos	35
5.2. Fotomosaicos	38
5.2.1. Fotomosaicos en blanco y negro	38

5.2.2. Fotomosaicos a color	52
5.2.3. Fotomosaicos con librería de R	59
5.3. Arte con el agente viajero	61
5.3.1. Arte con el agente viajero, caso Twitter	61
5.3.2. Arte con el agente viajero, complejidad y exploración de instancias más complejas	68
6. Conclusiones	78
A. Apéndice	86
A.1. Adecuando largo y ancho de pixeles para adaptarse a parámetros ingresados	86
A.2. Reduciendo dimensionalidad del vector de costos	87
Bibliografía	89

Capítulo 1

Introducción

Históricamente, la optimización matemática ha sido una disciplina altamente recurrida para la solución de problemas en varias áreas del conocimiento y la práctica (economía, informática, industria, logística, etc.). Sin embargo, fuera de la amplia esfera de problemas técnicos o puramente científicos abordables por la optimización hay una serie de modelos que concilian la optimización, particularmente del tipo lineal, con problemas de carácter estético más vinculados con ramas como el arte o la composición fotográfica. El presente trabajo abordará el problema de reproducir o replicar con buena similaridad una imagen mediante el empleo, ya sea de trazos continuos a lo largo de un cuadro, o mediante imágenes diversas que, posicionadas estratégicamente, lograrán construir un producto que, a ojo humano, tendrá una forma fácilmente asociable con la imagen objetivo. Se presentarán dos modelos de optimización lineal entera: el primero, un problema lineal con variables de decisión en tres dimensiones (siendo éstas asociadas a renglón, columna y tipo de imagen); y el segundo, una aplicación clásica en la Investigación de Operaciones con larga historia en el

campo; éste es el problema del agente viajero (TSP por sus siglas en inglés, *Traveling Salesman Problem*) para este trabajo se traza un camino hamiltoniano similar a la imagen con una historia y varios estudios relacionados que datan del siglo XIX, en el que William Hamilton y Thomas Kirkman lo plantearon (González-Santander, 2020).

En el Capítulo 3, se explica el planteamiento de los fotomosaicos en escalas de grises como un problema de programación lineal y la posterior extensión del problema al color (para este efecto, se desarrolla un paréntesis para dilucidar el concepto del color promedio a través de una técnica estadística, el k-medias); después, en el Capítulo 4, se explica el Problema del Agente Viajero (TSP), una formulación de programación lineal y las heurísticas que se han usado para obtener solución aproximada para este problema, después se muestran aplicaciones al conectar ciudades para replicar una imagen; posteriormente, el Capítulo 5 se enfoca en desarrollar los resultados y comentar los procedimientos que fueron seguidos en el cómputo de las soluciones de los problemas y las correspondientes imágenes desplegadas. Finalmente, en la conclusión (Capítulo 6) se dan algunas observaciones finales y probables alcances futuros.

Capítulo 2

La programación lineal

2.1. Definición general

La programación lineal es un área de la optimización que busca minimizar o maximizar una función lineal sobre múltiples variables, de tal manera que dichas variables cumplan un conjunto de restricciones (ecuaciones o inecuaciones), también lineales. La expresión estándar de un problema lineal es la siguiente:

$$\min_{x_i} \sum_{i=1}^n c_i x_i \quad (2.1a)$$

$$\text{s.a. } \sum_{i=1}^n a_{ij} x_i \leq b_i \quad \forall j \in \{1, \dots, m\} \quad (2.1b)$$

$$x \geq 0 \quad (2.1c)$$

Las desigualdades en la expresión (2.1b), especifican un espacio geométrico cerrado, más específicamente, un poliedro convexo al cual se le denomina región de factibilidad.

La disciplina y teoría de la programación lineal se manifiesta en problemas que buscan maximizar un beneficio o minimizar un costo bajo circunstancias que restringen la disponibilidad de recursos disponibles. Dicho beneficio o costo depende de las cantidades de los recursos; en estos casos recursos y costos se relacionan en un sistema de ecuaciones o inecuaciones lineales. Algunos de sus primeros usos se remontan a la Segunda Guerra Mundial en aplicaciones de planificación de gastos y recursos para reducir costos propios y aumentar las pérdidas del enemigo. Otros usos consisten en problemas de flujos de redes en informática y de mercancías en comercio; así como economía y gestiones de inventario, portafolios financieros y suministro de alimentos.

Se puede garantizar la resolución de los programas lineales en la medida de que la región factible sea no vacía y acotada en la dirección del gradiente de la función objetivo. El algoritmo Simplex fue el primero que se propuso para resolver estos problemas. Gráficamente la solución se puede representar como el contacto de un poliedro convexo con las curvas de nivel de una función objetivo cóncava (Fernandez, 2011). Sin embargo, al no haber convexidad estricta de la región factible ni concavidad estricta de la función, no se puede garantizar la unicidad de las soluciones; es más, para programas lineales no discretos puede haber hasta una infinidad de soluciones. Un caso que ejemplifica la infinidad de soluciones es cuando una de las restricciones de la función objetivo es paralela a las curvas de nivel. Lo que sí se puede asegurar es que cuando menos un vértice de la región factible es una solución óptima. La gráfica 2.1 da un ejemplo de un programa lineal en el espacio 2-dimensional donde las curvas de nivel van desplazándose en valores de la función objetivo en el vértice crecientes por cada iteración hasta dar con el vértice óptimo. Esta solución gráfica

ejemplifica un concepto que extiende su alcance y validez a más dimensiones y ante variedades de tamaño arbitrario de m restricciones.

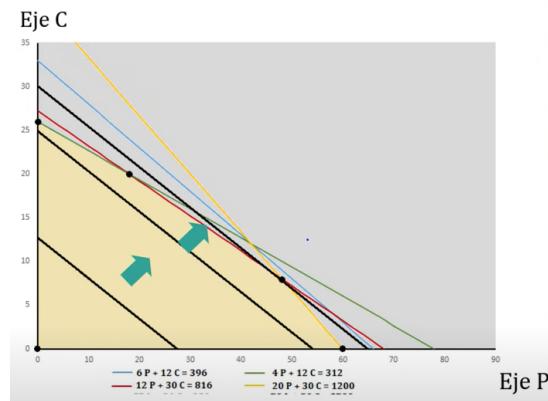


Figura 2.1. Ejemplo gráfico de un problema de programación lineal en 2 dimensiones (rectas representan iteraciones de soluciones factibles mejoradas por método Simplex).

Las variaciones de esta clase de problemas extendidas a más dimensiones pueden resolverse mediante el método Simplex, que incorpora variables de holgura para convertir las desigualdades en igualdades y ejecuta operaciones lineales simples sobre la función objetivo y las restricciones para encontrar el óptimo del problema. Este método irá llevando, a quien lo resuelve, de vértice en vértice hacia mejores soluciones (mejores valores de la función objetivo) cada vez hasta llegar a uno de los óptimos globales.

2.2. Programación entera

Algunos problemas requieren soluciones enteras, es decir, incorporar al planteamiento (2.1) la restricción $x_i \in \mathbb{Z} \forall i \in \{1, \dots, n\}$,

o equivalentemente, $x \in \mathbb{Z}^n$. En ciertos casos, esta restricción puede aplicarse solo a un subconjunto de variables con cardinalidad menor a n , aunque dichos casos no se abordan en este trabajo. La adición de la restricción de integralidad complica la búsqueda del óptimo en un problema lineal. Una estrategia intuitiva es identificar la solución entera más cercana a la del problema relajado, que no incluye la exigencia de integralidad en la región factible. Sin embargo, el valor óptimo del problema relajado es igual o superior al del problema entero asociado, y no hay garantía de que las soluciones relajada y entera sean próximas, ni de que técnicas simples como el redondeo conduzcan a la solución óptima entera. Considérese un problema de programación lineal dado por $\max z = 5x_1 + 3x_2$, sujeto a $x_1 + x_2 \leq 4$, $2x_1 + x_2 \leq 6$, $x_1, x_2 \geq 0$. La solución relajada óptima es $(x_1, x_2) = (2.5, 1.5)$, con $z = 17$. Al imponer la restricción de integralidad ($x_1, x_2 \in \mathbb{Z}$), la solución óptima entera es $(2, 2)$, con $z = 16$. La solución entera más cercana al óptimo relajado, como $(2, 1)$, con $z = 13$, es factible pero no óptima, ilustrando con esto la no-optimalidad de un recurso como el redondeo.

Ante las dificultades mencionadas, se vuelve necesario explorar distintas zonas de la región factible usando algoritmos, el más conocido de éstos siendo el de ramificación y acotamiento.

En su capítulo para cursos de optimización, Goic (2009) define los elementos más básicos para la definición matemática de esta clase de problemas. En el mismo, establece la importancia de elegir de manera estratégica las coordenadas a explorar ya que, si bien esta elección es indistinta en función de encontrar el óptimo (la ramificación y acotamiento *per se* encuentra el óptimo), el iterar sobre las coordenadas en un buen orden permite ahorrar, dependiendo del problema en cuestión, iteraciones del algoritmo que implicarían costos

computacionales sustantivamente elevados. En la formulación de Goic, el planteamiento parte inicialmente de la resolución del problema lineal relajado y, dada esta primera aproximación, va incorporando dos restricciones sobre alguna variable no entera (no importa mucho cuál): una restricción con el valor de esta variable por encima del 'techo' del primer óptimo y otra por debajo del 'piso' del mismo, eligiéndose de entre ambas alternativas la restricción con la mejor solución. En caso de que esta última solución sea entera, se detiene la ejecución de los pasos dando con un óptimo; de otra forma, se procede a la elección de una nueva variable no entera a ramificar. Dada la solución más reciente, el algoritmo reproduce y evalúa las ramificaciones de distintas variables fraccionales hasta dar con una solución con todas sus variables enteras. En términos de formulación matemática, lo que se hace es transitar de un problema, P (el problema original) a un problema P_0 (el problema relajado), el cual sería el primer nodo en el conjunto total de ramas a explorar, de tal manera que, como definición inicial, $L = \{P_0\}$

$$(P) \min_{x_i} z = \sum_{i=1}^n c_i x_i \quad (2.2a)$$

$$\text{s.a. } Ax \leq b \quad (2.2b)$$

$$x \geq 0, x \in \mathbb{Z}^n \quad (2.2c)$$

$$(P_o) \min_{x_i} z = \sum_{i=1}^n c_i x_i \quad (2.3a)$$

$$\text{s.a. } Ax \leq b \quad (2.3b)$$

$$x \geq 0, x \in \mathbb{R}^n \quad (2.3c)$$

Luego, se deriva un subproblema, P_k , el cual elige una coordenada fraccional-no entera (i) del óptimo de P_0 , llámesele x^* , i.e., $f = x_i^*$ y optimiza los siguientes problemas subyacentes.

$$(P_k^-) \min z = \sum_{i=1}^n c_i x_i \quad (P_k^+) \min z = \sum_{i=1}^n c_i x_i \quad (2.4)$$

$$x \in \mathbb{R} \quad (2.5)$$

$$x_i \leq \lfloor f_i \rfloor \quad x_i \geq \lfloor f_i \rfloor + 1 \quad (2.6)$$

$$Ax \leq b \quad Ax \leq b \quad (2.7)$$

$$x \geq 0 \quad x \geq 0 \quad (2.8)$$

Los cuales se llaman respectivamente, rama inferior y rama superior. De tal manera que, si existe factibilidad en ambos problemas, se incorporan como nodos P_k^+ y P_k^- de la lista L . Iterativamente, y dependiendo de la integralidad de las coordenadas en la solución de la ejecución corriente, se irá corriendo esta misma lógica sobre las ramas del árbol de problemas descartando o “podando / acotando” problemas (lo cual se refleja en una reducción del conjunto L) sin solución factible o de solución entera y actualizando las soluciones óptimas integrales que existan de manera temporal en los nodos hasta el momento explorados (este óptimo se define como \hat{Z}). El algoritmo deja de correr cuando el conjunto de problemas, L , a ramificar sea vacío.

La complejidad computacional de un algoritmo se refiere al tiempo o espacio requerido para ejecutarse en función del tamaño de la entrada, n . Se le suele expresar por la notación *Big O* que describe la

relación tamaño - tiempo; algunos casos comunes de orden son $O(1)$ (no dependiente de n), $O(n^a)$ (polinomial de orden a) o $O(2^n)$ (exponencial). El hecho de que una solución hallada por método Simplex a un problema lineal sea de una complejidad determinada y que el algoritmo de ramificación y acotamiento resuelve sucesivamente varios algoritmos de este tipo, vuelve a la complejidad computacional del problema lineal entero naturalmente mayor al problema lineal relajado. Se trata de una complejidad que puede pasar de un orden polinomial específico a uno polinomial más grande o inclusive exponencial. Cuando se escala la proporción en la dimensión del vector de variables de decisión y dimensión de restricciones, suele volverse complejo abordar esta clase de problemas, por lo cual se recurre a las heurísticas, que no son sino métodos aproximados para encontrar soluciones óptimas o cercanas a la óptima de manera eficiente, sin garantizar la optimalidad global en la solución.

Capítulo 3

Fotomosaicos

3.1. Fotomosaicos: concepto y contexto histórico

Una imagen digitalmente desplegada parte comúnmente de la lectura / interpretación de un archivo JPEG o PNG. Ambas clases de archivos pueden traducirse como un arreglo numérico de píxeles, entendiendo al pixel (abreviación de *picture element*) como la unidad mínima de información en la imagen. Esto se ve con mayor detalle más adelante en este capítulo. En la mercadotecnia han abundado técnicas de publicidad que consisten en reproducir con elevada similaridad una imagen de interés conformada por un conjunto de imágenes yuxtapuestas que, individualmente vistos, no presentan asociación con la imagen aproximada. Un fotomosaico —o mosaico fotográfico— es una imagen dividida en secciones rectangulares, las cuales pueden entenderse o interpretarse como submatrices de la matriz: en el presente texto se habla indistintamente también de “rectángulos” o “subáreas”. En el fotomosaico cada sección se reemplaza por otra

imagen que se le asocia por algún criterio de similaridad (o distancia), mediante este arreglo se logra hacer reconocible una imagen. Para construir un fotomosaico, se requieren tres insumos: una foto objetivo, un conjunto de mosaicos a usar para una dada subsección y un procedimiento o algoritmo que asigne a cada rectángulo, un mosaico. La idea del fotomosaico tiene una historia reciente en el ámbito de la computación, cuando en fechas similares dos personajes implementaron procedimientos con los cuales podía llevarse a cabo esta tarea. En 1993 Joseph Francis, diseñador, desarrolló un *software* trabajando en el departamento de animación por computadora de la agencia de publicidad R/Greenberg. Por otra parte, en 1995, Robert Silvers, como estudiante de ingeniería en el MIT, desarrolló otro algoritmo que logró gran impacto y audiencia cuando construyó la imagen de la actriz Marilyn Monroe a partir de portadas de la revista LIFE como parte de una edición para el sexoagésimo aniversario de LIFE, haciendo extensivo su desarrollo a un *software* cuya patente registró años más tarde (Fraga, 2024).

El programa desarrollado por Silvers se vio reflejado en réditos económicos para él, principalmente en el campo de la mercadotecnia. Los algoritmos, planteamientos matemáticos / computacionales que planteó no se hicieron públicos con el objetivo de explotar la patente. Sin embargo, en los años sucesivos se implementaron otros procedimientos que desarrollaban fotomosaicos. Algunos ejemplos son, en el campo informático del *open source*, lenguajes como *R* y Python han desarrollado módulos que permiten generar fotomosaicos con los tres tipos de insumos ya mencionados; y todo esto, como da a entender el nombre de *open source*, por un costo cero para el usuario. La NASA, asimismo, en 2014 generó un fotomosaico de una vista del planeta tierra conformada por más de 36,000 fotos de personas tomándose

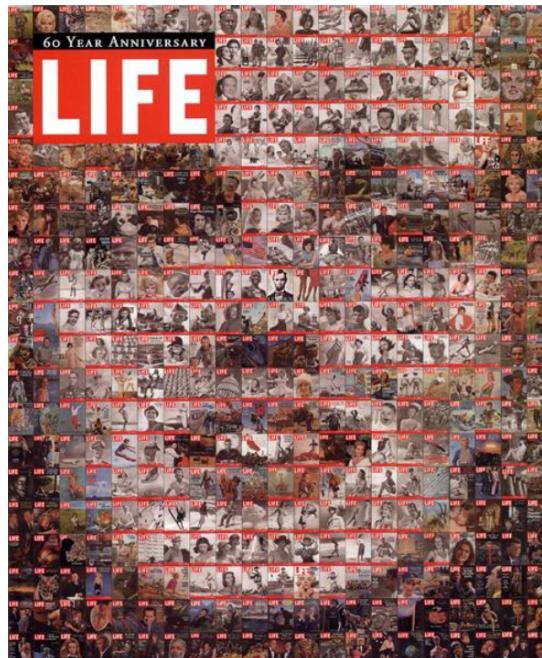


Figura 3.1. Marilyn Monroe formulada como conjunto de portadas de revistas LIFE; portada de aniversario 60. Por Robert Silvers, primera implementación relevante de algoritmo patentado.

selfies totalizadas en 3.2 mil millones de píxeles (o giga-píxeles) como parte de un programa orientado a concientizar sobre el cambio climático (Cole, 2014).

3.2. Fotomosaicos en escalas de grises

En la introducción se hace referencia a dos maneras de replicar imágenes con optimización lineal que se desarrollarán en el presente estudio. El primer modelo es el de los fotomosaicos en blanco y negro.

