

La transformada Mojette

Arturo Moreno, Gerardo Miramontes, Ernesto García, Claudia Sifuentes, Rafael Magallanes, Enrique de la Rosa,

The Mojette transform

Recibido: agosto 12, 2011

Aceptado: octubre 23, 2011

Palabras clave: Proyecciones; transformadas; Mojette

Abstract:

We present an introduction to a not very well known transform. Although this transform is not well known, there are other transforms and operations that are quite similar to it and for that reason it may be confused with some other method. The Mojette transform is a projection method very similar to the Radon transform, but with some important differences.

Keywords: Projections; transforms; Mojette

EN 1917 un matemático de nombre Johann Radon (1887-1956) publicó un artículo en el que aseguró que se podía reconstruir una función desconocida, si se conocían todas las integrales de línea de esa función en el plano. La operación con la cual se puede hacer esa reconstrucción se llama transformada de Radon. Se aplica principalmente en tomografía computarizada, es decir en la obtención de imágenes transversales.

Como lo que se obtiene es una reconstrucción, también se le llama reconstrucción tomográfica. Esta reconstrucción se obtiene a partir de las proyecciones que se recogen en

un detector cuando se hace incidir una fuente de energía (un haz) sobre un cuerpo u objeto, como se explica más adelante.

Para explicar a grandes rasgos qué es la reconstrucción tomográfica, partimos de la Figura 1 [1], donde se muestra que las proyecciones de un objeto a determinado ángulo θ están conformadas por una serie de integrales de línea. Por simplicidad puede entenderse que la integral es una suma, cuyo resultado representa la atenuación del haz al pasar a través del objeto. La imagen desconocida es $f(x,y)$, y la información recolectada es $p(r)$, donde r es la posición de rayos paralelos, desde una proyección a un ángulo θ .

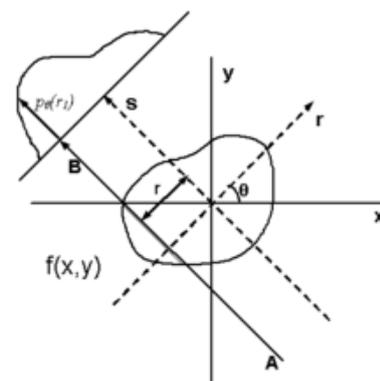


Figura 1. Geometría de haz paralelo. Cada proyección esta formada por el conjunto de integrales de línea a través del objeto.

Suponiendo que el objeto produce una atenuación exponencial, lo cual es típico en tejidos, entonces

$$I = I_0 e^{-\int f(x,y) ds} \quad (1)$$

donde I es la intensidad detectada, I_0 es la intensidad de la fuente, $f(x)$ es el coeficiente de atenuación en la posición x . En general la atenuación total p de un rayo en posición r a un ángulo de proyección θ está dada por

$$p(r, \theta) = \ln(I/I_0) = - \int f(x, y) ds \quad (2)$$

de acuerdo a la Figura 1 el punto $p(x, y)$ será proyectado en el ángulo θ por

$$x \cos(\theta) + y \sin(\theta) = r. \quad (3)$$

Sustituyendo podemos escribir[1]

$$p(r, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - r) dx dy \quad (4)$$

donde $p(r, \theta)$ es llamada la transformada de Radon (TR).

Para obtener $f(x, y)$ se requiere una transformación inversa. Sin embargo la transformada inversa de Radon es inestable cuando se tiene ruido en los datos.

Cuando se tienen datos discretos, por ejemplo en una imagen digital, entonces se trabaja con la versión discreta de la transformada, donde las integrales se aproximan por sumas.

Otros métodos de proyección

De la sección anterior, puede verse que la proyección obtenida es una representación del objeto o de la imagen bajo estudio. Esta representación se puede emplear en otros campos, como por ejemplo en sistemas automáticos de reconocimiento de patrones, o para almacenamiento de datos.

Cuando se tienen datos discretos, por ejemplo en una imagen digital, se requiere la versión discreta de la transformada de Radon

Recientemente, se han sugerido métodos basados en la transformada de Radon para extraer una representación de una imagen. Por ejemplo, [2] muestra cómo la transformada de Radon representa eficientemente y describe símbolos compuestos de segmentos de línea.

