

Retoque fotográfico mediante reconstrucciones geométricas

Ariel Nowik, Joaquin Mestanza, Rocio Parra, Martina Máspero, Marcelo Regueira
22.05 - Análisis de Señales y Sistemas Digitales - Grupo 1

ITBA: Instituto tecnológico de Buenos Aires
Ciudad de Buenos Aires, Argentina

Resumen—En este trabajo se estudiaron diversos métodos de retoque de imágenes para eliminar elementos no deseados presentes en diversas fuentes. Finalmente se procedió a realizar una implementación en función de las técnicas analizadas seguida de un análisis de sus ventajas y desventajas.

I. INTRODUCCIÓN

El problema elemental a resolver consiste en la eliminación de un objeto no deseado en una imagen. Naturalmente no es posible “adivinar” lo que se encuentre por detrás, ya que requiere información adicional, la cual en principio no es accesible, solo se dispone de la imagen. Por lo tanto la idea es, de algún modo asimilar la zona de la imagen a reemplazar con el resto de la misma. En lo que continua de este trabajo describiremos con un mayor detalle diversos métodos para llevar a cabo este proceso.

II. LÍNEAS DE INVESTIGACIÓN

En primera instancia, se realizó una investigación acerca del estado del arte de este problema que se busca resolver. Este tema es extensivamente explicado en *Image inpainting: Overview and recent advances* por Christine Guillemot y Olivier Le Meur [3].

II-A. Definición del problema de inpainting

En primer lugar se procede a definir matemáticamente el problema que en este trabajo se aspira a resolver.

Una imagen I puede definirse de la siguiente manera

$$I(\vec{x}) : R^n \rightarrow R^m \quad (1)$$

Donde I es una función que mapea un pixel $\vec{x} = (x, y)$ a un color R^m . Como la imagen es bidimensional, $n = 2$, y utilizando un esquema RGB se tiene $m = 3$.

En el problema de inpainting se entiende que la imagen I fue degradada por un operador M , lo cual genera una imagen nueva F que tiene la imagen con algunos píxeles cambiados (que pertenecen a una región U conocida). Se suele escribir entonces:

$$F = M(I) \quad (2)$$

Lo que se busca en principio es reconstruir la imagen original I a partir de F conociendo M (qué píxeles fueron removidos y se quieren reconstruir). Este problema en el sentido estricto no tiene solución, se suele decir entonces

que no es un problema bien definido con una única solución. Entonces el objetivo es lograr una aproximación de I (llamémosla I') que logre ser lo más parecida posible a I .

Otra consecuencia de este mal condicionamiento del problema es cómo definir el éxito a partir de I' y de F . Dado que se desconoce la imagen original I , que sería el resultado ideal, no existe una métrica cuantitativa que determine la calidad de una reconstrucción dada, y por lo tanto se debe recurrir a evaluaciones subjetivas. En general, lo que se espera como resultado es una imagen que parezca físicamente plausible, y que parezca natural al ojo humano.

II-B. Primeras ideas

Existen una gran variedad de técnicas para afrontar el problema. Dentro de todas ellas se distinguen dos subgrupos principales, que son los métodos basados en difusión, y los métodos basados en ejemplares. Como no existe ninguna ventaja determinante de una variante y la otra en este trabajo práctico se procedió a investigar ambas familias de métodos. Acorde al tipo de imagen puede ser más conveniente un método u otro, incluso en algunos casos se puede usar una mezcla de ambos.

II-B1. Métodos basados en difusión: La primera categoría de métodos se denomina difusión, y la idea fundamental proviene de propagar información con restricciones usando ecuaciones diferenciales parciales. La idea consiste en mediante una ecuación propagar sobre la región U con píxeles cambiados (del exterior al interior) información que permita reconstruirla de una manera acorde. En este método juega un papel fundamental qué ecuación utilizar para realizar la propagación, en particular sobre qué direcciones propagar más, y sobre cuáles menos. En general este método suele dar resultados de mayor efectividad cuando la imagen tiende a tener patrones homogéneos sin cambios de textura muy abruptos. Más adelante se describirá una variante del método con mayor detalle.

II-B2. Métodos basados en ejemplares: Una segunda categoría de métodos para resolver el problema de inpainting se basa en asumir que la imagen I posee ciertas probabilidades de estacionariedad y auto similaridad entre sus distintas regiones. Partiendo de este supuesto, se procede a rellenar la parte de la imagen que falta a partir del uso de ejemplares, o parches, de la región que se tiene.