Parte del supuesto de que se cuenta con una imagen en blanco y negro con una gama de escalas de grises (o *greyscales* en inglés). Se asocia una menor escala de gris a una mayor luminosidad o blancura. Hay una serie de pasos estándar que deben de seguirse si se buscan generar las instancias de programación lineal para fotomosaicos (o para resolver con TSP como se verá en el Capítulo 4). La secuencia de procedimientos se detalla a continuación:

- La imagen fuente se divide en una cuadrícula de $m \times n$ bloques rectangulares, donde cada bloque tiene k píxeles de ancho por h píxeles de alto. La imagen compuesta resultante, o fotomosaico, tiene dimensiones de $\mu' \times \nu'$ píxeles, donde $\mu' = k \cdot m$ y $\nu' = h \cdot n$. Cada bloque se identifica por un par ordenado $(i, j) \in \{1, \dots, m\} \times \{1, \dots, n\}$, representando su posición en la cuadrícula. Además, cada píxel de la imagen fuente se asocia a una escala de grises en el rango discreto de 0 (negro puro) a 255 (blanco puro), denotada como $\psi_{i'j'}$.
- Se obtiene la escala de grises promedio de cada rectángulo en la entrada (i, j) , es decir $\psi_{ij} = \frac{1}{hk} \sum_{i'=1}^k \sum_{j'=1}^h \psi_{i'j'}$. Notar que aquí (i, j) denota posición del rectángulo, y (i', j') denota posición del pixel (unidad mínima de información en imágenes digitales) dentro de cada uno de tales rectángulos. Se divide el promedio generado entre 255, para efectos de determinar un cálculo de distancias más sencillo (acotado entre niveles de 0 a 1), $\beta_{ij} = \psi_{ij}/255$. Esta será una escala de luminosidad del rectángulo.
- Supóngase que se cuenta con un conjunto $F = \{1, \dots, cf\}$ de mosaicos (con cardinalidad cf), cada rectángulo representa un una fotografía (según la bibliografía, se les llama fichas, fotos,

mosaicos / *tiles* en el planteamiento original de Robert Bosch). Cada mosaico de tipo $f \in F$ cuenta con una intensidad lumínica con valor b_f . Supóngase que se cuenta con al menos un rectángulo con el mosaico de tipo f impreso, pero no se puede usar más de u_f veces. Bajo este planteamiento, se espera que el total sea de $\sum_{f \in F} u_f = U \geq m n$ mosaicos disponibles.

- Se posicionan los mosaicos de acuerdo a la solución óptima del problema presentado a continuación, (3.1), y se despliega la imagen resultante de colocarlos, dicha imagen es el fotomosaico.

En (3.1) se muestra el planteamiento del vector de costos y el conjunto de restricciones introducido para el artículo realizado en la *Orbelin College*, “Opt Art” (Bosch, 2006).

$$\min_x z = \sum_{f \in F} \sum_{i=1}^m \sum_{j=1}^n (b_f - \beta_{ij})^2 x_{fij} \quad (3.1a)$$

$$\text{s.a. } \sum_{i=1}^m \sum_{j=1}^n x_{fij} \leq u_f \quad \forall f \in F \quad (3.1b)$$

$$\sum_{f \in F} x_{fij} = 1 \quad \forall (i, j) \in \{1, \dots, m\} \times \{1, \dots, n\} \quad (3.1c)$$

$$x_{fij} \geq 0 \quad \forall (f, i, j) \in \{1, \dots, cf\} \times \{1, \dots, m\} \times \{1, \dots, n\} \quad (3.1d)$$

Llámese $z^* = \arg\min z$ al valor de la función objetivo en el óptimo del problema (3.1). La expresión (3.1a) minimiza la suma de diferencias cuadráticas—o bien, de los cuadrados de las distancias euclídeas en \mathbb{R}^1 —entre las luminosidades del mosaico y el rectángulo en el área subyacente sobre la cual se coloca el mosaico f

en la posición (i, j) (Possani, 2012). La función de costos será de tipo tridimensional: $c_{fij} = (b_f - \beta_{ij})^2$. La condición en (3.1b) garantiza que no se excedan las dotaciones disponibles de cada figura. La restricción (3.1c) asegura que haya exactamente una foto en la posición (i, j) . Cabe destacar que, si bien este problema es de programación lineal entera (los valores admitidos por el modelo son dicotómicos: 0 o 1), el mismo se puede plantear como un problema con variables de decisión continuas y aún así se deriva de éste un resultado óptimo entero. Bosch, en su artículo *Opt Art* asegura esta integralidad, la razón de la misma es que el problema en cuestión es una instancia específica de una categoría de planteamientos de optimización combinatoria catalogados como “El Problema de Asignación” que busca asignar un número de objetos (tareas, recursos) a un número igual de destinatarios (personas, máquinas o, en este caso, rectángulos ordenados), de tal manera que se minimice el costo total de la asignación. La matriz de restricciones de “El Problema de Asignación” es totalmente unimodular, lo cual quiere decir que cada una de sus submatrices cuadradas son invertibles, con dichas submatrices teniendo un determinante igual a 0 o ± 1 . En programación lineal, esta propiedad garantiza que las soluciones básicas del problema relajado sean enteras. Esto es una noticia buena, ya que añadir la propiedad de “entero” o “integralidad” implica en la solución complejidades sustanciales que dificultan mucho la tarea de optimización; el algoritmo de ramificación y acotamiento para problemas enteros expresado en el Capítulo 2 ilustra esta aseveración, pues la ramificación podría iterarse hasta mn veces.

El planteamiento referido cuenta con tres dimensiones en su vector de costos, c_{fij} : la primera corresponde al mosaico de tipo f (que va a contrastarse en escala de grises con el rectángulo en cuestión) y la

segunda y tercera a la coordenada del rectángulo (horizontal y vertical, respectivamente). Una variación a este problema consiste en fichas de dominó con un par de valores ordenados $(w, z) \in \{0, \dots, 9\} \times \{0, \dots, 9\}$ que sustituyen a la dimensión de fichas f , donde $w < z$ y cada ficha puede cubrir dos rectángulos yuxtapuestos vertical u horizontalmente. La sustitución de f por el par ordenado de valores de las fichas implica pasar de tres a cuatro dimensiones en el vector de costos. Ahora bien, la orientación de la ficha de dominó implica una nueva dimensión, la quinta: o que serían las cuatro posibles orientaciones de la ficha de dominó. De tal forma que el vector de costos resultantes sería de tipo c_{wzjo} . Adicional se tiene una restricción de tipo (3.1c) pero que considera que los mn espacios deben llenarse una única vez. Este planteamiento fue introducido por Bosch en su artículo “*Constructing Domino Portraits*” de 2004; en el mismo, se muestran fotomosaicos del programa resultante, de los cuales en Figura 3.2 se despliega un ejemplo de retrato de Marylin Monroe que sigue la línea de las representaciones de la artista (tal como el del fotomosaico de LIFE mencionado antes en este capítulo). Naturalmente el problema planteado para mosaicos con fichas de dominó tiene más complejidad combinatoria que el planteamiento de esta sección (3.1), ya que se tienen dos dimensiones más en los costos, así como más restricciones. Si bien el desarrollo matemático de este problema no es objetivo de este escrito, sí ayuda a comprender el tipo de variaciones que puede uno creativamente ingenierarse sobre el modelo de fotomosaicos al transformar el vector de costos y las especificaciones de las restricciones.

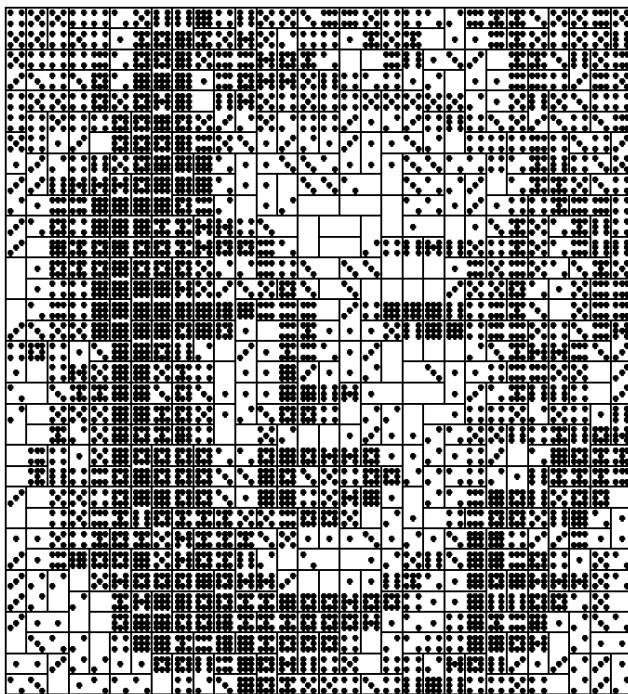


Figura 3.2. Ejercicio de Monroe por Robert Bosch usando 9 juegos (con 55 fichas por juego)

3.3. Fotomosaicos en la triada RGB, o espacios de color

3.3.1. Programa lineal

Una pregunta interesante es ¿cómo se construiría un fotomosaico usando el principio de optimización lineal planteado al inicio del capítulo, pero ahora en lugar de optimizar la asignación de escalas de grises, hacerlo en un contexto de color? Para ello, vale la pena cuestionarse cómo se representa una imagen a color. En los niveles primarios de educación es común mostrar empíricamente (con

acuarelas, digamos) cómo los colores se producen mezclando rojo, amarillo y azul, a esto se le llama en el modelo RYB, identificado como un modelo de tipo sustractivo (de hecho es el más común dentro de esta clase). Un modelo sustractivo parte del principio de que “el color de un objeto depende de las partes del espectro electromagnético que son reflejadas por él, o dicho de otro modo, de las partes del espectro que no absorbe” de tal manera que a más colores se superponen, más oscuro es el resultado, esto implica que se resta luminosidad, de esto se deriva el nombre de “sustractivo” (Castañeda, 2005). La manera en que el ojo humano interpreta un objeto específico depende no necesariamente de la luz que el objeto en sí emane, sino de las longitudes de onda que deja de absorber. Por el contrario, la representación de un objeto por una pantalla depende de la emisión de la luz realizada por el dispositivo, de ahí que se ha de emplear un modelo opuesto, *i.e.*, el de la síntesis aditiva del color. Una ventaja que tiene la misma es que el color digital de un pixel procesado por un dispositivo puede ser representado como una incorporación de valores numéricos de cada pixel que, en su conjunto, forman un espacio 3-dimensional. El modelo aditivo mas común en la informática y la fotografía parte de los colores primarios rojo, verde y azul (RGB, por sus siglas en inglés). A este espacio de color 3-dimensional también se le llamará triada RGB. Cabe mencionar que, así como en las escalas de grises, en las imágenes a color cada componente puede representarse convencionalmente en escalas continuas entre 0 y 1 o discretas de 256 valores (estos son dos ejemplos: el modelo se elige dependiendo de los requerimientos del problema en curso). Una extensión del principio de la solución de fotomosaicos sería minimizar la suma de los cuadrados de las distancias euclidianas, pero esta vez ya no en \mathbb{R}^1 , sino en \mathbb{R}^3 . Para ello, la fórmula enunciada en (3.1a) puede modificarse

incorporando tres componentes: la primera dimensión, r (que antes existía, sin embargo se omitía / obviaba al tratarse de una única dimensión), la segunda, g y, finalmente, b .

Dicho esto, podemos ocupar los conceptos definidos inicialmente en este capítulo, pasando así de β_{ij} y b_f a β_{ij}^A y b_f^A donde el superíndice A refiere a “aditivo”; para este caso, $\beta_{ij}^A = (\beta_{ijr}, \beta_{ijg}, \beta_{ijb})$ es el color del pixel “promedio” (concediendo que pudiera llamársele así y, en efecto, se puede comprobar que semejante pixel existe) y $b_f^A = (b_{fr}, b_{fg}, b_{fb})$ es el color del mosaico f . Habiendo definido esto, generamos el problema lineal siguiente:

$$\min_x z = \sum_{f \in F} \sum_{i=1}^m \sum_{j=1}^n [(b_{fr} - \beta_{ijr})^2 + (b_{fg} - \beta_{ijg})^2 + (b_{fb} - \beta_{ijb})^2] x_{fij} \quad (3.2a)$$

$$\text{s.a. } \sum_{i=1}^m \sum_{j=1}^n x_{fij} \leq u_f \quad \forall f \in F \quad (3.2b)$$

$$\sum_{f \in F} x_{fij} = 1 \quad \forall (i, j) \in \{1, \dots, m\} \times \{1, \dots, n\} \quad (3.2c)$$

$$x_{fij} \geq 0 \quad \forall (f, i, j) \in \{1, \dots, c_f\} \times \{1, \dots, m\} \times \{1, \dots, n\} \quad (3.2d)$$

Podemos simplificar la expresión del vector de costo unitario $(b_{fr} - \beta_{ijr})^2 + (b_{fg} - \beta_{ijg})^2 + (b_{fb} - \beta_{ijb})^2$ como $d^2(b_f^A, \beta_{ij}^A)$ con d la norma euclídea en \mathbb{R}^3 . Este costo, que antes se encontraba en el intervalo de valores $[0, 1]$ ahora los toma en $[0, 3]$. Vale la pena observar que, si bien la expresión parece más complicada debido a cómo se expresa el coeficiente de costos, en realidad el problema es en esencia idéntico con dos salvedades: la primera, que el cálculo de los costos es más largo; la segunda, que encontrar mosaicos que cubran de

manera razonablemente similar el total del espectro de colores que hay en la imagen grande a representar ya no es tan sencillo, debido a que se cuenta con tres dimensiones. Así —adelantándonos a la implementación y la elección de mosaicos— si en el planteamiento en escalas de grises se deseaba tener 10 imágenes con escalas de grises promedio homogéneamente distribuidas sobre el rango $[0, 1]$ (separadas por un *grid* de 0.1) ahora se necesitarían $10^3 = 1000$ imágenes con escalas promedio distintas para cumplir la tarea análoga; por lo tanto la cardinalidad del conjunto mosaicos F , debe de aumentar considerablemente si se desea generar una representación legible de una imagen con tonalidades diversas, afectando por consiguiente la complejidad computacional de la solución del problema. Como conclusión la base detrás del planteamiento es idéntico, pero incrementar la cardinalidad del conjunto de mosaicos a colocar implica incrementar los requerimientos computacionales de la resolución.

3.3.2. Idea del color promedio

En el subcapítulo anterior se anticipaba que sí existe el concepto de un color dominante o promedio en una imagen. En efecto, desde la década de 1970 existe una teoría en torno a la cuantificación del color que busca comprimir una variedad de tonos de color a un valor y que dió luz a nociones que hoy se manejan recurrentemente, relacionados, por ejemplo, con la popular extensión de archivo digital, JPEG. Hay dos conceptos centrales en torno a la cuantificación del color. El primero es que éste se usa para reducir costos computacionales, en especial en lo referente a dispositivos con una capacidad limitada, al reducir optimalmente la información desplegada por un archivo de imagen (tal es el caso del JPEG). El segundo, es que da al analista los elementos para entender las variedades de color, ayudando a

categorizar el color de una cosa o producto, o bien a derivar una paleta optimizada de una imagen.

Relacionado con la segunda cuestión, destaca el algoritmo de k-medias (universalmente conocido como *k-means* por su denominación en inglés) que permite usar exactamente la misma dimensión (renglon por columna, $\mu' \times \nu'$ en términos del planteamiento explicado) de pixeles pero donde la totalidad de valores de los mismos en el hiperplano RGB sea limitada y asignada manualmente por el usuario, como un hiperparámetro. Antes de ahondar en este enfoque, conviene hacer una breve recapitulación de dicho algoritmo, elemental en el estudio del aprendizaje supervisado, cuyo objetivo es segmentar el universo en un agrupamiento que opera en función de la proximidad vista como distancia euclídea (Chong, 2021). En síntesis, para un conjunto de n observaciones en x , en un espacio euclídeo (indistintamente de su dimensión, para el análisis del espacio de color RGB la dimensión es 3), el objetivo del algoritmo es seleccionar, primeramente, el hiperparámetro k o número total de *clusters* / segmentos a generar; posteriormente se enfoca en seleccionar los mejores centroides, asignando a cada punto con su centroide más cercano (lo cual implica una asignación de *cluster*). De tal manera que el promedio de la suma cuadrática de la distancia de cada punto a su correspondiente centroide se vea minimizado. El algoritmo se ejecuta en bucle; sin embargo, en el resultado final computado, el centroide debe de ser igual al promedio, μ_i del i-ésimo cluster generado. Visto matemáticamente, se optimiza la expresión (3.3), la cual se encuentra en el artículo “*K-means clustering algorithm: a brief review*” de Chong (2021).

$$z^* = \underset{S}{\operatorname{argmin}} \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2 \quad (3.3)$$

Donde $S = \{S_1, \dots, S_k\}$ constituye la segmentación óptima de *clusters* y la unión de los mismos compone el universo total de observaciones, $\bigcup_{i=1}^k S_j = \{x_1, \dots, x_n\}$. Supóngase inicialmente el caso con un *cluster*, *i.e.* $k = 1$. En esta casuística hay un único centroide μ_1 , y puede considerarse un ejercicio trivial donde el color “promedio” es el promedio aritmético en cada una de las tres dimensiones de los *renglon* \times *columna* pixeles. Ejercicios empíricos pueden conducir a la conclusión de que este no es un muy buen valor para instanciar k , ya que en una imagen con cierta variedad en su gama de colores, este promedio aritmético estaría capturando cierto ruido de aquellos segmentos de la imagen que, si bien no son muy grandes, generan contraste, haciendo que el promedio computado se desvíe del color que tiene el objeto central de la imagen. Regresando al problema de la imagen, si se asigna $k = 2$ se tiene un tipo de imagen; un refinamiento ocurriría si se usa $k = 3$; y mucho mayor sería éste si se pasa a $k = 10$. Dada la k , el algoritmo va a definir esta cantidad de centroides para agrupar la totalidad de valores de pixeles. Se desplegará cada centroide reemplazando al pixel original dando lugar a una imagen con la misma dimensión (renglón por columna) pero menor variedad en sus colores (Aqil Burney, 2014). El caso trivial es cuando k equivale al total de distintos valores de pixeles que hay en el espacio RGB, pues entonces el resultado es la misma imagen original. En la figura 3.3 se desarrolla la simplificación con k-medias, en un ejemplo con tres despliegues, cada uno con una k distinta.

El primer caso analizado corresponde a una imagen de flores blancas, presentada en la Figura 3.3a, con el objetivo de identificar el



(a) Imagen original



(b) 2 tipos de píxeles



(c) 3 tipos de píxeles



(d) 10 tipos de píxeles

Figura 3.3. Simplificación de imágenes basado en selección de píxeles-centroides del modelo k-medias (cuatro configuraciones para k)

tono de blanco predominante. La imagen tiene una resolución de 517×517 píxeles, lo que equivale a $517^2 = 267,289$ píxeles en total. Un análisis de agrupación (`COUNT(*) GROUP BY r, g, b`) revela que solo existen 58,168 combinaciones únicas de valores r, g, b (en el rango $[0, 1]$), lo que indica una repetición significativa de colores. De estas combinaciones, solo 22 aparecen en más de 500 píxeles cada una, representando más del 0.2 % del total de píxeles. Notablemente, todas estas 22 combinaciones tienen valores de $r, g, b > 0.90$, lo que confirma que corresponden a diferentes tonalidades cercanas al blanco puro. Esto sugiere la posibilidad de agrupar estas tonalidades en un único clúster, junto con otras combinaciones de r, g, b en una vecindad cercana en el espacio RGB. Para explorar esta idea con un enfoque geométrico más formal, los experimentos en la Figura 3.3 aplican un algoritmo de segmentación k-medias para reducir la diversidad de

colores a $k = 2, 3$, o 10 clústeres.