Transformada Mojette

Entre las transformaciones lineales propuestas y usadas para compresión de imágenes, en años recientes, se encuentra la transformada Mojette (TM), que es una especie de transformada de Radon discreta [3, 4, 5, 6].

La transformada Mojette [7, 8, 9, 10, 11] proyecta una imagen digital original en 2D $I = \{I(i, j); i = 1, \dots, N; j = 1, \dots, M\}$ en un conjunto de K proyecciones discretas en 1D con $P = \{M_k(l); k = 1, \dots, K; l = 1, \dots, L_K\}$.

La TM es una TR discreta exacta definida para un conjunto $S = \{(p_k, q_k), k = 1, \dots, K\}$ de ángulos de proyección específicos.

$$M_k(l) = \text{proj}(p_k, q_k, b_l) = \sum_{(i, j) \in L} \delta(b_l - iq_k - jp_k) \quad (5)$$

donde

$$\delta(x) = \begin{cases} 1, & \text{si } x = 0 \\ 0, & \text{si } x \neq 0. \end{cases} \quad (6)$$

Dada una imagen 2D $N \times M$, $I(i, j); i = 1, \dots, N$ y $j = 1, \dots, M$, las proyecciones están dadas por

$$P_k = \sum I(i, j)(\theta_k - iq_k - jp_k), \quad (7)$$

lo que define líneas de proyección q_k, p_k , en la dirección k . De la definición de la TM tenemos:

$$\text{proj}_{p_i, q_i}(m) = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} I(x, y) \delta(m + q_i x - p_i y) \quad (8)$$

donde $\delta(\cdot)$ es la delta de Kronecker, $I(x, y)$ es el valor del pixel, p_i y q_i definen el ángulo de proyección θ_i con desplazamiento x y y , y además $\theta_i = \tan^{-1}(q_i/p_i)$.

La fórmula de Katz [8] da el número de elementos B_i para la proyección definida por p_i y q_i para una imagen de $P \times Q$ pixeles:

$$m = B_i = (Q - 1)|p_i| + (P - 1)|q_i| + 1 \quad (9)$$

donde m será el número de líneas y cuáles pixeles son sumados para una imagen de tamaño $P \times Q$, dados los valores p y q .

Por ejemplo, usando tres ángulos de proyección tenemos (p_i, q_i) con $i = 1, 2, 3$. Si $p_1 = 1$ $q_1 = 0$, dirección $(1, 0)$ y usando una imagen de tamaño 3×3 , tenemos m con tres valores diferentes, como se muestra en la Tabla 1. Así, para $m = 1$, se suman los pixeles $(1, 1)$, $(2, 1)$ y $(3, 1)$, es decir la primera fila.

Tabla 1. Cálculo de m para la dirección $(1,0)$

pixel x,y	$m = y p_1 - x q_1$
1,1	$m=1-0 = 1$
2,1	$m=1-0 = 1$
3,1	$m=1-0 = 1$
1,2	$m=2-0 = 2$
2,2	$m=2-0 = 2$
3,2	$m=2-0 = 2$
1,3	$m=3-0 = 3$
2,3	$m=3-0 = 3$
3,3	$m=3-0 = 3$

Tabla 2. Cálculo de m para la dirección $(1,1)$

pixel x,y	$m = y p_1 - x q_1$
1,1	$m=1-1 = 0$
2,1	$m=1-2 = -1$
3,1	$m=1-3 = -2$
1,2	$m=2-1 = 1$
2,2	$m=2-2 = 0$
3,2	$m=2-3 = -1$
1,3	$m=3-1 = 2$
2,3	$m=3-2 = 1$
3,3	$m=3-3 = 0$

Usando 45 grados, $(p, q) = (1, 1)$, entonces la TM entregará cinco elementos, como se muestra en la Tabla 2.

Ahora m tendrá cinco casos diferentes 0, 1, 2, -1, y -2. Entonces, cuando $m = 0$ se suman los pixeles (1,1), (2,2) y (3,3), correspondiendo a la diagonal principal. Si $m = 1$, solamente se suman los pixeles (1,2) y (2,3). Si $m = 2$, solamente se toma el pixel (1,3), correspondiendo a la esquina inferior izquierda.