Estos métodos, entonces, tienen a grandes rasgos dos criterios a desarrollar:

- cómo elegir por qué sección del contorno se comienza a rellenar, o qué región debe rellenarse a continuación
- cómo elegir el parche con el cual se rellenará dicha región

II-B3. Métodos híbridos: De los dos métodos vistos hasta ahora, basados en difusión o en ejemplares, ninguno es claramente superior al otro para todos los casos, sino que en ciertas circunstancias puede funcionar mejor uno o el otro.

En líneas generales, con los métodos basados en difusión se obtienen buenos resultados para huecos pequeños y esparsamente distribuidos. Lo mismo aplica para estructuras con bordes definidos. Por otro lado, los métodos basados en ejemplares funcionan bien para regiones con texturas o patrones regulares.

Como es razonable esperar que una imagen tenga tanto estructuras como texturas, surge una familia de algoritmos que busca aprovechar las ventajas de ambos métodos según corresponda. Se los conoce, por lo tanto, como métodos híbridos.

A grandes rasgos, existen dos formas de implementar esta idea:

- identificar qué partes de la imagen corresponden a estructuras y qué partes a texturas, rellenar por difusión o por ejemplares según corresponda, y combinar los resultados obtenidos
- combinar ambos métodos mediante el uso de una única función de energía (más allá del alcance de este trabajo)

III. DESARROLLO DEL MÉTODO DE DIFUSIÓN

En esta sección se procede a trabajar en profundidad el método de difusión iniciando por una introducción teórica y luego profundizando en una versión del método en particular, llamada difusión anisotrópica.

III-A. Filtros basados en EDPs

En la década del 70, un inconveniente que se tuvo que enfrentar en el procesamiento de imágenes fue que al aplicar transformaciones sobre las imágenes, la transformación deterioraba los bordes de la misma. Por ejemplo, si tenemos un gato y un perro en una imagen, queremos hacer difusión sobre las partes dentro de los mismos y dejar los bordes que los separan intactos. Para poder solventar este inconveniente se propusieron distintos métodos que utilizaban ecuaciones diferenciales parciales, las cuales permiten diferir qué zonas se va a procesar. El primer modelo presentado que apuntaba a esta solución fue hecho por Witkin en 1983, el cual consistía en convolucionar la imagen con un filtro gaussiano.

$$I(x, y, t) = I(x, y, 0) * G(x, y; \sigma_t)$$

Donde $I(x, y, 0)$ es la imagen original.

En 1984, Koenderink observó que las imágenes obtenidas podrían ser vistas como la solución a la ecuación de la difusión del calor:

$$\frac{\partial I(x, y, t)}{\partial t} = \Delta I(x, y, t) = \frac{\partial^2 I(x, y, t)}{\partial^2 x} + \frac{\partial^2 I(x, y, t)}{\partial^2 y}$$

Donde nuevamente la condición inicial es $I(x, y, 0)$, la imagen original.

III-B. ¿Por qué la ecuación del calor?

Los procesos de difusión están altamente relacionados con las leyes de conservación, es decir en sistemas donde hay una cierta concentración o temperatura I , sucede que hay un flujo de las regiones de mayor concentración a uno de menor concentración hasta que el sistema alcanza el equilibrio.

III-C. ¿Qué problemas conlleva utilizar el Laplaciano?

Como el operador laplaciano es un operador isótropo, la ecuación del calor no tiene direcciones preferentes, lo cual hace que sea igualmente aplicada en todas las direcciones del plano. Este tipo de flujo es llamado *difusión isotrópica*.

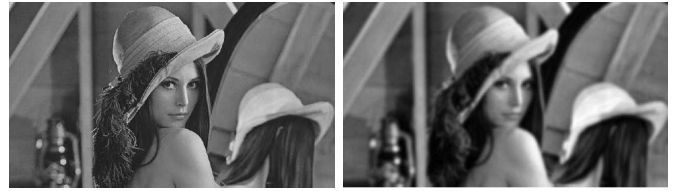


Figura 1. En la izquierda se encuentra la foto original y en la derecha se puede ver el resultado de haber aplicado difusión isotrópica

En la figura 1 se puede ver que la imagen se encuentra totalmente destruida.