Ahora, la asignación de la k depende de la clase de problema a abordar. En general para k-medias (en aplicaciones más allá de la cuantificación del color) existen varios criterios que permiten determinar una k que satisfaga alguna condición “óptima”; dos métodos comunes son el método del codo y el análisis de la silueta. Como la participación del algoritmo k-medias en este escrito es meramente auxiliar y complementario para ilustrar el problema de optimización lineal en la asignación de mosaicos de color, se desea minimizar la discusión de pormenores del algoritmo. Para el caso específico de interés, a más homogénea es una imagen, más sencillo se espera que una k “pequeña” será suficiente para desplegar la imagen con aceptable precisión. El lector puede imaginarse una imagen con fondo verde (un jardín) o blanco (retratos para identificación), como es común en la fotografía. La presencia del fondo y su contraste con el objeto central en imágenes complejas (considérese “ruido”) dará lugar a una pérdida de información si asignamos un k muy pequeña. Por lo tanto, para efectos de simplicidad, en el capítulo de implementación será importante que cada mosaico a asignar del conjunto F sea una imagen con colores homogéneos, para así elegir el valor del centroide con más población de pixeles asociada, llamarle el “pixel promedio” del mosaico y asegurar que una k pequeña (a saber, $k = 2$) sea suficiente para determinar rápida y eficientemente un color promedio. Tantas imágenes con atributos similares al de la flor blanca mostrada se pueden imaginar como colores de flores existan. Esto es, imágenes simples, uni-color, con un fondo nulo o chico para minimizar el ruido. En el capítulo de la implementación esta inventiva se llevará a cabo análogamente, precisamente con flores. Similarmente, bajo esta idea de color promedio determinado con k-medias existirá un “pixel



(a)

Blanco-gris
dominante

Figura 3.4. Color promedio de flores blancas derivadas de
 $k = 2$.

promedio” en cada uno de los mn rectángulos, que posicionados contiguamente, conforman la imagen objetivo.

En relación a la imagen de la flor blanca, la presentación de la figura 3.4a muestra el color dominante derivado de $k = 2$, este color ya se conoció anteriormente en la figura 3.3b, pues es aquél visualmente más abundante de ésta, o expresado con formalidad, el centroide determinado por el algoritmo con más pixeles asociados.

Capítulo 4

El problema del agente viajero (TSP)

4.1. Descripción general del problema

La teoría de grafos estudia las propiedades y relaciones de un grafo que es un par (V, E) con V un conjunto de vértices y E conjunto de aristas que los une. Siendo la arista modelada, ya sea como un par dirigido, donde el orden de los vértices importa (digrafo o arco, representado por una flecha), o bien, no dirigido (representado por una simple arista o recta). Existen problemas de optimización planteados sobre grafos que pueden ser resueltos con programación lineal, siendo el problema del agente viajero (*Traveling Salesman Problem*, TSP) uno de los más interesantes y populares (Desrochers, 1991). Este problema se define dentro del marco de los caminos hamiltonianos: dicho tipo de ruta se enfoca en visitar cada vértice de un grafo exactamente una vez. El TSP plantea que un individuo (el agente viajero) debe visitar una cantidad n' de ciudades partiendo de una ciudad origen y finalizando

el recorrido en dicha ciudad. El objetivo es minimizar la distancia total recorrida. Existe un conjunto de parámetros necesario para definir el problema: la distancia que hay entre cada par de ciudades existente; para ello defínase $c_{i,j}$ como el valor de la distancia que hay entre las ciudades (o vértices) (i, j) para cada par $(i, j) \in \{1, \dots, n'\}^2$. Explicado este problema, se puede desarrollar intuitivamente una formulación matemática mediante la programación lineal entera en 4.1.

$$\min_{x_{i,j}} \sum_{i=1}^{n'} \sum_{j=1}^{n'} c_{i,j} x_{i,j} \quad (4.1a)$$

$$\text{s.a. } \sum_{i=1}^{n'} x_{i,j} = 1 \quad \forall j \in \{1, \dots, n'\} \quad (4.1b)$$

$$\sum_{j=1}^{n'} x_{i,j} = 1 \quad \forall i \in \{1, \dots, n'\} \quad (4.1c)$$

$$\sum_{i \in S} \sum_{j \in S} x_{i,j} \leq |S| - 1 \quad \forall S \subset \{1, \dots, n'\} \quad (4.1d)$$

Donde S es el conjunto potencia generado con $\{1, \dots, n'\}$. El problema lineal esbozado en 4.1 se define en tres tipos de restricciones. Para entenderlo, defínase antes el conjunto de variables de decisión como $\{x_{i,j} | (i, j) \in \{1, \dots, n'\}^2\}$ con dos asignaciones de valor posibles: 0 y 1. 1 describe que el agente realizó un recorrido directo partiendo de la ciudad i hacia la ciudad j con una distancia correspondiente de $c_{i,j}$; 0, por el contrario, indica el caso opuesto: no se viaja de i a j . En términos técnicos, el agente viajero habría ocupado (o no) el arco dirigido (i, j) .

Así, la primera expresión (4.1a) es la función objetivo: el recorrido total para arribar a las n' ciudades. (4.1b) exige que de cada ciudad se

parta hacia una única ciudad; por su parte la condición (4.1c), análogo a (4.1b), requiere que a cada ciudad se arribe desde un único punto de partida. La factibilidad del *tour* generado por el agente se termina de construir en el conjunto de ecuaciones (4.1), ya que sin éste podrían haber *tours* donde a cada ciudad se llega una única vez y cada ciudad parte hacia un único destino, pero dos o más subconjuntos resultado de la partición de $\{1, \dots, n'\}$ generan cada uno un recorrido o *tour* cerrado. Supongamos que se generara una solución óptima con las restricciones (4.1b) y (4.1c) cumplidas, pero que constara de tres *subtours*: T_1, T_2, T_3 con $T_1 \cup T_2 \cup T_3 = \{1, \dots, n'\}$. Poniéndonos imaginativos, esto querría decir que el agente terminó de recorrer T_1 , sugiriendo que el agente pasó de T_1 a T_2 sin un arco de conexión, algo inviable en un tour continuo, luego hizo lo propio de T_2 a T_3 . En este escenario se viola el supuesto de que el agente parte de un punto y termina en el mismo tras recorrer todas las ciudades. Resumidamente, la restricción (4.1) elimina la posibilidad de lo que llamaremos *subtours*. Sin embargo, al tratarse de una restricción que explora cada uno de los $2^{n'}$ posibles subconjuntos generables con $\{1, \dots, n'\}$ (también caracterizado como “conjunto potencia”), es una restricción con una elevada complejidad computacional asociada que debe simplificarse con adecuadas alternativas, como se hace en las restricciones de Miller, Tucker & Zemlin introducidas en 1960 en el artículo “*Integer programming formulations and traveling salesman problems*”. Sumado al planteamiento descrito, el TSP tiene como conjunto factible los $n'!$ posibles ordenamientos de $\{1, \dots, n'\}$, de manera que los programas que lo estudian operan sobre el campo de “algoritmos de optimización combinatoria”.

A continuación se retoma y explica el enfoque propuesto por Miller, Tucker & Zemlin (indistintamente también referido MTZ).

Introduzcanse $n' - 1$ variables enteras, $u_i, i \in \{2, \dots, n'\}$ que representan la posición / orden en que se visita la ciudad j , de tal manera que el mapeo de $i \neq 1$ a u es inyectivo. Se formulan las $(n' - 1)^2 - n'$ restricciones 4.2a, de subtour y las $n' - 1$ restricciones 4.2b, especificadas para variables de orden. Las cuales sustituyen eficientemente a la expresión que explora el conjunto potencia de ciudades pasando a explorar el producto cartesiano cuadrado de las mismas (el cual es significativamente más pequeño). Complementariamente, este enfoque permite relajar las restricciones de integralidad:

$$u_i - u_j + (n' - 1)x_{ij} \leq n' - 2, i \neq j \quad (4.2a)$$

$$1 \leq u_i \leq n' - 1 \quad (4.2b)$$

La idea detrás de estas restricciones es que, si el agente viaja de la ciudad i a la ciudad j entonces el orden en que se visita la ciudad j debe ser mayor al orden en que se visita la ciudad i . En caso contrario, se tendría $x_{ij} = 0$, por lo cual es trivial deducir que se cumple $u_i - u_j \leq n' - 1$, pues $u_i, u_j \in \{2, \dots, n'\}$. Por lo cual la restricción garantiza que existe una solución factible con un único ciclo hamiltoneano (no hay multiplicidad desubtours); no obstante, las restricciones no garantizan el óptimo global, de donde se derivan alternativas de mejora para fortalecer las restricciones, alcanzando con las mismas un mejor óptimo. Las restricciones MTZ son una alternativa que reemplaza el número exponencial de restricciones de eliminación de subtours por un número polinomial (del orden de $O(n'^2)$), lo cual representa un ahorro computacional sustantivo desde una formulación más compacta; esto, sin eliminar la complejidad

NP-Dura inherente al problema. Este reforzamiento de la restricción puede extenderse a problemas de ciclos hamiltonianos más generales (enfoque más allá del TSP), tal como desarrolla Desrochers (1991) en su artículo que estudia problemas de rutas vehiculares (VRP's por sus siglas en inglés) “*Improvements and extensions to the Miller-Tucker-Zemlin subtour elimination constraints*”.

Aun con esta reducción con las restricciones de Miller, Tucker & Zemlin, el ritmo en que escala el tiempo de procesamiento computacional el programa lineal (*i.e.*, la complejidad) vuelve necesario el ocupar heurísticas para hacer el cálculo con esfuerzos relativamente moderados. Esto ha sido comprobado tanto desde el enfoque empírico como desde la ciencia de la computación, demostrándose que no existe algoritmo alguno que resuelva el problema en tiempo polinomial; o, dicho en otros términos, se trata de un problema NP-Duro. En el artículo de Nilsson (2003) se explica un enfoque en la construcción de algoritmos para la resolución del TSP. Parte de los principios de *i)* construcción de tours, *ii)* mejora de tours, *iii)* tiempo de cómputo (asociado a la complejidad) y *iv)* la lejanía de la solución respecto a la cota inferior de *Held-Karp* (asociado a la optimalidad del algoritmo). La *Held-Karp* es el valor óptimo del problema TSP sin ocuparse las restricciones de integralidad en la solución, haciéndolo resolvible vía Simplex, o bien, aproximable con algoritmos. De tal manera, Nilsson discute los algoritmos con los cuales se construye una solución o tour factible, abordando su complejidad y distancias promedio resultantes respecto a la cota *Held-Karp*. Esta distancia varía desde el 25 % en el algoritmo de vecino más cercano hasta el 10 % en el algoritmo de Christofides.

Como segundo paso, se procede a mejorar el tour elegido en la anterior iteración, existiendo algoritmos de *2-opt*, *3-opt*, *k-opt* y sus

correspondientes derivaciones. Para el $2\text{-}opt$ se seleccionan dos ciudades del tour y se reconectan siempre y cuando el tour resultante tenga una mejor solución. Este movimiento puede ser también interpretado como un reverso a un segmento del tour inicial. La selección de dos ciudades para este reverso tiene la ventaja de solo contar con una posible combinación en la reconexión de nodos, tal como lo muestra el gráfico 4.1. Nilsson estima una desviación respecto a la cota de *Held-Karp* de 5% para el $2\text{-}opt$; para el $3\text{-}opt$ (el cual sigue una lógica similar pero con dos alternativas para la reconexión de nodos) se reduce a 3%. Si bien los grados de mejora a un tour realizados desde este enfoque son relevantes, el cómputo sigue siendo complejo pues cada iteración implica una evaluación de cada pareja posible asociable a cada ciudad. Una manera efectiva de reducir esta complejidad es cotejar a cada ciudad versus sus m vecinos más cercanos, con m' el hiperparámetro a elegir. Esta variación o aceleración del algoritmo $2\text{-}opt$ requiere cierto nivel de sensibilidad del problema para elegir adecuadamente la m y simplifica la optimización permitiendo reducir la complejidad del algoritmo por un grado polinomial (2 a 1), o de $O(n^2)$ a $O(m'n')$ (Nilsson, 2003).

Finalmente, el principio de mejora por $k\text{-}opt$ es extendible a una completa gama de variaciones que se le derivan, tales como el algoritmo de Lin-Kernighan, que elige la k óptima para cada iteración, o al algoritmo de búsqueda tabú, el cual asigna costos a optar por determinados movimientos en la elección de ciudades catalogados como “movimientos tabú”. Asimismo, hay enfoques de probabilidad en los plantamientos heurísticos que implican simulaciones aleatorias y replican comportamientos biológicos; algunos ejemplos son los Algoritmos Genéticos y la Optimización de Colonia Hormiguera.

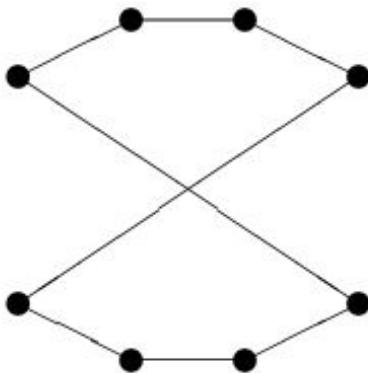


Figura 4.1. Única representación del movimiento *2-opt*

4.2. Arte con el agente viajero, o arte con línea continua

La segunda implementación con una imagen a emular es vía el TSP: parte del supuesto de que se cuenta con una imagen en blanco y negro, desplegado y expresado como escala de grises. Procedemos de la siguiente manera, basado en los pasos indicados por Bosch (2003) en su artículo “*Continuous line drawings via the traveling salesman problem*”

- Particionamos la imagen en rectángulos conformados por pixeles. Cada rectángulo será de k pixeles a lo largo por h pixeles a lo ancho. Los rectángulos se identificarán en función de sus coordenadas o posición: horizontal —o columnas— siendo n el total de éstos y vertical —o renglón— totalizando m . De tal manera que la matriz grande cuenta con dimensiones $\mu' \times \nu'$ con $\mu' = k m$ y $\nu' = h n$; asimismo, obtenemos una escala de grises para cada pixel en el rango discreto de 0 a 255 (empleando la

conocida convención de 256 escalas mencionada en 3.3.2).

- Obtenemos la escala de gris promedio de cada rectángulo y fijamos un parámetro que ponderará el nivel de saturación máxima de ciudades viable para cada rectángulo: $\gamma \in [4, 9] \cap \mathbb{N}$ ¹. De tal manera, para el rectángulo $(i, j) \in \{1, \dots, m\} \times \{1, \dots, n\}$ la escala de gris del pixel promedio será de ψ_{ij} , con valores entre 0 y 255. Entonces constrúyase

$$g_{ij} = \lfloor \gamma - \gamma \psi_{ij} / 255 \rfloor \quad (4.3)$$

Ésta será una escala de oscuridad del rectángulo y el parámetro γ introducido por Bosch implica variaciones en el grado de detalle de la imagen a desarrollar o saturación de ciudades.

- Dividimos la imagen en $m \times n$ rectángulos; en cada uno situamos uniformemente g_{ij} ciudades. El conjunto total de ciudades construidas será de cardinalidad τ , definida como

$$\tau = \sum_{i=1}^m \sum_{j=1}^n g_{ij} \quad (4.4)$$

- Computamos las distancias (o “costos”) entre las τ distintas coordenadas en una matriz $C = [c]_{ij} \in \mathbb{R}^{\tau \times \tau}$. Resolvemos el TSP sobre estas distancias (se recomienda usar alguna heurística para hacer el problema escalable). Se conectan las ciudades por aristas

¹Este rango de $[4, 9]$ no tiene otra justificación más allá de la visual, *i.e.* se define a la luz de ejecuciones empíricas del planteamiento y las determinó Bosch en el artículo. Cabe destacar que valores superiores a 9 podrían generar imágenes sobre-saturadas de puntos

dando lugar a una representación de aristas en el plano cartesiano; este es el resultado final.

Siendo que ahora el planteamiento se esboza sobre τ ciudades podemos entender que ahora la n_* introducida en la sección anterior tiene un determinado valor en el problema del arte con línea continua $n' = \tau$.

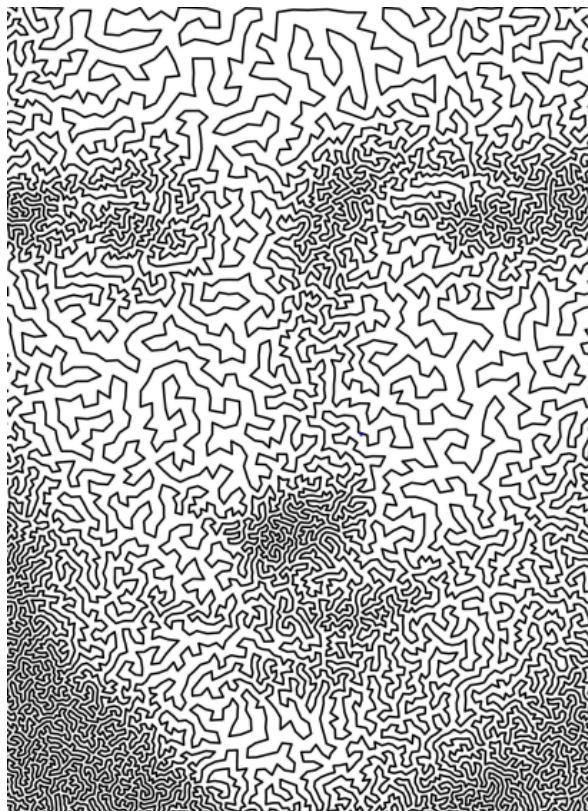


Figura 4.2. Rostro de la Monalisa reproducido por Robert Bosch usando el problema del agente viajero

Capítulo 5

Implementación y resultados

5.1. Partición de la matriz en rectángulos

Una imagen en escala de grises se entiende como un despliegue visual en el que cada unidad básica mostrada en la pantalla —mejor conocida como pixel— se ve representada por una cifra que cuantifica la cantidad de luz (o información de intensidad lumínica) presente. Representando la mínima unidad al negro y, la máxima, al blanco. La escala tiende a volverse discreta, ya que contar con un continuo estándar de 0 a 1 (o 0 % a 100 %) representaría un reto computacional y de almacenaje considerable. Algunas imágenes en escala de grises tempranas cuantificaban hasta 16 valores, lo cual implicaba un almacenaje de 4 bits por pixel; en la medida de que la rama de la fotografía digital evolucionó, aumentó a hasta 256 intensidades, implicando un peso de 8 bits por pixel. Si bien actualmente hay variaciones adicionales con mayor o menor número de valores (también

discretos), los cuales se eligen en función de la aplicación de la imagen, un estándar recurrentemente usado es el de 256 valores. Por eso, y en aras de respetar convenciones, se usará esta escala discreta, o bien, dependiendo del caso, la escala continua de 0 a 1. Para pasar de la escala $[0, 1]$ a la escala $\{0, \dots, 255\}$ hacemos $e_{255} = \text{round}(255e_1)$; para pasar en sentido inverso se usará $e_1 = \frac{1}{255}e_{255}$.