Sobre la diagonal principal tenemos: $m = -1$ y se suman los pixeles (2,1) y (3,2). Para $m = -2$ solamente el pixel (3,1), el de la esquina superior derecha. Las Figuras 2 y 3 ilustran el concepto de la TM sobre una imagen 3×3 pixeles.

Diferencias entre la TM y TR

Retomando la información de la literatura [12]-[6]:

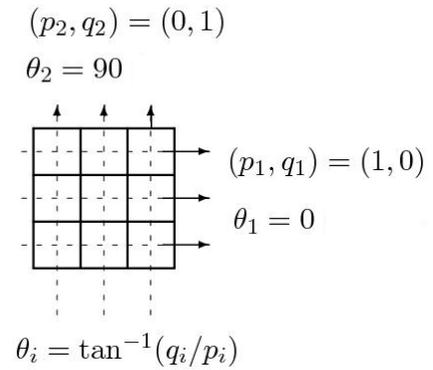


Figura 2. TM a dos ángulos de proyección: 0 y 90 grados.

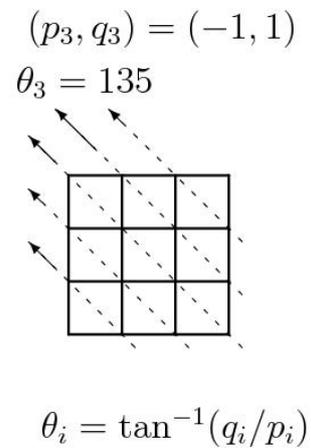


Figura 3. TM a un ángulo de 135 grados.

1. La TM es una forma discreta exacta de la TR. Además la TM tiene una inversa exacta a partir de un número finito de proyecciones discretas (en la TR se usan técnicas de regularización).
2. Otra **diferencia muy importante** es que el espaciamento (tasa de muestreo) de las proyecciones depende del ángulo de proyección y del tamaño de la imagen (la línea de proyección pasa por el centro del pixel).

De la definición de la TM se tiene que

$$\mathcal{M}\{f(x,y)\} = f(x,y)\delta(m + qx - py), \quad (10)$$

donde $\delta(\cdot)$ es la delta de Kronecker. Así la TM es un conjunto de proyecciones donde cada elemento de la transformada corresponde a la suma de los pixeles centrados en la

línea de proyección $m = y p_i - x q_i$, o en forma de ecuación

$$proj_{p_i, q_i}(m) = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} f(x, y) \delta(m + qx - py), \quad (11)$$

donde (x, y) define la posición del pixel.

La diferencia número 2 de la TM es muy importante para el caso de tareas de reconocimiento o de compresión de imágenes, ya que el tamaño de la transformada (número de elementos) puede hacerse grande comparado con el número de elementos de la matriz original de la imagen.

Por ejemplo:

Si tenemos una imagen 3×3 . La dirección de proyección se define por (p_i, q_i) , donde p_i da el desplazamiento en x y q_i el desplazamiento en y , con $\{i = 1, \dots, I\}$, donde I indica el número de proyecciones. Entonces cuando usamos ángulos de 0 y 90 grados tenemos: dirección (1,0), y (0,1), como se ilustra en la Figura 2. Pero si ahora tomamos un ángulo de 135 grados, tenemos el resultado que se muestra en la Figura 3. Es decir, en las dos primeras proyecciones el número de elementos resultantes en cada proyección es de 3. En el caso de la tercera proyección se tienen 5 elementos, en lugar de los 3 anteriores.

Se sabe que para un bloque de dimensión $P \times Q$, la proyección en la dirección (p_i, q_i) estará compuesta de B_i elementos dados por:

$$B_i = (Q - 1)|p_i| + (P - 1)|q_i| + 1 \quad (12)$$

donde i nuevamente indica el número de proyección.

En efecto, para el arreglo de 3×3 , con (1,1) tenemos

$$(3 - 1)(1) + (3 - 1)(1) + 1 = 5. \quad (13)$$

La transformada Mojette es una forma discreta exacta de la transformada de Radon y tiene una inversa exacta a partir de un número finito de proyecciones discretas

Criterio para reconstrucción completa

Se puede determinar [8] si el número de proyecciones es suficiente para reconstruir de forma única la imagen original.