III-D. Difusión anisotrópica

Hummel en 1986 establece que para obtener un buen resultado en el realce de la imagen, el flujo del filtro con el que se implemente satisfaga el principio del máximo. Desde ese entonces, aparecieron propuestas tales como PDE no lineales en las que no difunda el flujo uniformemente a pesar de no cumplir con el principio del máximo. Estos filtros son llamados de difusión anisotrópica. Pietro Perona y Jitendra Malik fueron quienes en 1990 presentaron un suavizado multiescala preservando bordes basados en 3 criterios:

- Causalidad: Disminuir la resolución de la imagen no puede dar nuevos detalles.
- Localización inmediata: los límites de las regiones deben ser distinguibles y coincidir con los de la imagen original
- Suavizado según zona: suavizado debe ser mayor dentro de una región que en sus fronteras

Este modelo se basa en la segunda ley de Fick, con un coeficiente de conductividad variable $c(x, y, t)$ que controla la difusión en Ω

$$\begin{cases} I_t = \text{div}(c(x, y, t) \cdot \nabla I) \\ I(x, y, 0) = I_0(x, y) \end{cases}$$

Donde

$$\text{div}(c(x, y, t)\nabla I) = \nabla \cdot (c(x, y, t)\nabla I) = c\Delta I + \nabla c \cdot \nabla I$$

Lo deseable según Perona y Malik es que el coeficiente de difusión c dependa de la norma del gradiente en cada punto. De esta forma, los gradientes con normas pequeñas (zonas homogéneas) tengan coeficiente alto (filtrado fuerte, por ende suavizado). Mientras que para una norma grande, el coeficiente sea pequeño con tal de que no se deterioren los bordes, es decir que el proceso de difusión sea lento.

Assumiendo $s = |\nabla I|$, dos propuestas que surgieron para c fueron:

Función de Lorentz:

$$c(s^2) = \frac{1}{1 + \frac{s^2}{K^2}}$$

Función de Leclerc:

$$c(s^2) = e^{-\frac{s^2}{2K^2}}$$

Donde K es una constante que se elige experimentalmente para controlar la sensibilidad ante los bordes.

El filtro que se aplica es isotrópico localmente pero la diferencia es que al adaptarse de distinta forma a regiones homogéneas y no homogéneas, adopta el nombre de anisotrópico (aunque técnicamente es difusión no lineal).

Con todo esto en mente, se procedió a realizar una serie de pruebas con un algoritmo que cumpla con estas especificaciones.



Figura 2. En la izquierda se encuentra la foto original y en la derecha se puede ver el resultado de haber aplicado difusión anisotrópica tras 40 iteraciones

Se observa en la figura 2 que a medida que avanza el algoritmo ciertos bordes prevalecen. Sin embargo, uno de los objetivos finales del proyecto es poder aplicar difusión en una región desconocida por lo que se optó por seguir investigando sobre alternativas que realicen inpainting usando difusión con EDPs.

El problema a abordar con inpainting fue desarrollado por Bertalmio et al. en el año 2000, cuyo trabajo está basado en el uso de un modelo anisotrópico que propaga los Laplacianos de la imagen de la vecindad próxima hacia adentro de la región. La dirección de la propagación está dada por la dirección de los isofotos estimados mediante la componente perpendicular del gradiente en cada punto.

La ecuación a resolver numéricamente es:

$$\frac{\partial I}{\partial t} = \Delta(\nabla I)\Delta I^\perp$$

Donde I es la intensidad de la imagen dentro de la región la cual luego de varias iteraciones llega a un estacio estacionario ($\Delta(\nabla I)\Delta I^\perp = 0$), lo que implica que los Laplacianos permanecen constantes (no hay variaciones según el gradiente ∇) en las direcciones de ∇I^\perp del isofoto. Esto se puede interpretar como que la información de la imagen es propagada hacia la región de forma tal de preservar la dirección de los isofotos. Los métodos de difusión tienden a prolongar estructuras (por ejemplo, con los isofotos) que llegan a la región a rellenar. Es por esta razón que estos métodos sirven para situaciones donde se quiere propagar fuertes estructuras o rellenar huecos pequeños. Sin embargo, esta técnica no sirve para imágenes texturizadas, especialmente si la región a rellenar es grande. A pesar de que se intenta preservar los bordes, después de unas iteraciones de difusión, la región aparece suave con un efecto de blur cuando la región a restaurar es grande.