En este apartado se describirá cómo se procede para formular y computar los dos enfoques para reproducir imágenes: el de TSP y el de los fotomosaicos, explicados brevemente en los capítulos 3 y 4. Para iniciar los ejercicios se ha decidido usar una imagen de la extinta red social Twitter (ahora llamada X tras ser adquirida por otro grupo empresarial), cuyo símbolo era un pájaro, y la cual se muestra en 5.1. Una librería / paquete del reconocido lenguaje *open source* de programación matemática, R, encargada de procesar y graficar las imágenes de tipo png tiene por nombre “png”. La función *readPNG(.)* lee la imagen de entrada y la transforma en un arreglo numérico con tres ejes: el primero representa la cantidad de píxeles por fila (μ), el segundo por columna (ν), y el tercero codifica la intensidad de cada componente cromática del modelo RGB (rojo, verde y azul). De los elementos extrapolados del primer componente de la dimensión tres (*i.e.*, rojo), *readPNG(.)[, , 1]*, se extrae una matriz $\mu \times \nu$, la cual consta de valores que conforman la escala de grises de la imagen en el intervalo continuo de 0 a 1. Para graficar dicha imagen en el cuadrante del plano cartesiano $x - y$ (delimitado al rango $[0, 1]$ para ambas dimensiones) se usará la función *grid.raster(.)* perteneciente al paquete “png”.

Como siguiente paso, se usan las dimensiones de pixeles, m y n , para delimitar el tamaño de la matriz que permita generar una partición de la misma en rectángulos que sea factible; es decir, que



Figura 5.1. Despliegue con *grid.raster(.)*

sean rectángulos de dimensiones idénticas y que, posicionadas estratégicamente, generen la matriz grande (o matriz total). Para ello analizamos residuos de divisiones con aritmética modular para pasar de (μ, ν) a (μ', ν') ¹.

Posteriormente, se convierte la matriz de la escala $[0, 1]$ a la escala $\{0, \dots, 255\}$. Luego, se fracciona la matriz en un conjunto de $m \times n$ rectángulos mediante la función *matsplitter(.)* de R, la cual recibe como entradas la matriz a fraccionar, la cantidad de renglones de pixeles por rectángulo, $k = \mu'/m$ y la cantidad de columnas, $h = \nu'/n$. *matsplitter(.)* entregará una lista de $m \times n$ elementos. Los elementos 1 a n de la lista consistirán en los primeros n rectángulos superiores de la matriz, los elementos $n + 1$ a $2n$ consistirán en el segundo grupo de n rectángulos superiores. Así sucesivamente, hasta llegar a los n rectángulos inferiores, enumerados $n \times m - n + 1$ a $m \times n$. Entendemos que este tipo de ordenamiento es de clase *rowwise* (o por renglón), ya que se va llenando renglón por renglón, empezando por los renglones de arriba y continuando hacia abajo. Cada renglón se llena de

¹Véase más detalle sobre las propiedades de la regla de asignación para generar esta transformación en apéndice

izquierda a derecha. El otro tipo común de ordenamiento es *columnwise*, o por columna². Conociendo esto, podemos calcular la luminosidad promedio de cada elemento del output de *matsplitter(.)* computado como un promedio aritmético del total de pixeles que compone a cada rectángulo. Dicho conjunto de luminosidades puede sintetizarse en un vector de dimensión $m \times n$, bajo la ya mencionada lógica del ordenamiento *rowwise*.

5.2. Fotomosaicos

5.2.1. Fotomosaicos en blanco y negro

Construcción del vector de costos y la Matriz Perfecta

Posterior a reproducir los pasos descritos en la sección 5.1, se procede a convertir la escala del conjunto $\{0, \dots, 255\}$ a $[0, 1]$ dividiendo entre 255 (ver primer párrafo de la sección referida). Llámese h al vector resultante de extraer estas escalas promedio de los rectángulos ordenados con base en el esquema *rowwise*. Supóngase que las dimensiones de las imágenes / mosaicos que compondrán el fotomosaico son cuadradas, de valor $pi \times pi$. Entonces, generaremos tres arreglos importantes: uno de parámetros de costos, dos de escalas de grises. El primero será uno numérico 3-dimensional de tipo $cf \times m \times n$, llámese A , con $A[f, i, j] = c_{fij}$, donde el lado derecho es el mismo ponderador de costos de la expresión (3.1a). Uno de los arreglos en escala de grises será de dimensión $m \times n$, tendrá por nombre mp , por “Matriz Perfecta”; esto es un eufemismo por tratarse de la mejor representación factible de una imagen con las dimensiones de

²Arreglo por columna, o *columnwise* donde se llenan las columnas de izquierda a derecha, cada columna se llena de arriba a abajo

cantidades de rectángulos, $m \times n$, que se ingresan como *input* (ya que sustituye cada rectángulo por el valor en escala de grises del pixel promedio en el mismo); el otro, será un arreglo de $(m \text{ pi}) \times (n \text{ pi})$ y será *MF*, por “matriz final”, por tratarse del arreglo que representa el resultado final del fotomosaico. Además, se contará con una dotación de $\sum_{f=1}^{cf} u_f$ imágenes de solo cf tipos, cada uno matricialmente denotado mediante Mg_f , o “matriz de *grayscale*” de la imagen tipo f .

Mediante un bucle se convierte la información plasmada en el arreglo A (de tres dimensiones) en un *data frame*, $df3$ (de una dimensión) en cuya construcción se involucran los campos siguientes: t, i, j, f, c' ³. En este esquema, cada t se asocia a exactamente una combinación (i, j, f) y a cada terna (i, j, f) le corresponde una sola t . Por tanto, hay un mapeo biyectivo $\eta(i, j, f) = t$ y su inverso, η^{-1} . Este remapeo atiende un pendiente importante, permitiéndose que la librería elegida, (*lpSolve*), para computar el problema de programación lineal, admita los argumentos aportados ya que el vector de costos, $c_{fij} = (b_f - \beta_{ij})^2$, debe indexarse a una única dimensión, c'_t . A continuación se describe cómo se llenan tanto el arreglo de costos, c_{fij} como la “Matriz Perfecta” la cual, cabe mencionar, es el arreglo que representa a un fotomosaico con costo o valor de la función objetivo igual a 0.

1. Tomamos el número de iteración, t , y le asignamos su correspondiente coordenada matricial, $(i(t), j(t))$ deducida por el esquema *rowwise* (previamente descrito).
2. Obtener la escala $[0, 1]$ de luminosidad del elemento (i, j) , h_t ; es decir, el valor promedio del correspondiente rectángulo.

³La forma genérica de dicho arreglo 2-dimensional se explica y se esboza con mayor detalle en el apéndice

3. Obtener la diferencia cuadrática de las luminosidades; es decir, la que hay entre h_t y cada elemento del vector de luminosidades promedio de las cf figuras disponibles, llámese $lumP$. Esto es hacer $[(h_t - lumP_f)^2]_{f \in F} \in [0, 1]^{cf}$. Para cada f , asignar el valor de este vector a los elementos $A[i, j,]$ del arreglo A .
4. Asignar el valor h_t a la posición (i, j) de la “Matriz Perfecta”, mp .

Es importante considerar que el orden del mapeo i, j, f a t también debe guiar la construcción de la matriz.

Adaptación a módulo *lpSolve* y redimensionamiento

La mayoría de módulos en el *open source* percibe a los costos de un programa lineal como un vector de una única dimensión; el planteamiento desglosado en el capítulo 3 plantea tres dimensiones. Por lo tanto hay que construir un arreglo que descomponga toda la información en uno solo pasando de A a $df3$, la construcción de este arreglo se ve explicada en el Apéndice (ver A.1). Llevado al universo de programación en R , y con ayuda de esta $df3$ podemos definir el modelo en un objeto de clase “*lp*”, auxiliándonos en el módulo *lpSolve*. Dicho objeto se basa en (a) un vector de pesos o costos, c' , (b) una instrucción de optimización, *io*: “minimiza” o “maximiza” (en este caso se minimiza), (c) una matriz RM de restricciones con dimensiones $re \times (m \ n \ f)$ que representa a las re restricciones existentes sobre las variables, (d) un vector de lados derechos, $rh \in \mathbb{R}^{re}$ con el valor contra el cual será contrastada la combinación lineal de x , $RM[t,] \cdot x$, (e) un arreglo de caracteres, *sign*, con el cual, $\forall t \in \{1, \dots, re\}$, la restricción t -ésima será $RM[t,] \cdot x \text{ ``sign}[t]'' rh_t$, con $sign \in \{<, \leq, >, \geq, =\}$.

Holgura y variabilidad: Construcción del vector de disponibilidades

Hay que construir un vector, u que cuantifique las dotaciones disponibles de las cf figuras. Estas dotaciones pueden responder a tantas clases de formulaciones como pueda imaginarse el programador; como una observación trivial pero conveniente, siempre será necesario definirlo tal que la región factible sea no vacía. El planteamiento de u aquí especificado es propuesta original del presente escrito. Para ello, nos valemos de dos parámetros: el primero, hd , de “holgura”, que nos dirá qué tantos más de los mosaicos estrictamente necesarios tendremos. Ejemplo: si $hd = 0.1$, tendremos aproximadamente 10% más de los mosaicos exactamente suficientes para llenar los mn espacios habilitados. Para garantizar la viabilidad del problema, hd debe ser estrictamente no negativo ya que, de no cumplir esto, no se tendrían los mosaicos suficientes para llenar los rectángulos habilitados, dando lugar a un problema con región factible vacía.

Si deséaramos que las disponibilidades de cada mosaicos fueran las mismas para todas, tendríamos $\lceil \frac{mn(1+hd)}{cf} \rceil$ imágenes de cada tipo. Sin embargo, para darle mayor variabilidad a estas disponibilidades introducimos el segundo parámetro vd , que denota variabilidad, y nos permitirá tener una dotación del mosaico f extraída aleatoriamente (utilizando distribución uniforme) de entre el siguiente conjunto de enteros consecutivos,

$u_f \in \{\lceil \frac{mn(1+hd)}{cf}(1 - vd) \rceil, \dots, \lceil \frac{mn(1+hd)}{cf}(1 + vd) \rceil\}$. A mayor $vd \in [0, 1]$, y suponiendo una correcta simulación aleatoria, mayor será la varianza de las disponibilidades de cada tipo de mosaico en el vector u . Implicando que habrá una mayor diferencia de dotaciones del tipo de mosaico más común (o poblado) versus el menos común.

Con lo mencionado hasta este punto se han generado los parámetros necesarios y suficientes para definir el modelo. Sigue discutir cómo se despliega la solución y evaluar la eficiencia de la misma.

Identificación de solución para cada rectángulo y despliegue del resultado

Si bien el problema no fue restringido a soluciones integrales, el resultado es entero y consta únicamente de 0's y 1's, lo cual soporta la afirmación de la sección 3.2 de que la matriz de restricciones es unimodular. Si extraemos únicamente los valores $T = \{t|x_t = 1\}$ se valida que las $m \times n$ combinaciones (i, j) son cubiertas exactamente una única vez cada una y que cada figura en F se ocupa en una cantidad de ocasiones que no excede su correspondiente límite definido en el vector u . Habiendo supuesto que la dimensión de las cf figuras es de $pi \times pi$ y sea $MG_f \in \mathbb{R}^{pi \times pi}$ la matriz de escala de grises de la figura f , entonces procedemos a llenar el fotomosaico, o matriz final MF , mediante el siguiente *loop*: a través de $m \times n$ iteraciones hacer lo siguiente únicamente a través de las soluciones $\kappa \in T$.

1. Extraemos el κ -ésimo elemento. Sobre éste se asigna una (i, j, f) correspondiente, definida en función de η^{-1} , el mapeo biyectivo mencionado en 5.2.1.
2. Ahora hay que definir la posición de la figura, f en la matriz final. Definimos el pixel superior izquierdo de la figura f sobre la misma, $MF[rI, cI]$. Sígase que $rI = (i - 1)pi + 1$ y $cI = (j - 1)pi + 1$
3. Sigue definir al resto de la figura que llenará la posición (i, j) . Se asigna el valor de MG_f a la submatriz-rectángulo de pixeles con el

extremo superior izquierdo en (rI, cI) ; esto es, $MF[\{rI, \dots, rI + pi - 1\}, \{cI, \dots, cI + pi - 1\}] = MG_f$

Se puede validar que este procedimiento llena por completo y con valores factibles a la matriz MF , y de manera óptima desde el enfoque de la distancia en escalas de grises.

El siguiente paso es ejecutar los pasos descritos hasta este punto de la sección usando una gama de mosaicos de dimensiones idénticas $pi \times pi$ y buscando emular el logo de Twitter. Para ello, se generó un código de R que lee todos los archivos png de una subcarpeta y extrae sus dimensiones (pixeles). Posteriormente, elige la mínima dimensión (de entre columnas y renglones) del conjunto de imágenes, acota a todas las imágenes de manera que cumplan con esta dimensión a lo largo y ancho y guarda su luminosidad promedio. Además, define un intervalo de luminosidad, $i = 1/cf$, con $cf \in \mathbb{N}$ sobre el cual se seleccionará la imagen. Esto es, si $i = 0.05$ entonces las figuras se clasificarán en $cf = 20$ categorías, a saber, $\{[0, 0.05), [0.05, 0.10), [0.10, 0.15), \dots, [0.90, 0.95), [0.95, 1]\}$. Para el primer ejercicio extraemos las imágenes de un conjunto de 60 fotos de una cuenta de Instagram. Al computar su luminosidad se encuentra que las 20 categorías no son todas cubiertas, estando las imágenes acotadas en luminosidades entre 0.30 y 0.75. Para hacer que la cantidad de categorías cubiertas sea de 20 se oscurecen o iluminan tantas más imágenes como categorías falten por cubrir. El resultado de hacer la selección de 20 imágenes, ordenadas por luminosidad de menor a mayor, se muestra en 5.2.

Habiendo seleccionado las imágenes, se procede a reproducir el ejercicio de los fotomosaicos. Al hacerlo con $hd = 0.3$ y $vp = 0.3$, podemos desplegar la imagen resultante con dimensiones de

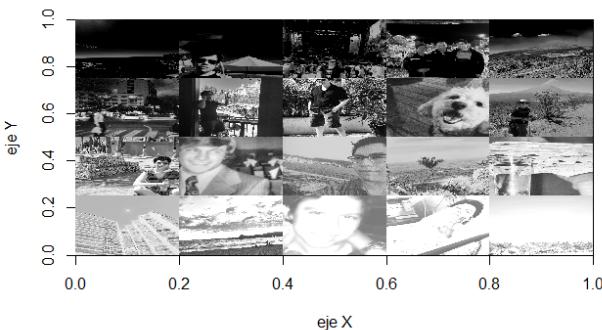


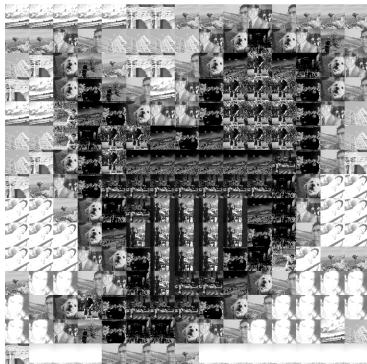
Figura 5.2. Catálogo de imágenes seleccionadas para el fotomosaico

rectángulos 15×15 y 20×20 en la figura 5.3 usando la función *grid.raster()*.

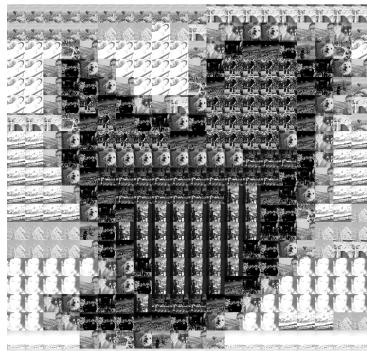
Para instancias con dimensiones mayores también hubo despliegues, los cuales se encuentran en 5.4.

Claramente el resultado ha mostrado ser visualmente eficiente. La matriz *mp* sirve como referencia para ver cómo se vería la imagen si se deseara desplegar la mejor representación posible de la imagen original contando exactamente con dimensiones de pixel $m \times n$. El resultado se muestra en figura 5.5.

Procedemos a evaluar la complejidad computacional del algoritmo de resolución del programa lineal. Para ello, registramos el tiempo que lleva resolver para la función *lp()* una determinada imagen conforme se agregan más dimensiones. Cabe recalcar que los tiempos registrados contabilizan el esfuerzo del algoritmo de *lp()* y no los procesamientos (que pueden ser bastante robustos) para construir las matrices y vectores que definen el problema, así como para desplegar el resultado final con



(a) Fotomosaico: 20 tipos de figuras, 15×15 mosaicos

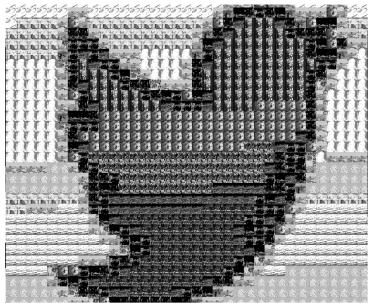


(b) Fotomosaico: 20 tipos de figuras, 20×20 mosaicos

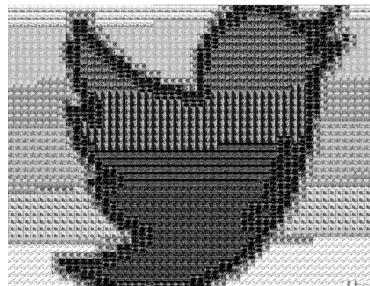
Figura 5.3. Soluciones gráficas al problema de los fotomosaicos con $dp = 0.3$ y $vp = 0.3$; dimensiones 15×15 y 20×20

grid.raster().