El criterio de Katz indica que un bloque de tamaño $P \times Q$ puede ser reconstruido de manera única si

$$P \leq \sum_{i=1}^I |p_i|, \quad \text{o bien} \quad (14)$$

$$Q \leq \sum_{i=1}^I q_i.$$

Obsérvese que q_i siempre debe ser positivo, ya que \tan^{-1} es periódica en π .

Así pues, cuando se requiere cumplir con este criterio, se observa que el tamaño de la TM es mayor al número de elementos (píxeles) de la imagen original. El ejemplo dado en [12] indica que con un bloque 2×8 , donde $P = 2$, $Q = 8$ (en la referencia invirtieron el orden de P y Q), P define la altura de la imagen, o sea 2. Entonces cuando $Q > P$ (nuevamente invertido en la referencia), la imagen es reconstruible usando P proyecciones, es decir con sólo 2 proyecciones. Usando las direcciones (0,1), (-1,1), y (1,1) se tienen tres vectores de tamaño 8, 9, y 9. Al cumplir con el criterio de reconstructibilidad se tomarían 2 de esos 3 vectores, quedando una representación con 17 elementos, lo cual es mayor al tamaño original de 16 píxeles.

Por supuesto que luego se pueden aplicar algunas técnicas para reducir la cantidad de datos observando correlaciones y usando transformada Wavelet, por ejemplo.

El asunto es que el tamaño de la proyección depende en mucho del ángulo, ya que cambia la tasa de muestreo, como se dijo antes.

Por ejemplo, se dice que si se tiene un arreglo 48×48 y se desea una redundancia de 33%, donde

$$R = \frac{\sum \text{bins}}{\sum \text{pixels}} \quad (15)$$

donde “bins” son los elementos de la TM o tamaño de la proyección, se recomienda usar 4 proyecciones con direcciones $(\pm 15, 2)$, $(\pm 21, 2)$, teniendo

$$\sum_i |p_i| = 15 + 15 + 21 + 21 = 72, \quad (16)$$

con lo cual se cumple el criterio de Katz, pero lo que no indican es que el costo será el tamaño de la TM, ya que para la primera proyección se tendrá

$$B_1 = (Q - 1)|p_1| + (P - 1)|q_1| + 1 \quad (17)$$

$$= (48 - 1)(15) + (48 - 1)(2) + 1$$

$$= 705 + 94 + 1 = 800.$$

Para las otras proyecciones se tendrá $B_2 = 800$, $B_3 = 1082$, y $B_4 = 1082$. Si tomamos las cuatro proyecciones nos da un total de 3764. La redundancia queda de 66% y no de 33% como dice en el artículo.

Con una sola proyección sí se tiene una reducción en los datos pues 48×48 es 2304 y B es 800 o 1082. Pero se requiere más de una proyección y el criterio de Katz nos dice que sean al menos 3. Si tomamos 3 proyecciones el criterio se cumple también y R queda como 16% o 28% según se tomen $2 \times 800 + 1082$, o $2 \times 1082 + 800$.

Un caso de reconocimiento de caracteres

Aunque en la literatura se explica la Redundancia R como un factor de confiabilidad en la transmisión de datos, eso implica cumplir con el criterio de Katz, lo cual a su vez implica una TM de gran tamaño. Pero, en el caso de reconocimiento de caracteres, la reducción del vector de características es muy importante, ya que reduce la carga computacional del reconocimiento. De una imagen de tamaño 16×16 se tienen 256 elementos en total. Se pueden usar sólo dos proyecciones (1,0) y (0,1) que son las de menor tamaño. Así con $P = 16$, $Q = 16$ y con direcciones (1,0) y (0,1) tenemos en cada caso

$$\begin{aligned} B_1 &= (Q-1)|p_1| + (P-1)|q_1| + 1 \\ &= (16-1)(1) + (16-1)(0) + 1 = 16, \end{aligned} \quad (18)$$

$$\begin{aligned} B_2 &= (Q-1)|p_1| + (P-1)|q_1| + 1 \\ &= (16-1)(0) + (16-1)(1) + 1 = 16. \end{aligned} \quad (19)$$

Quedando nuestra representación con un vector de 32 elementos, que ya es muy manejable. Si incluyéramos otra proyección ($i = 3$), digamos a 45 grados (1,1) tendríamos

$$\begin{aligned} B_3 &= (Q-1)|p_1| + (P-1)|q_1| + 1 \\ &= (16-1)(1) + (16-1)(1) + 1 = 31. \end{aligned} \quad (20)$$

Tendríamos pues un vector de características de 63 elementos (teníamos 32).