Figura 3. En la izquierda se encuentra la foto original y en la derecha se puede ver el resultado de haber aplicado el algoritmo desarrollado por Bertalmio tras 200 iteraciones, debajo de las mismas está la imagen con zoom en la zona donde se aplicó difusión

Se puede observar que el algoritmo utilizado no es óptimo, pues, en la práctica se encontró que había que regular distintas constantes que afectaban la convergencia del algoritmo, tales como Δt que regula la velocidad de propagación, a su vez se utilizó la función de Leclerc (mencionada en este trabajo) para regular la actualización de I_t dado que grandes gradientes afectaban en gran medida a la convergencia del algoritmo (controlado por la constante K). Se deduce que, si bien el algoritmo implementado a grandes rasgos cumple lo deseado (la propagación de los isofotos), la manipulación de los meta-parámetros no fue la adecuada para esta imagen en específico.

IV. DESARROLLO DEL MÉTODO BASADO EN EJEMPLARES

En esta sección se procede a trabajar en profundidad el método basado en ejemplares, y luego se profundiza en una mejora en particular de una parte del método, llamada patch match.

IV-A. Descripción global del método

La ejecución del método de ejemplares (en todas sus variantes) consiste en una secuencia de pasos a ejecutar de manera

iterativa para reconstruir la zona U desconocida de la imagen. Estos pasos son tres

- Cálculo del borde
- Selección del punto del borde a eliminar
- Selección de rectángulo correspondiente para el reemplazo

En general las distintas líneas de investigación buscan optimizar cada una de estas etapas para lograr por un lado una mayor efectividad, y por el otro una mayor eficiencia.

IV-A1. Cálculo del borde: Es indispensable para el funcionamiento del algoritmo calcular el borde de la región U a eliminar, ya que el algoritmo siempre copia un ejemplar del exterior al borde del interior de la región U . Hay diversos algoritmos que calculan bordes, forman parte del área de la geometría computacional, el Convex Hull (wikipedia: https://en.wikipedia.org/wiki/Convex_hull) por ejemplo es un algoritmo que aplica para resolver un problema similar. Generalmente al ser un problema conocido existen infinitas implementaciones ya realizadas de código libre para resolver este problema.

IV-A2. Cálculo del gradiente de la imagen: Para el funcionamiento del algoritmo fue necesario el cálculo de los gradientes de la imagen. Para ello se calcula primero la imagen en una escala de grises y luego se aplica el operador Sobel, el cual combina tanto derivada como suavización gaussiana.

IV-A3. Selección del punto del borde a eliminar: En cada iteración se necesita calcular de todos los píxeles del borde aquellos con una mayor prioridad para ser asimilados. En los papers citados se muestran fundamentos por los cuales es más apropiado poderar por un lado el gradiente de la imagen y por el otro la confianza. La fórmula que determina la prioridad está dada por

$$P_i = C_i \nabla F_i \perp \vec{n}_i \quad (3)$$

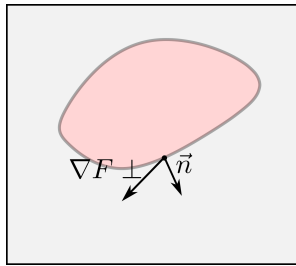


Figura 4. Gradiente y normal

Donde $\nabla F_i \perp$ es la perpendicular al gradiente de la imagen en el píxel i . Es el punto donde ocurre el cambio más suave en la intensidad del color de la imagen. El algoritmo realiza el producto escalar entre $\nabla F_i \perp \vec{n}_i$ lo cual consigue la componente normal de la perpendicular al gradiente, que nos indica que tan suave cambia la intensidad de la imagen en la dirección normal al borde de la región

IV-A4. Determinación del rectángulo a reemplazar: Se necesita calcular cuál rectángulo utilizar para reemplazar la zona del borde de la imagen. Para ello se debe desarrollar un criterio que determine cuál es el más apto. El método más sencillo consiste en realizar la norma de la diferencia comparando la región a reemplazar y el candidato. Se utiliza la siguiente ecuación.

$$\sum (R_{1i} - R_{2i})^2 + (G_{1i} - G_{2i})^2 + (B_{1i} - B_{2i})^2 \quad (4)$$

Se utilizó una distribución normal para elegir al azar los candidatos

IV-A5. Optimización mediante Patch Match: Un enfoque por fuerza bruta en la selección del rectángulo correspondiente puede resultar excesivamente costoso en tiempo, ya que implicaría potencialmente comparar una región de la imagen con todas las demás. Por lo tanto, existen diversos criterios que pueden tomarse en este sentido. Uno de ellos es desarrollado por Barnes, Shechtman, Finkelstein y Goldman, quienes desarrollaron un algoritmo con el nombre de PatchMatch [4]. El mismo hace uso de la asunción de ciertas propiedades de coherencia en la imagen para reducir el número de comparaciones que debe hacerse, sin perder la efectividad del algoritmo. Una vez elegido un lugar donde se quiere rellenar, PatchMatch comienza comparando con algunas regiones aleatorias de la imagen original, y a partir del resultado más satisfactorio que obtiene hace dos suposiciones:

- si se tiene un buen match para una región que se quiere rellenar, es probable que, si existe un parche que es un match aún mejor, se encuentre cerca del bueno (es decir, se asume que estamos en la zona correcta de la imagen, y sólo debemos terminar de acercarnos al match óptimo)
- si un parche es un buen match para una cierta región a rellenar, es probable que también sea un buen match para regiones cercanas

De esta manera, según los autores del algoritmo, se logra obtener muy buenos resultados con menos comparaciones. La limitación que tiene este algoritmo es que no puede dejarse de introducir comparaciones random incluso habiendo encontrado un “buen” match, puesto que se corre el riesgo de quedarse en un mínimo local, que no sea el mínimo global.

IV-A6. Pruebas del Método: Se procedió a ejecutar el método con algunos casos simples

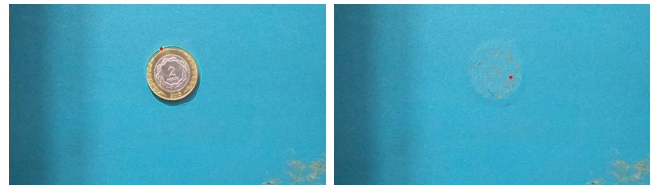


Figura 5. Resultados - ejemplo 1

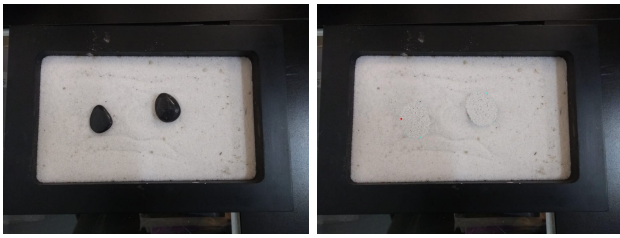


Figura 6. Resultados - ejemplo 2

Estas son algunas pruebas iniciales, las cuales mostraron algunos desperfectos en el metodo original, como por ejemplo que la region reemplazada se vea "sucia" tambien dieron una noci3n de la tardanza del algoritmo.

V. OBJETIVOS FUTUROS

Se identifican dos direcciones en las que se podr3a avanzar para mejorar el algoritmo propuesto.

En primer lugar, podr3a reducirse el tiempo de b3squeda de los *nearest neighbors*. Se consider3 utilizar el approach aplicado en Patch Match, pero el algoritmo descrito en este paper requiere que el orden de rellenado sea distinto al aplicado en este trabajo. Por lo tanto, se deber3an tomar las observaciones de las que hace uso Patch Match pero aplicarlas de forma tal que sea compatible con el orden de rellenado propuesto por Criminisi.

Por otro lado, siendo que se obtuvieron resultados parcialmente satisfactorios tanto utilizando el enfoque de Criminisi, basado en ejemplares, como con PDEs, basado en difusi3n, una posible l3nea para continuar ser3a desarrollar un m3todo h3brido que combine ambos algoritmos, separando las regiones con texturas y estructuras y aplicando Criminisi y PDEs respectivamente.

REFERENCIAS

- [1] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE T. Image Process.*, vol. 13, no. 9, pp. 1200–1212, Sep. 2004.
- [2] Pierre Buysens, Maxime Daisy, David Tschumperl3, Olivier L3zoray. Exemplar-based Inpainting: Technical Review and new Heuristics for better Geometric Reconstructions. *IEEE Transactions on Image Processing*, Institute of Electrical and Electronics Engineers, 2015, 24 (6), pp.1809 - 1824. [ff10.1109/TIP.2015.2411437](https://doi.org/10.1109/TIP.2015.2411437)ff. [ffhal-01147620f](https://arxiv.org/abs/1407.4011)
- [3] Guillemot, Christine & Le Meur, Olivier. (2014). Image Inpainting : Overview and Recent Advances. *Signal Processing Magazine, IEEE*. 31. 127-144. [10.1109/MSP.2013.2273004](https://doi.org/10.1109/MSP.2013.2273004).
- [4] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman, "Patch-match: randomized correspondence algorithm for structural image editing," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 24:1–24:11, July 2009.