Las gráficas 5.6 y tablas reflejan la relación que tiene la dimensión (suponiendo que hay tantos rectángulos a lo largo como a lo ancho) con el tiempo de cómputo requerido para generar la imagen de Twitter; también, la que hay entre el total de rectángulos (la dimensión al cuadrado) y el tiempo. Vale la pena mencionar que con la función *lpSolve*, *R* abortó al llegar a la ejecución de dimensión 60; sin embargo las gráficas desplegadas y sus estimaciones lineales, cuadráticas y cúbicas dan a entender que la ejecución es cuando menos de orden 3 en el número de rectángulos ingresados como parámetro objetivo. Esto se explica por las gráficas de *fit* (en particular 5.6b), donde es eviente que el comportamiento dista de ser lineal y tiene un ajuste cuadrático extraño (el que tenga su vértice en valores de rectángulos “negativos” es absolutamente contraintuitivo). No es sino



(a) Dimensiones 35×35

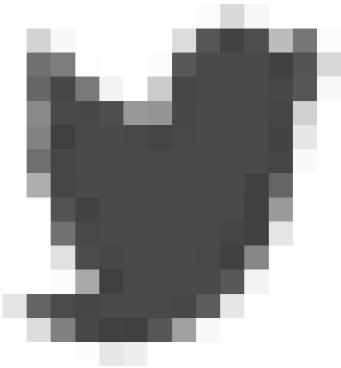


(b) Dimensiones 60×60

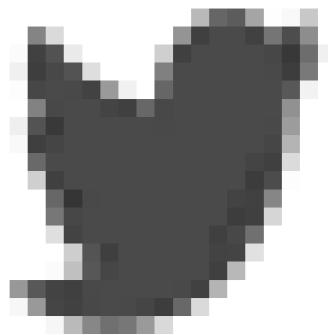
Figura 5.4. Soluciones gráficas al problema de los fotomosaicos con c; dimensiones 35×35 y 60×60

hasta el estimado cúbico que se distingue una relación asociable. Se tiene entonces complejidad $O(n^3)$.

La tabla de valores óptimos de la función objetivo, z^* , introduce una interrogante interesante; ¿qué tan bueno o malo es el valor de la función objetivo en el óptimo, es decir z^* , para las instancias trabajadas? La respuesta para estos ejercicios depende: es natural pensar que a mayor sea la cantidad de rectángulos a reproducir, más amplio será el rango de valores de la función objetivo ya que cada “costo” de desplegar el mosaico f en el rectángulo (i, j) es una diferencia cuadrada de valores, ambos entre 0 y 1, teniendo por lo tanto un valor máximo de 1. Por lo tanto, el rango factible de valores de la función objetivo es de 0 al número total de rectángulos, mn . De tal manera que, como podría intuirse, el valor óptimal de la función objetivo muy probablemente subirá en la medida que se amplían las dimensiones del problema. Sin embargo, el determinar si ello representa una buena o mala reproducción dependerá de la óptica con la que se aborde el problema. Si se toma la z^* resultante contra el peor



(a) Matriz perfecta,
dimensiones 15×15

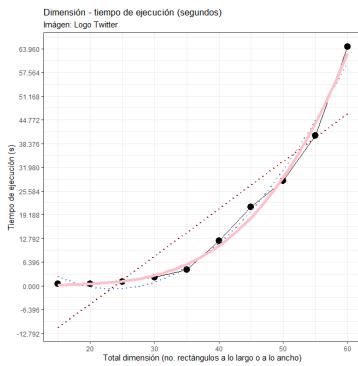


(b) Matriz perfecta,
dimensiones 20×20

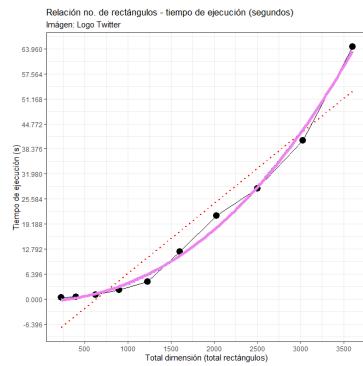
Figura 5.5. Matrices perfectas basadas en luminosidades promedio de submatrices-rectángulo

escenario con valor supremo de la función objetivo (aquél donde se iguala el costo al número total de mosaicos) y se mide ello como porcentaje, puede llevarse a cabo un ejercicio más equiparable con resultados evaluados en función de la dimensionalidad del conjunto de rectángulos. Visto desde esta métrica, en la cuarta columna se puede observar que para las diez instancias corridas de dimensionalidad variada, el resultado óptimo se encuentra entre 2.5% y 3.5% de su máximo posible.

Una cuestión interesante de abordar es el atributo de escasez de mosaicos en el problema, el cual se concentra en el vector de disponibilidades, u . Inspeccionando visualmente la imagen original, o inclusive la versión de mp de la misma, se puede observar que todos los contornos de la imagen (alrededor del pájaro) tienen la máxima luminosidad disponible (es decir, color blanco). Esto incluye, naturalmente, a los extremos inferior derecho y superior derecho. Sin



(a) Dimensión vs tiempo



(b) Rectángulos vs tiempo

Figura 5.6. Graficaciones tiempos de ejecución

embargo, al ver el fotomosaico en la figura 5.3, se observa claramente que las celdas inferiores derechas tienen una luminosidad muy elevada (cercana a 1), en tanto que la luminosidad de las superiores derechas es alta, pero claramente menor. Esto ocurre porque, si bien pudiera ser deseable que todos los contornos tuvieran el máximo de luminosidad, existe una disponibilidad limitada del mosaico más luminosidad, así como de todos en general. Por lo tanto, se recurre a otros mosaicos, lo menos oscuros posibles, que minimicen la suma de diferencias cuadráticas de luminosidades para cubrir el contorno del ave sin exceder los recursos existentes. Haciendo uso de los parámetros introducidos y explicados en el presente escrito, podríamos hacer el ejercicio sin limitar estos recursos. Basta indicar que $hd = m \cdot n$, queriendo decir que, si tuviéramos dimensiones $m \times n = 15 \times 15$, entonces los mosaicos disponibles son $15^2 = 225$ y, si asignamos $hd = 225$, entonces dispondremos de 22500% más de los mosaicos estrictamente necesarias. Además, si $vd = 0$ las disponibilidades se distribuirán homogéneamente, con lo cual garantizamos que cada tipo

Dimensión $n = m$	Total de mosaicos	Tiempo de ejecución (s)	Variables	Restricciones
15	225	0.45	4,500	245
20	400	0.52	8,000	420
25	625	1.12	12,500	645
30	900	2.36	18,000	920
35	1,225	4.42	24,500	1,245
40	1,600	12.12	32,000	1,620
45	2,025	21.28	40,500	2,045
50	2,500	28.29	50,000	2,520
55	3,025	40.49	60,500	3,045
60	3,600	64.45	72,000	3,620

Cuadro 5.1. Resultados de ejecución para diferentes tamaños de mosaicos.

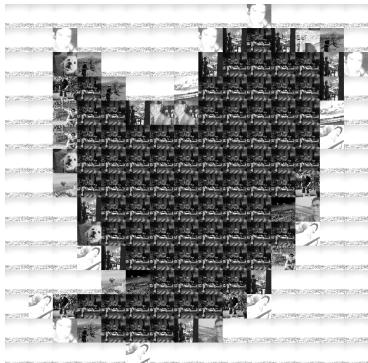
de mosaico, f , tiene el potencial de llenar por sí mismo todo el fotomosaico. Cabe mencionar que esta reparametrización tiene una alternativa: eliminar las restricciones del tipo definido en las expresiones 3.1a; es decir, las que usan al vector u . Hacer esto implicaría reducir la dimensionalidad de la matriz de restricciones, conllevando inherentemente ahorros en el cálculo de soluciones; sin embargo no se hará así, porque eso cambiaría drásticamente la estructura del código de la solución computada y podría perderse la sensibilidad del lector sobre la importancia de la holgura de disponibilidades que este escrito busca enfatizar al describir casos extremos. En 5.10 se muestra el resultado de hacer $hd = m n$ y $vd = 0$.

Claramente, estas imágenes tienen contornos más blancos y los fotomosaicos se asemejan a los resultados de usar la “Matriz Perfecta”,

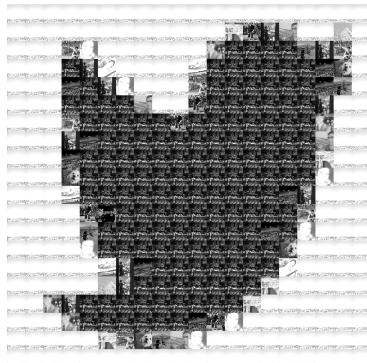
Dimensión $n = m$	Total de mosaicos	Valor óptimo, z^*	Valor z^* vs valor max
15	225	6.77	3.0 %
20	400	10.3	2.6 %
25	625	20.3	3.2 %
30	900	26.3	2.9 %
35	1,225	29.9	2.4 %
40	1,600	40.4	2.5 %
45	2,025	64.4	3.2 %
50	2,500	88.3	3.5 %
55	3,025	99.6	3.3 %
60	3,600	117	3.3 %

Cuadro 5.2. Valor óptimo z^* y su comparación porcentual con el valor máximo por varias dimensiones de rectángulos.

5.5, puesto que el costo total de diferencias cuadradas de luminosidades (valor de la función objetivo) se reduce al relajar las restricciones. Para estas instancias tenemos resultados de distancias de 0.30 y 0.59 representando en ambos casos 0.13 % y 0.15 % del supremo del rango de la función objetivo. Recordemos que en ejercicios previos con holguras de disponibilidad más amplias estos porcentajes eran del orden del 3 %; es decir, con niveles hasta 20 veces menores. De esta modificación estratégica de las restricciones realizadas en ejercicios de mismas dimensiones podemos dilucidar la sensibilidad del algoritmo diseñado y sus variaciones paramétricas sobre la calidad visual y cuantitativa de los resultados. No obstante lo anterior, es claro que no todos los 20 tipos de mosaicos disponibles se están empleando, lo cual es indicador de que la imagen carece notablemente de variedad en



(a) Fotomosaico: 20 tipos de figuras, 15×15 mosaicos



(b) Fotomosaico: 20 tipos de figuras, 20×20 mosaicos

Figura 5.7. Soluciones gráficas al problema de los fotomosaicos con $dp = m n$ y $vp = 0$; dimensiones 15×15 y 20×20

escalas de intensidad lumínica. Esto nos podría hacer perder interés en esta clase específica de instancias.

Finalmente, consideramos que vale la pena explorar la posibilidad de construir una imagen más elaborada, con mayor relieve y color que el logo de Twitter. Para ello, se propuso reproducir el ejercicio de los fotomosaicos con una imagen caricaturizada del connotado científico británico, Isaac Newton, la cual se puede encontrar en 5.8.

En ejercicios anteriores ya se demostró que el algoritmo computado para este trabajo soporta composiciones de imágenes de hasta 60×60 rectángulos. Por ello, se optó por una asignación poco por debajo de ese límite, de 50×50 . Reproduciendo de nueva cuenta la escala de gris del promedio en cada rectángulo, se genera el resultado de la matriz objetivo (o “Matriz Perfecta”).

En 5.10, se tienen dos resultados de nueva cuenta, de forma similar

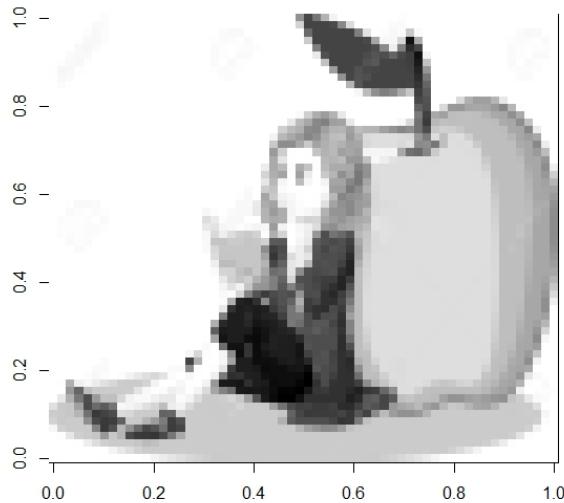


Figura 5.8. Imagen objetivo: caricatura de de Isaac Newton

a como se trabajó con la imagen del pájaro. Un ejercicio con holgura indefinida en las dotaciones de mosaicos y otro ejercicio que exige variabilidad en las asignaciones de mosaicos sin importar que exista un mosaico geométricamente más cercano a un dado rectángulo sobre el espacio de grises.

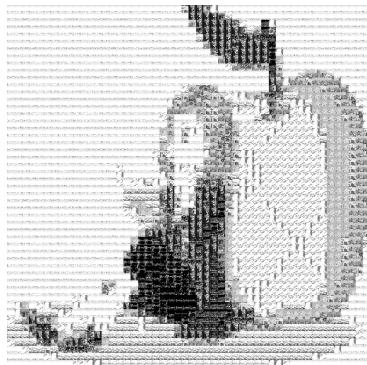
5.2.2. Fotomosaicos a color

En el capítulo de planteamiento del problema de los fotomosaicos se explicaba que, para la aplicación a color, el concepto en términos de programación lineal es idéntico y los únicos retos puntuales son: la elección de variedad de mosaicos a elegir, ya que éstos representan una dimensión superior al del escala de grises en proporciones cúbicas; la partición las mn submatrices-rectángulo en tres dimensiones (es decir, lo que hace la previamente referida función *matsplitter(.)* de *R*) y el posicionamiento del resultado final haciendo yuxtapuestos varios

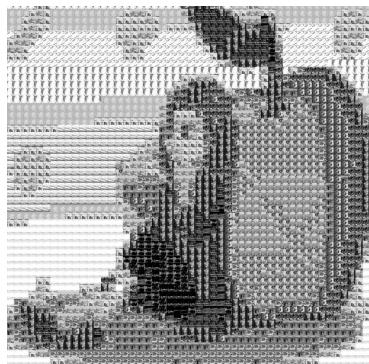


**Figura 5.9. Fotomosaico de Isaac Newton, Matriz Perfecta
50 × 50**

arreglos matriciales de dimensión tres. Sobre todo esta primera diferencia hace que no sea tan trivial el problema de reproducir el ejercicio de los fotomosaicos con cualquier imagen. Se trabaja entonces partiendo de un ejercicio relativamente limitada en cuanto a colores, digase la bandera de México, y a raíz de eso, selecciónense los colores a ocupar. La propuesta iniciada de la foto enfocada de una flor con fondo de intensidad medida en la figura 3.3 se extiende a otros colores perceptibles en la bandera nacional: rojo, verde, café; complementariamente añádese un rosa y un anaranjado para verificar que el algoritmo acierte al reservar su uso, ya que a simple vista no parecieran figurar como colores en el símbolo patrio. En 5.11, está la



(a) Fotomosaico Newton:
20 tipos de figuras,
Disposiciones no limitadas



(b) Fotomosaico Newton:
20 tipos de figuras,
Disposiciones limitadas

Figura 5.10. Soluciones gráficas al problema de los fotomosaicos con dimensiones 50×50 : Versión con holgura de $dp = m n$ y $vp = 0$; y sin holgura de $dp = 0.3$ y $vp = 0.3$

bandera a reproducir.

En 5.12 se desglosan los respectivos colores dominantes generados al reproducir k-medias con $k = 2$ sobre una selección manual de imágenes de flores realizada.

Habiendo seleccionado la imagen a computar y los mosaicos, se procede a generar el algoritmo de fotomosaico. Vale la pena sugerir empezar con dimensiones pequeñas, diganse m y n con $n > m$ ya que la bandera tiene una orientación horizontal. Así pues, sea $m = 15$ y $n = 20$, y posteriormente $m = 20$ y $n = 30$, se generan los resultados de 5.13.

La imagen se ha computado con éxito. Como se preveía, las asignaciones están concentradas principalmente en los tres colores especificados que integran la bandera (verde, blanco, rojo y café). Al



Figura 5.11. Imagen objetivo: bandera de la República Mexicana

computar el caso de 20×30 se observa que las flores naranjas están desplegadas en 3 ocasiones; esto se explica porque las fronteras entre dos objetos de distinto color tienen centroides que pueden estar cerca de un tercer color en el hiperplano RGB, tal es el caso de dos rectángulos ubicados entre el verde de la encina, el laurel y el nopal y el café del águila; así como un tercer rectángulo en el listón tricolor en la parte inferior del laurel. El rosa se aleja de las posibilidades de ser seleccionado por el modelo, en la medida de que es un color distante de los que componen la bandera.

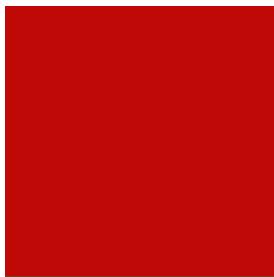
Extiéndase el ejemplo de la visualización de Isaac Newton a la versión a color. Las versiones degradadas de la imagen original, en sus versiones de 40×40 y 30×30 se presentan en 5.14, con dos despliegues que son fácilmente identificable.

El primer ejercicio se realiza con la instancia de 30×30 con las mismas flores que en el ejercicio de la bandera nacional (ver imagen 5.13) y los resultados se muestran en 5.15

Con algo de imaginación se puede identificar que el desarrollo de



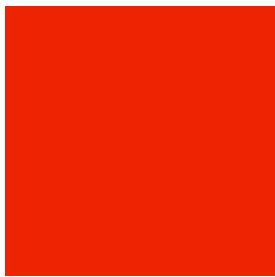
(a) Flores verdes



(b) Flores rojas



(c) Flores cafés



(d) Flores naranjas



(e) Flores rosas

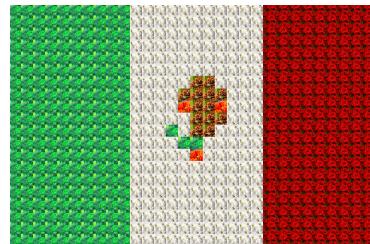
Figura 5.12. Desglose de color promedio o predominante de 5 tipos de flores que, junto con las de tipo blanco, conforman los mosaicos a asignar de forma óptima

esta instancia da un buen resultado. Ajusta el verde de la hoja de manzana con el verde, el rojo de la manzana con las flores rojas y naranjas y las flores cafés se posicionan en donde iría el pelo castaño del personaje. Sin embargo, no deja de ser un ejercicio limitado ya que la variedad disponible es de solo seis colores. De tal manera que se amplió la gama de flores de seis a veinte y se generaron las dos correspondientes ejecuciones manipulando los parámetros que definen al vector u , *i.e.*, vd y hd : la ejecución con disponibilidades de máxima holgura y la de variación forzada en el uso de las distintas fichas.

Se observa en los resultados en 5.16 que ambos ejercicios, en los



(a) Dimensiones 15×20



(b) Dimensiones 20×30

Figura 5.13. Resultado de fotomosaicos a color: bandera de la república mexicana a base de flores



(a) Dimensiones 30×30



(b) Dimensiones 40×40

Figura 5.14. Resultado de arreglos objetivos, versión Newton



Figura 5.15. Imagen objetivo: Newton en dimensiones 30×30

cuales se agregaron nuevos mosaicos, ocupan naturalmente una mayor variedad de éstos que en las instancias anteriores de Newton. Resulta llamativo cómo en el primer ejercicio el algoritmo identificó el color del pétalo con las flores grises más que con las propias verdes claras con las que sí se asociaban en el ejercicio 5.15, con dimensiones 30×30 y una variedad más limitada de mosaicos disponibles. Esto se debe a que el tipo de gris que representa al pixel “promedio” de la flor gris, se encuentra cercano a muchos colores, entre ellos este verde. El segundo ejercicio tuvo como resultado una seriación de colores variados plasmados en el fondo que originalmente es color blanco. El hecho de que se colocaran yuxtapuestas bajo un patrón horizontal siendo que no existe similaridad con el blanco de algunos de estos colores (rosa, azul, vino)



(a) Dimensiones 40×40



(b) Dimensiones 40×40

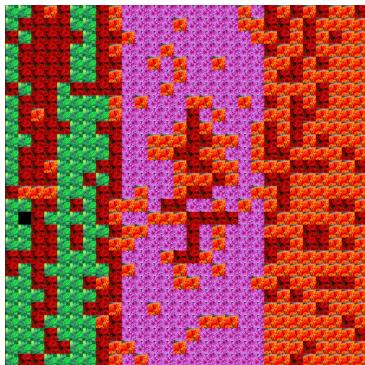
Figura 5.16. Resultado de arreglos objetivos, versión Newton

es un indicativo de que estas imágenes se ocuparon ahí como recurso bajo el objetivo de cumplir el empleo de los mosaicos de los veinte tipos de colores en similar cuantía, tal como exigen las restricciones del problema. Si bien el resultado es imperfecto, sí es visualmente atractivo e interesante visto como problema artístico y de programación. Además, analizando visualmente la imagen objetivo y los resultados, sí es factible relacionar que una imagen es un intento de replicar de forma creativa a la otra. El resultado en general de las instancias de Newton a color sí puede considerarse exitoso puesto que se produjeron imágenes con características similares.

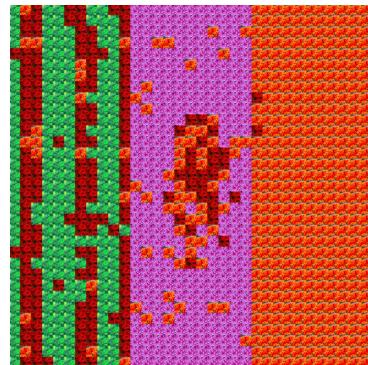
5.2.3. Fotomosaicos con librería de R

En el Capítulo 3 se menciona que existen módulos o librerías de lenguajes de programación *open source* con los cuales se pueden reproducir ejercicios con imágenes. Tal es el caso de R, que hace uso de su librería *RSimMosaic* para, vía el módulo *composeMosaicFromImageRandomOptim*, generar la imagen objetivo

con tres insumos mínimos: imagen objetivo, conjunto (folder) de imágenes-mosaicos a colocar y nombre del archivo jpg a entregar como salida - resultado. A diferencia del algoritmo que se definió y programó en este trabajo, esta librería se fija como objetivo usar tantas imágenes como sea posible. Además, iterativamente se ejecuta una acción sobre cada subrectángulo, aleatoriamente elegido, que consiste en asignarle una etiqueta mediante un algoritmo de clasificación supervisada llamado “k-vecinos más cercanos” donde cada elemento, o pixel, en el espacio 3-dimensional se asocia con sus k coordenadas más cercanas, cada una de éstas eligiendo por su categoría y optándose por el voto mayoritario. Tanto el componente de selección y clasificación aleatorio, como el hecho de que no se garantiza un resultado óptimo en cuanto a distancia euclidiana (u otra medida) hacen que este programa opere sobre una heurística, las cuales por definición buscan soluciones sub-optimales. Una de las ventajas de esto es que se puede desplegar una solución sin un costo computacional tan elevado. Una desventaja se da en cuanto a la solución ya que, como se puede observar, el resultado es menos bueno cualitativamente. De esto podría derivarse otro reto, el de identificar qué hiper-parámetros pueden ajustarse para hacer una mejor reproducción de la imagen de la bandera (número de vecinos, reemplazar mosaicos en el proceso de selección, etc). Vale la pena rescatar que este algoritmo procesa cada pixel de la imagen tal como se lo entrega el archivo jpg, siendo esta una tarea compleja en imágenes de más de 400×400 o mayor dimensión. De manera que tuvo que hacerse el ejercicio de reducción de dimensiones desarrollado anteriormente de agrupar un conjunto de pixeles, asignando manualmente dimensionalidades $m \times n$ y asignarles su color promedio (ver sección 5.1 donde se define la construcción del arreglo *matsplitter()* y los cálculos derivados). Haciendo esto, el algoritmo



(a) Dimensiones 30×30



(b) Dimensiones 35×35

Figura 5.17. Resultado de fotomosaicos a color: bandera de la república mexicana a base de flores

logra computarse rápidamente, generando imágenes tales como se muestra en 5.17.

Vale la pena notar que este algoritmo se fija como objetivo desplegar la mayor variedad posible de mosaicos haciendo que en la misma se haga uso de los colores naranja y rosa, que tienen poca relación con la bandera de México. Además la asignación es, como dice el nombre, aleatoria, implicando que sucesivas reproducciones del mismo código darán un resultado diferente.

5.3. Arte con el agente viajero

5.3.1. Arte con el agente viajero, caso Twitter

Tras haber seguido los pasos de la sección 5.1 lo siguiente consiste en convertir la escala del conjunto $\{0, \dots, 255\}$ a la escala $\{0, \dots, \gamma\}$, empleando la ecuación (4.3) desglosada en la sección 4.2 de la presente tesis. Así pues, obtenemos un vector de dimensión $m \times n$, llámese g .

Para el primer ejercicio, asígnese $\gamma = 7$. Bajo esta γ , las escalas primera a treintaiseisava de luminosidad coinciden con la escala 7 en oscuridad (o siete “ciudades”). Las siguientes luminosidades (por orden) corresponden a seis ciudades; sucesivamente, hasta llegar a la escala 255, que corresponde a cero ciudades. Es obvio que situar más puntos, o ubicar más aristas que conecten éstos en un rectángulo generará una mayor oscuridad en el mismo, de ahí que se considere a γ como una escala de oscuridad. Lo siguiente consta en ubicar las ciudades en el plano $x - y$, haciendo un loop a través de los $m \times n$ rectángulos, representados como matrices.

1. La asignación del k -ésimo elemento de la lista $\{1, \dots, m \times n\}$ a una coordenada de matriz (i, j) se rige por el siguiente pseudocódigo, el cual tiene como base el tipo de arreglo *rowwise*.

```

mod2 = kmod(n)
si mod2 = 0
entonces i = mod2
e.o.c.
i = mod2
j =  $\frac{(k-i)}{n} + 1$ 
```

2. Una vez que tenemos la asignación $k \mapsto (i, j)$, $(i(k), j(k))$, el siguiente paso es plasmar los puntos que asemejarán a la imagen objetivo sobre el rectángulo $[0, 1] \times [0, 1]$ en el plano $x - y$. Para ello, vale la pena considerar que el extremo superior izquierdo de la imagen corresponde al elemento $(1, 1)$ de la matriz y a la coordenada $(0, 1)$ del plano $x - y$. Partiendo de aquí, deducimos la posición del extremo superior izquierdo de la coordenada genérica (i, j) de la matriz. La coordenada x corresponde al pixel con la columna j , representando un paso de $\frac{(j-1)}{n}$ a la derecha de

$x = 0$. La coordenada y , con el renglón i , corresponde su ubicación en el eje y a un paso de $\frac{(i-1)}{m}$ unidades abajo de $y = 1$. Bajo esta formulación, derivamos que cada rectángulo (asociado a una submatriz) tendrá un ancho $\frac{1}{n}$ y un largo de $\frac{1}{m}$. En este esquema se puede demostrar que para todo par de coordenadas de matriz de tipo $\{(i, j), (i + 1, j)\}$ o $\{(i, j), (i, j + 1)\}$ existirá una relación de adyacencia entre sus correspondientes rectángulos. Además, las esquinas de la figura que envuelve a este conjunto de $m \times n$ rectángulos serán las coordenadas $\{(0, 1), (1, 1), (0, 0), (1, 0)\}$. Con esto, garantizamos que los rectángulos mencionados conforman una partición del cuadrado $[0, 1] \times [0, 1]$.

3. El siguiente paso es usar el valor entero $g_k = g[k]$ para colocar esa cantidad de ciudades en el rectángulo asociado a la coordenda $(i(k), j(k))$. Una alternativa inmediata es colocarlas aleatoriamente, lo cual muy probablemente permitiría reproducir la imagen con buena similaridad. Sin embargo, se propuso hacer una asignación de coordenadas más razonada que obedece a la idea de hacer una buena distribución de las ciudades a lo largo y ancho de su correspondiente rectángulo. Con una “buena” distribución se busca dar a entender que se quiere evitar una excesiva concentración de las ciudades en una zona del rectángulo, dejando a otras zonas muy blancas (o sin ciudades). La idea propuesta, pues, es generar un *grid* de coordenadas $\{x_0, x_1, \dots, x_{10}\}$ donde x_0 y x_{10} corresponden a los extremos izquierdo y derecho del rectángulo y $x_{l+1} - x_l = q \quad \forall l \in \{0, \dots, 9\}$ ($q = \frac{1}{10n}$); otro grid para las coordendas y : $\{y_0, y_1, \dots, y_{10}\}$, con $y_{p+1} - y_p = r \quad \forall p \in \{0, \dots, 9\}$ ($r = \frac{1}{10m}$) con y_0 y y_{10} extremos

superior e inferior. Con esto, ya habremos construido el *grid* 2-dimensional $G = \{(x_i, y_i) | (i, j) \in \{0, \dots, 10\}^2\}$ que consiste en 121 coordenadas. La propuesta principal es iterativa sobre el total de rectángulos. Para cada uno de éstos, sobre el conjunto de coordenadas generadas por el *grid*, aplicar un k-medias con g_k centroides - *clusters*. Por lo tanto en una “buena” asignación de ciudades en función de los criterios mencionados anteriormente, se pueden guardar las coordenadas de estos centroides e incorporarse los mismos a la selección de ciudades correspondiente al algoritmo k-medias aplicado sobre los rectángulos anteriores (si los hubiere).

Al iterar este proceso las $m n$ veces, se obtienen las $\sum_{k=1}^{m n} g_k = \tau$ ⁴ ciudades objetivo y podemos resolver el TSP sobre éstas. Los resultados de escribir todas las ciudades como puntos bajo los ejercicios con 15×15 rectángulos (606 ciudades) y con 50×50 (5,765 ciudades) se desglosan en la imagen que se nombra “puntillismo”, 5.18, por su similitud con la técnica artística que despuntó de manera importante en Francia a inicios del siglo XX y consiste en generar una obra con puntos diminutos (el tradicional puntillismo se trabaja sobre lienzo y con óleos de colores).

Existe un tipo de archivo cuya extensión es “.tsp”. Éste se basa en el conjunto de ciudades sobre el que se ha de resolver una instancia del problema lineal. Dicha solución puede procederse mediante distintos métodos heurísticos y puede realizarse mediante diversos *softwares*, tales como R, Python, Java y Concorde. En aras de practicidad, y en vista de que hasta el momento la construcción de los rectángulos y las ciudades se ha llevado a cabo con R, se usará este lenguaje para

⁴esta es una expresión equivalente a la ecuación ec.(4.4) , pero en la versión que mapea $(i, j) \rightarrow k(i, j)$

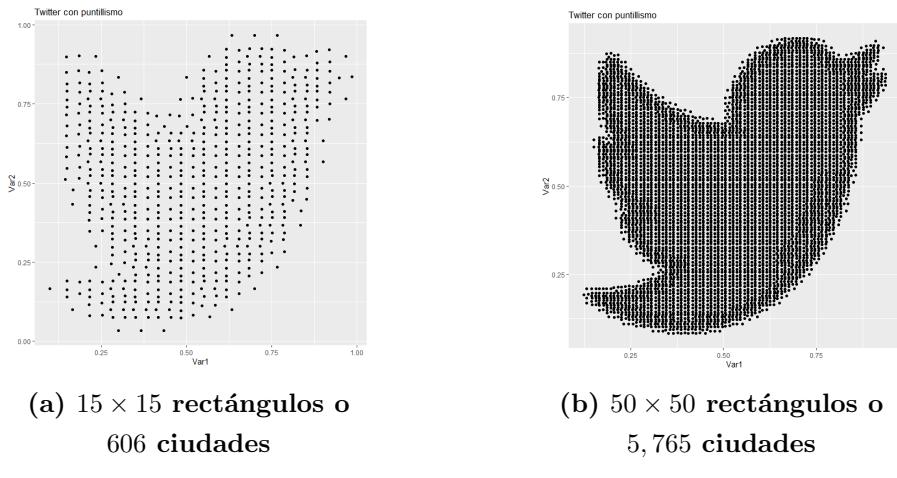


Figura 5.18. Gráfica de los puntos sobre el cuadrante $x - y$

manipular tanto objetos como archivos de tipo TSP. Para este efecto se usará la librería homónima cuya función, *ETSP(.)* debe su nombre a *Euclidean TSP* y convierte el *data frame* de 2 columnas (una por coordenada) y τ renglones (uno por ciudad) en un problema del agente viajero. El *Euclidean* obedece a que, en otros potenciales ejercicios del agente viajero, las coordenadas podrían ser de otra métrica sobre un espacio normado o incluso, sobre coordenadas terrestres. La función encargada de resolver un ETSP es *solveTSP(.)*. Si bien se le puede ingresar el método o algoritmo a emplear como entrada de la función, el que se usa por *default* es el algoritmo de inserción arbitraria con refinamiento de tipo *two opt*. Como dice el término “heurística”, el *two opt* no garantiza un máximo global pero sí una ejecución rápida y localmente óptima de la solución. El tipo de datos del *output* es doble: uno es de la clase *tour* y otro es entero. La salida entera consiste en el orden de las ciudades que se anidan para construir el *tour*. Con ello se puede graficar la solución con ayuda de *ggplot(.)*. Los fotomosaicos

resultantes con dimensiones 15×15 y 50×50 se despliegan en 5.19.

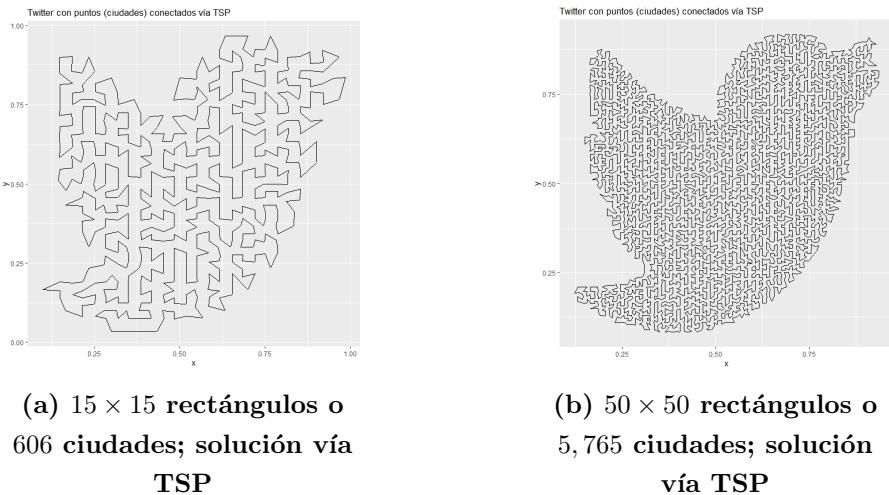


Figura 5.19. Aristas unen ciudades eficientemente, heurística 2-opt, cuadrante $x - y$. Ciudades colocadas con k-medias

En el punto 3 de los pasos del instructivo en la subsección 5.3.1 se describe la lógica de la disposición de ciudades determinística y regular, garantizando una cobertura equitativa del espacio $[0, 1] \times [0, 1]$. Aquí hay una buena dispersión en el sentido de que cada región del rectángulo tiene una representación proporcional. Sin embargo, alternativamente se puede contemplar el enfoque de aleatorizar las coordenadas de puntos del rectángulo, respetando las escalas de color definidas por rectángulo (y por lo tanto, manteniendo el número de ciudades). Para ello se ocupó la función de distribución uniforme desarrollada por R , $runif()$ dos veces por rectángulo: una vez para la dimensión x y otra para la dimensión y . Los resultados aleatorios de puntillismo y conexión de ciudades para dimensiones 15 y 50 con 606 y 5,765 ciudades respectivamente se despliegan en los

conjuntos de imágenes 5.20 y 5.21.

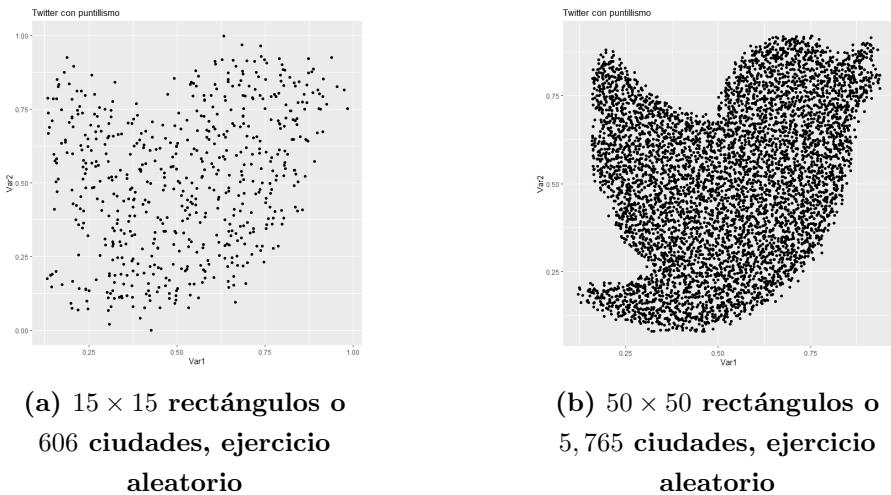
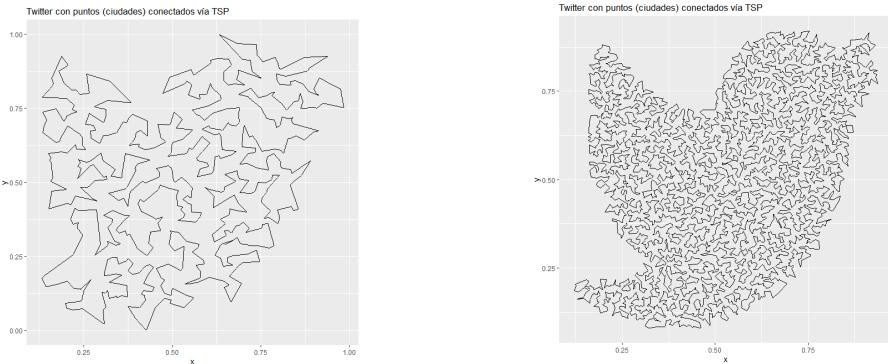


Figura 5.20. Gráfica de los puntos sobre el cuadrante $x - y$, ejercicio aleatorio

El nivel de detalle en la dimensión 50 da pie a que la calidad y propiedades de los ejercicios sean visualmente similares entre los casos de dispersión regular regida por k-medias frente al de distribución aleatoria (contrastar imagen 5.19b contra 5.21b). La diferencia importante se detecta al comparar ambos tipos de planteamientos con dimensión 15: se observa cómo la conexión entre puntos en el caso aleatorio no tiene una forma visiblemente tan bien definida; es decir, no es trivial deducir que se trata de una silueta de un pájaro, cosa que en el caso de dimensión 50 sí se identifica de forma más evidente (contrastar imagen 5.19a contra 5.21a).

Finalmente, como se ve en 5.22, el ejercicio se hace extensivo a la versión más grande que se logró correr del *TSP* con 85 rectángulos y 15,311 ciudades y donde el logo de Twitter es muy fácilmente



(a) 15×15 rectángulos o
606 ciudades, ejercicio
aleatorio; solución vía TSP

(b) 50×50 rectángulos o
5,765 ciudades, ejercicio
aleatorio; solución vía TSP

Figura 5.21. Aristas unen ciudades eficientemente, heurística 2-opt, cuadrante $x - y$. Ciudades colocadas aleatoriamente.

distingüible.

5.3.2. Arte con el agente viajero, complejidad y exploración de instancias más complejas

Un siguiente paso interesante consiste en evaluar la potencia computacional del solver del tsp en *R* que se ha ocupado (módulo *tsp*, función *solve_TSP(.)*) al registrar el tiempo que lleva resolver una determinada imagen conforme se agregan más dimensiones y, naturalmente, más ciudades.

La tabla 5.3 asocia la dimensión única ($m = n$) con el número de ciudades en la imagen de Twitter; asimismo describe la relación del número de ciudades con el tiempo de solución del TSP; claramente esta relación es de una complejidad superior a la lineal. Vale la pena mencionar que la columna de restricciones supone la técnica de

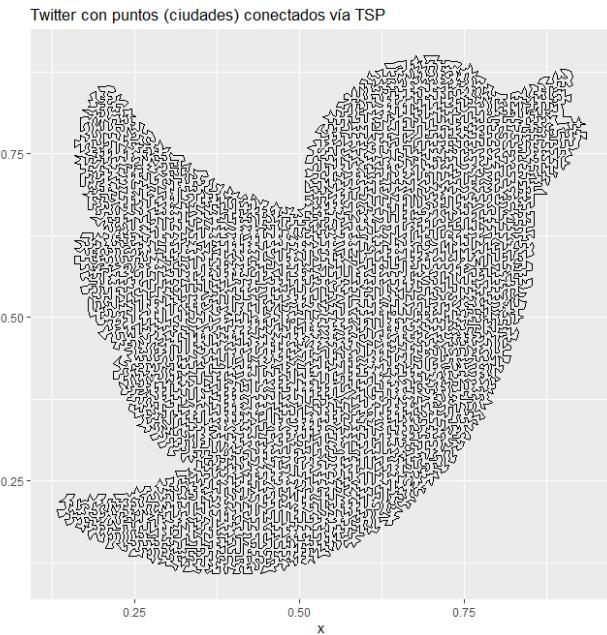


Figura 5.22. 85×85 rectángulos o 15,311 ciudades; solución vía TSP

reducción de restricciones a los subtours de Miller-Tucker-Zemlin que asume que son n^2 y no 2^n de estas restricciones. Si bien las restricciones de MTZ no se implementan en la heurística *2-opt*, se mencionan para dimensionar el tamaño del reto al que habría que enfrentarse en un escenario de búsqueda de soluciones exactas y no subóptimas.

Dimensión $m = n$	Total de ciudades	Tiempo de ejecución (s)	Variables	Restricciones
10	286	0.00	81,796	82,368
15	606	0.04	369,664	370,880
20	1,051	0.16	1,104,601	1,106,703
25	1,670	0.83	2,788,900	2,792,240
30	2,284	2.46	5,216,656	5,221,224
35	3,058	5.83	9,351,364	9,357,480
40	3,894	15.67	15,163,236	15,171,024
45	4,723	29.56	22,306,729	22,316,175
50	5,765	61.31	33,721,249	33,732,863
55	7,360	182.34	54,169,600	54,184,320
60	8,271	221.77	68,409,441	68,425,983
65	9,616	309.78	92,467,456	92,486,688
70	10,987	433.37	120,714,169	120,736,143
75	13,149	1,096.58	172,896,201	172,922,499
80	15,327	2,268.22	234,916,929	234,947,583
85	15,311	1,922.58	234,426,721	234,457,343

Cuadro 5.3. Tiempos de ejecución de la función solve_TSP(.) a las instancias definidas por la dimensión.

Algo que resalta al observar la tabla 5.3 es que el número de ciudades crece respecto a la dimensión a tasas crecientes en una curva “suave” hasta antes de pasar de 80 a 85; vale la pena observar la gráfica en 5.23 para concluir que es en este intervalo donde ésta empieza a manifestar quiebres. Éstos podrían deberse a que, en la medida de que el largo y ancho en rectángulos de la imagen se aproximan a los pixeles reales de la misma van a existir transiciones de dimensiones donde algunos pixeles reciben asignaciones de escalas de grises mayores en tanto que otros disminuyen. En el balance, a pesar de que se tienen más rectángulos, serán más los rectángulos que reducen su métrica de oscuridad que aquellos que la aumentan. Esto al grado de que en algunas variaciones de dimensión se disminuya el total de ciudades resultante (la gráfica muestra cómo eso ocurre tres veces en el rango analizado). Otra de las consideraciones es que, tal como se explica en el apartado de partición de rectángulos, 5.1, la dimensión original se va recortando a lo largo y a lo ancho, de tal manera que ambas sean divisibles entre el *input* entregado. Esto da lugar a que las matrices objetivo difieran en dimensión. Ejemplo, la dimensión original de la imagen de Twitter es de 450×550 en tanto que para ajustarse a dimensiones 80 y a 85 se tuvo que pasar a 400×480 y 425×510 respectivamente. Por consiguiente los rectángulos tienen formas, oscuridades y estructuras distintas contribuyendo a conjuntos de ciudades con cardinalidades poco predecibles.

Similar a lo realizado en la subsección 5.2.1 de implementación para fotomosaicos, 5.2.1, se reproducen las gráficas asociadas a la tabla anterior en 5.24. Se añaden los correspondientes ajustes polinomiales. A diferencia de el caso de los fotomosaicos donde bastaban los ajustes cúbicos, en este ejercicio fue hasta llegar a los ajustes de orden 6 y 4 para dimensión y rectángulos respectivamente, que se logró un ajuste

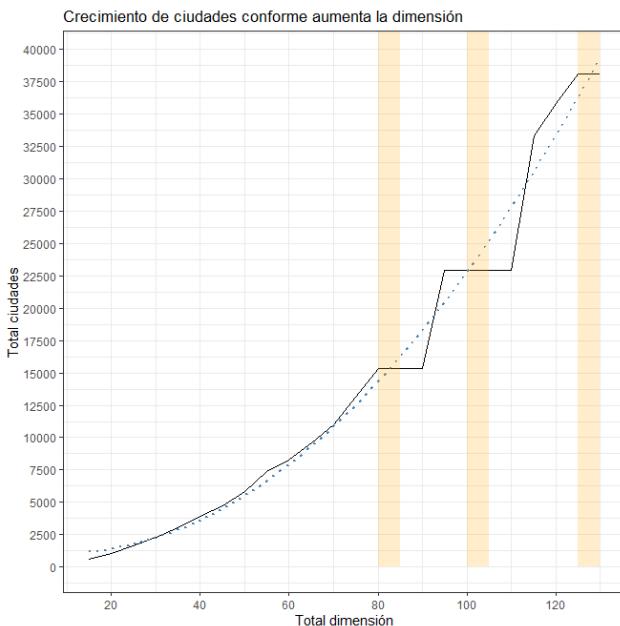
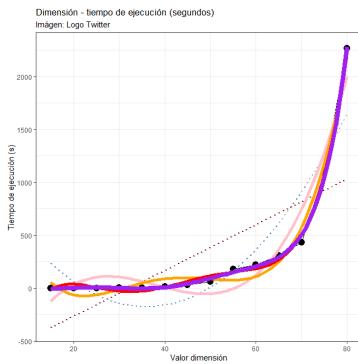


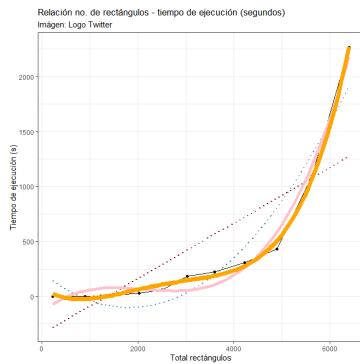
Figura 5.23. Ciudades generadas vs dimensión (con $m = n$).
Rangos donde se reduce la cantidad de ciudades se resalta en amarillo

suave, consistente e intuitivo.

Sin embargo existe un componente que vale la pena precisar si lo que se busca es comparar el tiempo de ejecución de las instancias de programación lineal genérico vs el del módulo que trabaja el TSP. Y es que el problema aquí versa sobre las ciudades distribuidas sobre los rectángulos (y sus correspondientes distancias mutuas) y no sobre el número de rectángulos en sí. Se puede decir que las ciudades crecen polinomialmente en función de la dimensión de rectángulos, tal como se observa en la gráfica 5.23. Interesa por lo tanto conocer el comportamiento del tiempo de ejecución conforme crecen las ciudades.



(a) Tiempo vs dimensión



(b) Tiempo vs rectángulos

Figura 5.24. Graficaciones tiempos de ejecución

La gráfica 5.25 demuestra cómo, similar al caso de la relación rectángulos vs tiempo, basta el polinomio de grado 4 para estimar una relación en ciudades vs tiempo. En síntesis, en notaciones *Big O*, hablamos de un algoritmo de orden $O(n^3)$ para fotomosaicos frente a uno $O(n^p)$ con $p \geq 4$ para el TSP; esto aún cuando el primero opera un problema de optimización exacta, mientras que el segundo opera una heurística, *2-opt* (¡y una de las más sencillas de entre las documentadas en este escrito!). Esta confronta basta para dimensionar la diferencia en complejidades de un enfoque TSP, mucho más grande versus su contraparte.

Una cuestión que vale la pena abordar es la relativa a si el algoritmo programado es capaz de reproducir imágenes de un grado más elevado de detallado visual, o se limita a ser aplicable únicamente a imágenes sin relieve, ni dimensión, ni variedad de colores. Para este efecto se buscará reproducir una foto personal del autor de este escrito en su escritorio de casa, la foto tendrá por nombre “*home office*”; se aprecia en 5.26.

Lo que sigue es reproducir la imagen. Para esto, se hizo uso de un

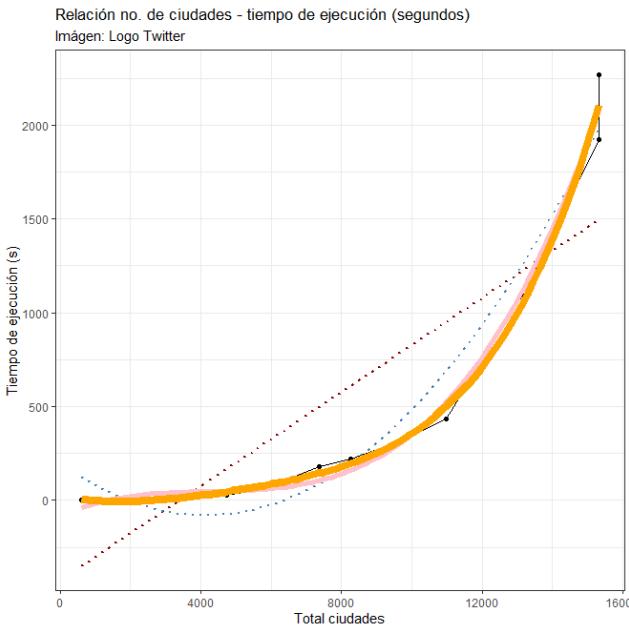


Figura 5.25. Relación Tiempo tsp vs ciudades

truco: atenuar los colores más claros. Es decir, a partir de un parámetro determinado de escala de luminosidad en adelante se aclarará la escala al máximo valor posible, es decir, 1. Para esta imagen el parámetro elegido es de 0.70. Esto permitió que la imagen resultara más distingible, sobre todo en lo que concierne a las facciones faciales: cara, ojos, boca y al artículo electrónico que se tiene sobre las orejas (audífonos). La razón de este recurso puede aclararse si se muestra el resultado de ejecutar el programa sin hacer esta modificación de la luminosidad. Ambos resultados se muestran en los dos pares de imágenes indexados como 5.27 y 5.28, respectivamente.

Ahora hagamos la instancia del Newton que se mencionó en los apartados del ejercicio de los fotomosaicos. Se reproduce procesando



Figura 5.26. Selfie en pandemia

en dimensiones 50×50 . El resultado genera una imagen que se distingue con claridad como una manzana tanto por la forma de la cáscara como por la hoja, acompañada de una figura. La parte del hombre se identifica según el observador, cuando el escritor preguntó, las respuestas fueron variadas. Algunos empezaron por identificar primero las piernas para luego dilucidar que el resto de ese segmento de imagen es de una figura humana; otros, en tanto, lograron identificar la forma completa, ya concibiendo el libro que sostiene. Esto, fue tras entrecerrar los ojos.

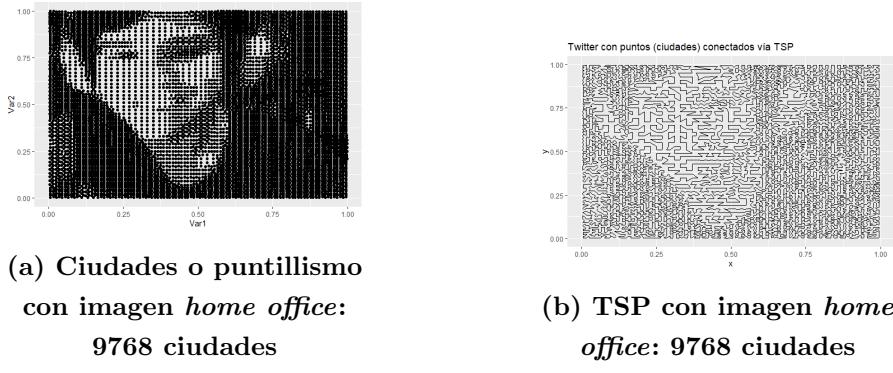


Figura 5.27. Puntillismo y TSP para autorretrato (sin agudizar brillos)

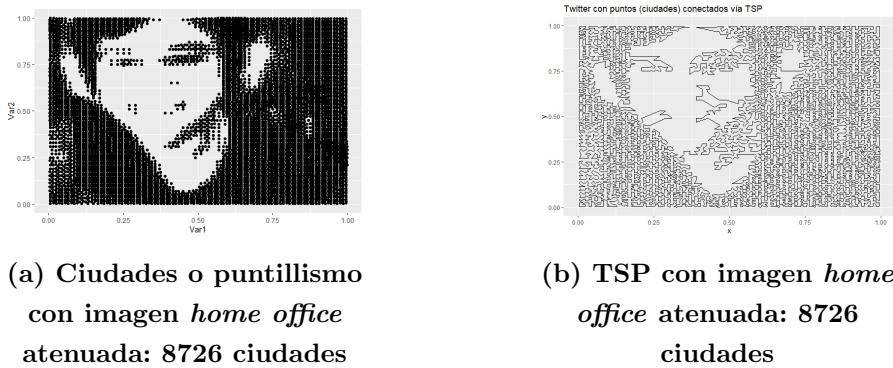


Figura 5.28. Puntillismo y TSP para autorretrato (agudizado brillos)

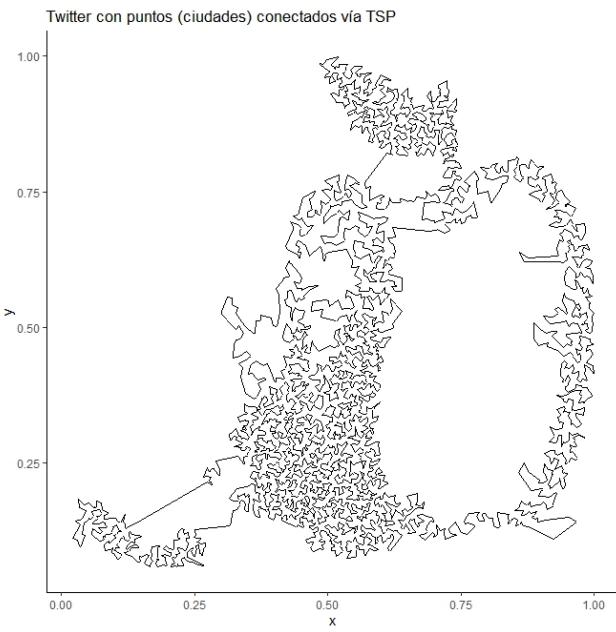


Figura 5.29. Ejercicio de Newton con el problema del agente viajero, versión 50×50

Capítulo 6

Conclusiones

En el presente trabajo se presentó una aplicación de la optimización en la generación de fotomosaicos. Primero, se revisaron modelos de programación lineal para figuras en escalas de grises, posteriormente con color y por último se describió cómo generar una plantilla para hacer un fotomosaico al resolver un problema del agente viajero que nos dé un tour hamiltoneano que se asemeje a la figura subyacente. Es interesante resaltar que esta aplicación de la optimización es en el ámbito del diseño o arte más que en el de decisiones para la gestión empresarial, que es la aplicación donde se ha centrado más esta disciplina.

Analizamos la manipulación de las escalas de grises, y del modelo aditivo de color de tipo RGB para poder darle un valor a los mosaicos que se van a usar y a los rectángulos a los cuales van a suplir; asimismo, se aplicó una variante de este modelo de escalas de grises por rectángulo para definir la dimensionalidad del conjunto de ciudades a visitar para el agente viajero. Se determinó que la interpretación en arreglos de pixeles de las imágenes dan pie a plantear una variedad de

problemas de programación lineal, cuyos resultados se pueden ver a simple vista. Esto permite plantear nuevos fotomosaicos con base en la creatividad del proponente (un ejemplo de estas variantes poco profundizado en este texto y propuesto por Bosch es el planteamiento con fichas de dominó, el cual incorpora un componente de orientación y deriva en buenos resultados visuales, como fue mostrado en la sección 3.2). Tanto los fotomosaicos como el TSP para replicar la imagen objetivo son formas creativas de interpretar el arreglo numérico de pixeles partiendo de las dimensiones de renglones y columnas que se entreguen exógenamente; asimismo, la segmentación en k-medias de una imagen resultó una buena aplicación del modelo RGB para sintetizar un color principal, o promedio. En particular la aplicación de fotomosaicos es interesante en el rubro de la mercadotecnia para demostrar cómo se altera la percepción del observador en función de la lejanía con que se contempla una imagen (a mayor lejanía, el observador logra dilucidar mejor la foto a replicar). El vínculo entre estas ramas radica en que la mercadotecnia se centra en la manipulación de la percepción. De manera complementaria, la concepción y entendimiento de los pixeles permite procesar diversos algoritmos con aplicaciones fundamentales en la era tecnológica actual, tales como la identificación de formas / siluetas, detección de colores, análisis de archivos de video, identificación de patrones de movimiento, etc.

Para dar un marco teórico de los dos tipos de problemas desarrollados en este trabajo fue necesario hacer una revisión general de la programación lineal, de la complejidad que implica pasar de un problema en un espacio continuo de soluciones a uno entero ocupando el algoritmo de ramificación y acotamiento; finalmente, se identificó una aplicación de un problema lineal entero donde, aún con la

ramificación y acotamiento resulta infactible resolver instancias en escalas muy grandes de variables, tal es el caso del problema TSP que requiere de heurísticas para su solución: algunas heurísticas conocidas de construcción y mejora de tours, que fueron mencionados y discutidos en la bibliografía de Nilsson.

Los programas de R para particionar las imágenes objetivo que se codificaron para este proyecto, *matsplitter(.)*, resultaron útiles y robustos a las dos clases de problemas abordados: fotomosaicos y TSP. Fungieron como punto de partida para obtener escala de gris promedio ayudando a hacer más eficiente y modular el código ya que todos los planteamientos compartían esta lógica.

Las instancias trabajadas de los fotomosaicos en *grayscale*s resultaron efectivas para replicar la imagen objetivo con relativa simplicidad, ya que previamente se generó un *grid* con imágenes promedio con una variedad que aproxima de buena manera al espectro completo de escalas de grises usando 20 mosaicos (en el rango de 0 a 1 con *grid* de 0.05); sin embargo, para la versión a color debió hacerse una selección manual de flores de colores que no abarcan satisfactoriamente el espacio tridimensional. La selección funcionó para el conjunto de imágenes objetivo que se establecieron, la bandera de México y la representación de Newton y su manzana, ya que se eligieron mosaicos manualmente, pensando pragmático en cumplir un objetivo, conciliando visualmente colores de flores versus colores en los segmentos de la imagen. En el espacio RGB 3- dimensional, significa un reto computacional relevante el crear un *grid* con respectivas imágenes asociadas que satisfaga la variedad exhaustiva que se realizaba con simplicidad en el ejercicio con escalas de grises. Esto, ya que el salto de una a tres dimensiones es cúbico, haciendo muy difícil pasar de hacer 10 a $10^3 = 1,000$ mosaicos, o de 20 a $20^3 = 8,000$. En

contraste con la librería de R especializada en fotomosaicos (*RSimMosaic* y su función, *composeMosaicFromImageRandomOptim*), los resultados del problema de fotomosaicos desarrollado aquí fueron visiblemente más estéticos, ya que la librería no trabaja un problema de minimización de distancias sobre cada rectángulo, sino de asignación de mosaicos mediante heurísticas de rápida —no necesariamente eficiente— solución. De manera que, en el planteamiento aquí desarrollado, la similaridad, establecida por un buen criterio matemático, se ve bien reflejada. Frente a esta observación, podemos establecer un posible alcance futuro para este proyecto: optimizar y documentar el código para poder construir la correspondiente librería para uso común en R, haciendo accesible este algoritmo para todo el público.

Para el enfoque del TSP se hizo una introducción teórica, pasando por restricciones factibles que mejoran el cómputo en términos de complejidad, así como algoritmos de construcción y mejora de tours, donde existen heurísticas de diseños muy variados y con distintos grados de complejidad y cercanía con el óptimo, explorando la cota de *Held-Karp* que aproxima los niveles óptimos. En lo referente a estas heurísticas del TSP, exhaustivamente estudiadas en el campo de la investigación de operaciones, la reflexión final conduce a balancear el *trade-off* entre economizar tiempo de cómputo por un *software* (hacer que éste sea factible) y obtener una buena cercanía de la solución al óptimo global.

Para definir las posiciones de las ciudades se inspeccionaron dos principales propuestas de disposición en el cuadrado $[0, 1] \times [0, 1]$: sobre cada rectángulo, una disposición regular basada en k-medias y otra basada en dispersión aleatoria. Estéticamente, las evidencias se inclinaron por el primer enfoque, pues los ejercicios con disposición

regular fueron más claros y asociables a la imagen subyacente.

Sobre el cómputo del TSP, la heurística de mejora de tours elegida para ejecutar las instancias definidas en este texto fue la *2-opt*. Respecto a los tiempos de ejecución, estos fueron relativamente razonables para las proporciones del problema, llegando al nivel de 2,000 segundos (media hora aproximadamente) para el procesamiento de 15,000 ciudades, como se ve en el logo de Twitter con las dimensiones de 85×85 rectángulos. Los despliegues con TSP tuvieron buenas aproximaciones a las imágenes objetivo, incluso variando sus grados de elaboración: pasando por la imagen del logo de Twitter, el Newton y el autorretrato. La complejidad de este modelo, no obstante, se demostró en las gráficas correspondientes (fig 5.23), con un tiempo de procesamiento que es notablemente más sensible ante cambios en dimensionalidad de rectángulos elegidos que lo que resultó ser en la solución de fotomosaicos (ambos enfoques mostraron contrastes de complejidades con cuando menos un grado polinomial de diferencia). También, si observamos la tabla de complejidad correspondiente al ejercicio de Twitter para ambos planteamientos (ver 5.1 y 5.3), la máxima dimensionalidad en común de rectángulos procesados fue de 60×60 , con el TSP implicó una resolución del problema de 221 segundos versus los 64 para fotomosaicos, esta referencia en tiempos puede dar noción de la significativa diferencia entre los tiempos de cómputo de ambos problemas. Esto tiene que ver, primero con la restricción de integralidad que sí se tiene en TSP y en los fotomosaicos no (ya que, como se menciona, la unimodularidad de este problema permite que no sea necesario agregar integralidad sobre las variables de decisión para que ésta se cumpla en el óptimo global); y segundo, con la creciente complejidad inherente al problema de TSP comentada en este escrito, lo que lo hace un problema NP-Duro. Esta complejidad

es principalmente atribuible a las restricciones de no-subtours. Aún con heurísticas o relajación de restricciones para resolverse de manera computacionalmente viable, no se logró ejecutar el ejercicio de manera más económica a lo expuesto en las tablas de complejidad.

Si bien el procesamiento de imágenes como arreglos matriciales trae consigo una amplia de variedad de aplicaciones en múltiples ramas como la ingeniería, la medicina, la industria o el transporte, esta ciencia computacional también puede aplicarse en la producción de arreglos visualmente estéticos. Como indica Robert Silvers en su obra “*Photomosaics*”, los resultados de esta clase de problemas de optimización con un enfoque visual nos da una linda conciliación o “matrimonio entre arte y tecnología; de fotografía y las computadoras, de la belleza y la ciencia”.

Bibliografía

- [1] Silvers, R. (1997). Photomosaics. New York: Henry Holt and Company.
- [2] Cole, S. (2014). Image: NASA releases Earth Day “global selfie” mosaic of our home planet. Isla de Man: PhysOrg.
- [3] Possani, E. (2012). Puentes, agentes viajeros y mosaicos: modelando usando gráficas y programación lineal. Ciudad de México: ITAM.
- [4] Bosch, R. (2004). Constructing Domino Portraits. Ohio: Oberlin College.
- [5] Bosch, R. (2003). Continuous line drawings via the traveling salesman problem. Ohio: Oberlin College.
- [6] Bosch, R. (2006). Opt Art. Ohio: Oberlin College.
- [7] Castañeda, W. (2005). Color. Caldas: Universidad de Caldas.
- [8] Aqil Burney, SM. (2014). K-Means Cluster Analysis for Image Segmentation. Sind: University of Karachi.
- [9] Goic, M. (2009). Análisis Post Optimal y Algoritmo de Ramificación y Acotamiento. Santiago: Universidad de Chile.

- [10] Nilsson, C. (2003). Heuristics for the Traveling Salesman Problem. Suecia: Linkoping University.
- [11] Chong, B. (2021). K-means clustering algorithm: a brief review. Shanxi: Academic Journal of Computing & Information Science, Vol. 4.
- [12] González-Santander, G. (2020). Tres métodos diferentes para resolver el problema del viajante. Baobab Soluciones. [https://baobabsoluciones.es/blog/2020/10/01/problema-del-viajante/#:~:text=El%20problema%20del%20viajante%20\(por,Su%20origen%20no%20est%C3%A1%20claro.](https://baobabsoluciones.es/blog/2020/10/01/problema-del-viajante/#:~:text=El%20problema%20del%20viajante%20(por,Su%20origen%20no%20est%C3%A1%20claro.)
- [13] Fernandez, V; Zelaia, A. (2011). Investigación operativa. Programación lineal. Vizcaya: Universidad del País Vasco.
- [14] Fraga, A. (2024). Uncalibrated: Drawing, Photography and the Raw Images of Astronomy. Yorkshire del Oeste: The University of Leeds.
- [15] Desrochers, M; Laporte, G. (1991). Improvements and extensions to the Miller-Tucker-Zemlin subtour elimination constraints. Quebec: Operations Research Letters, Volume 10.

Apéndice A

Apéndice

A.1. Adecuando largo y ancho de píxeles para adaptarse a parámetros ingresados

Para la regla de asignación consistente en hacer la μ' múltiplo de m y n múltiplo de n se requiere que la cantidad de renglones de la nueva matriz acotada sea 0 módulo m y, la cantidad de columnas, 0 módulo n . Si bien puede haber más de una combinación renglón \times columna consistente con las propiedades mencionadas, para nuestra regla de recortar la matriz mayor se escogerá una nueva dimensión $\mu' \times \nu'$ que corresponda al máximo valor factible que no exceda $\mu \times \nu$.

$$\begin{aligned} DIM = & \{(\mu', \nu') = (\max(a), \max(b)) | 0 = a \text{mod}(m), \\ & 0 = b \text{mod}(n), a \in [0, \mu] \cap \mathbb{N}, b \in [0, \nu] \cap \mathbb{N}\} \end{aligned} \quad (\text{A.1})$$

Es trivial notar que se requiere que $m \leq \mu$ y $n \leq \nu$. El acotamiento arriba expresado se realiza de la manera más simétrica posible. Esto es, si $\mu' < \mu$, se recortarán $\epsilon = \lfloor \frac{1}{2}(\mu - \mu') \rfloor$ renglones por arriba y

$\delta = (\mu - \mu') - \epsilon$ renglones por abajo . Así, se tendrá $\epsilon - \delta \in \{0, 1\}$ (0 si $\frac{1}{2}(\mu - \mu')$ es par, 1 e.o.c), lo cual quiere decir que, en el peor caso, se recortará un pixel más por un lado que por otro. Análogamente, si $\nu' < \nu$, se recortarán $\rho = \lfloor \frac{1}{2}(\nu - \nu') \rfloor$ columnas por la izquierda y $\sigma = (\nu - \nu') - \rho$ por la derecha (siendo también la diferencia de, a lo mucho, uno).

A.2. Reduciendo dimensionalidad del vector de costos

La forma genérica que ayuda a convertir el arreglo 3-dimensional de costos en los fotomosaicos en uno de 2 se muestra a continuación

Cuadro A.1. Tabla de costos para combinaciones de f, i, j

t	f	i	j	costo (c')
1	1	1	1	A[1,1,1]
2	1	1	2	A[1,1,2]
\vdots	\vdots	\vdots	\vdots	\vdots
n	1	1	n	A[1,1,n]
$n + 1$	1	2	1	A[1,2,1]
$n + 2$	1	2	2	A[1,2,2]
\vdots	\vdots	\vdots	\vdots	\vdots
$2n$	1	2	n	A[1,2,n]
\vdots	\vdots	\vdots	\vdots	\vdots
$(m - 1)n + 1$	1	m	1	A[1,m,1]
$(m - 1)n + 2$	1	m	2	A[1,m,2]
\vdots	\vdots	\vdots	\vdots	\vdots
mn	1	m	n	A[1,m,n]
$mn + 1$	2	1	1	A[2,1,1]
$mn + 2$	2	1	2	A[2,1,2]
\vdots	\vdots	\vdots	\vdots	\vdots
$mn + n$	2	1	n	A[2,1,n]
$mn + n + 1$	2	2	1	A[2,2,1]

Continúa en la siguiente página

t	f	i	j	costo (c')
$mn + n + 2$	2	2	2	A[2,2,2]
\vdots	\vdots	\vdots	\vdots	\vdots
$mn + 2n$	2	2	n	A[2,2,n]
\vdots	\vdots	\vdots	\vdots	\vdots
$mn + (m - 1)n + 1$	2	m	1	A[2,m,1]
$mn + (m - 1)n + 2$	2	m	2	A[2,m,2]
\vdots	\vdots	\vdots	\vdots	\vdots
$2mn$	2	m	n	A[2,m,n]
\vdots	\vdots	\vdots	\vdots	\vdots
$(cf - 1)mn + 1$	cf	1	1	A[cf,1,1]
$(cf - 1)mn + 2$	cf	1	2	A[cf,1,2]
\vdots	\vdots	\vdots	\vdots	\vdots
$(cf - 1)mn + n$	cf	1	n	A[cf,1,n]
$(cf - 1)mn + n + 1$	cf	2	1	A[cf,2,1]
$(cf - 1)mn + n + 2$	cf	2	2	A[cf,2,2]
\vdots	\vdots	\vdots	\vdots	\vdots
$(cf - 1)mn + 2n$	cf	2	n	A[cf,2,n]
\vdots	\vdots	\vdots	\vdots	\vdots
$(cf - 1)mn + (m - 1)n + 1$	cf	m	1	A[cf,m,1]
$(cf - 1)mn + (m - 1)n + 2$	cf	m	2	A[cf,m,2]
\vdots	\vdots	\vdots	\vdots	\vdots
$m \cdot n \cdot cf$	cf	m	n	A[cf,m,n]

Como se muestra, el ordenamiento de $df3$ sigue la jerarquía f, i, j .

La estructura descrita redefine el problema, para este caso, pasando de uno sobre variables de decisión indexadas sobre tres componentes x_{fij} a su análogo 1-dimensional, x_t . El ejercicio requiere de una reasignación $(f, i, j) \mapsto t(f, i, j)$ definida por $t = (f - 1)(m \cdot n) + m(i - 1) + j$ que representa el orden de los renglones que sigue $df3$. Ahora, defínanse en R los parámetros requeridos en el párrafo anterior

1. $ce = df3["costo"]$

2. $io = "min"$
3. Restricciones $RM \in \mathbb{R}^{re \times m n cf}$ con $re = cf + m n$, los primeros cf renglones representando a las restricciones de disponibilidad de los mosaicos que hay, y las restricciones restantes representando el correcto llenado de cada celda genérica (i, j) . La matriz consta únicamente de valores 0 y 1. Los 1's se asignan fácilmente conociendo, con la ayuda de la asignación $t(.)$ qué posiciones corresponden a cada f para el primer conjunto de restricciones; similarmente, para el segundo conjunto de restricciones basta reconocer qué posiciones le corresponden a cada celda (i, j) .
4. $rh_f = u_f \quad \forall f \in \{1, \dots, cf\}$

$$rh_t = 1 \quad \forall f \in \{cf + 1, \dots, cf + m n\}$$
5. $sign_f = "\leq" \quad \forall f \in \{1, \dots, cf\}$

$$sign_t = "=" \quad \forall f \in \{cf + 1, \dots, cf + m n\}$$

Foto-mosaicos
y optimización para el arte (Opt-Art)
escrito por Julio César Espinosa León,
se terminó de imprimir en marzo de 2025
en los talleres de Tesis Martínez.
República de Cuba 99, colonia Centro,
Ciudad de México.