Conclusiones

La TM no reduce el número de datos si se requiere reconstruir la imagen (invertir la TM). El criterio de Katz indica

el número de proyecciones y direcciones necesarias para ello. Esto significa que el número de elementos de cada proyección sea variable y en aumento.

Si tenemos la condición

$$N \leq \sum_{i=1}^I |p_i|, \quad (21)$$

podríamos investigar cómo encontrar las direcciones que den el menor tamaño en B . Si seleccionamos una $I = 3$

$$|p_1| + |p_2| + |p_3| \geq N. \quad (22)$$

Tomando las proyecciones (1,0) (0,1) y (p_3, q_3) , sólo nos quedaría por encontrar esta última, ya que (1,0) y (0,1) dan el menor tamaño, posiblemente. Así si $N=16$,

$$\begin{aligned} p_3 &\geq N - |p_1| - |p_2| \\ p_3 &\geq 16 - 1 - 0 = 15. \end{aligned} \quad (23)$$

De todas maneras esto daría un tamaño para $B_3 = 241$ elementos!

Consideramos que lo que se debe investigar es qué propiedades tienen las dos primeras proyecciones (a 0 y 90 grados). Además sin cumplir con el criterio de Katz, ver si estas propiedades permiten el reconocimiento de algún alfabeto pequeño. Quizá todo el alfabeto de caracteres alfanuméricos no sea adecuado, si no más bien un tamaño reducido y de ciertas características.

Bibliografía

- [1] http://es.wikipedia.org/wiki/Reconstrucción_tomográfica
- [2] O. Ramos Terrades, E. Valveny, "Radon Transform for Linear Symbol Representation", *Proc. of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)*. Vol. 1, pp. 195-199, Edinburgh, Scotland, 3-6 August, 2003.
- [3] N. Normand, J. P. Guédon, "La transforme Mojette: une representation recordante pour l'image", *Comptes Rendus Academie des Sciences de Paris, Theoretical Comp. SCI. Section*, pp. 124-127, 1998.
- [4] Serfózó P., Vászárhelyi J., Analysis of Mojette Transform Implementation on Field Programmable Gate Array, 7th International Symposium of Hungarian Researches on Computational Intelligence, Nov. 24-25, pp. 255-266, 2006.

- [5] B. Parrein, N. Normand, J.P Guédon, “Multiple Description Coding Using Exact Discrete Radon Transform”, *Proc. of the Data Compression Conference (DCC’01)*, 2001.
- [6] József Vásárhelyi, Péter Serfözö, “Analysis of Mojette Transform Implementation on Reconfigurable Hardware,” *Dagstuhl Seminar Proceedings 06141, Dynamically Reconfigurable Arcuitectures (2006)*.
- [7] Guedon, J. P. – Parrein, B. – Normand, N., *Internet Distributed Image Databases, Int. Comp. Aided Eng., Vol. 8*, pp 205 –214, 2001.
- [8] Katz, M., *Questions of uniqueness and resolution in reconstruction from projections*, Springer Verlag, Berlin, 1977.
- [9] A. Kingston, B. Parrien, and F. Asstrusseau, “Redundant Image Representation via Multi-scale Digital Radon Projections,” *International Conf. on Image Processing*, United States, April 2008.
- [10] A. Kingston, S. Colosio, P. Campist, F Asstrusseau, “Lossless Image ompresion and Selective Encryption Using a Discrete RAdon Transform,” *International Conf. on Image Processing*, United States (2007).
- [11] Turán, J. Ovsenik, L. Benca, M. Turán, J. Jr., *Implementation of CT and IHT Processors for invariant Object Recognition System, Radioengineering, Vol. 13, No 4*, pp. 65-71, 2004.
- [12] J. P. Guédon and N.Normand, “The Mojette Transform: The First Ten Years”, *Discrete Geometry for Computer Imagery*, Poitier, France, 2005.

Acerca del autor o autores

Los autores son profesores de la Facultad de Ingeniería Eléctrica de la Universidad Autónoma de Zacatecas, López Velarde 801, Centro. Zacatecas, Zac, 98000. México. Cubren los programas académicos de Ingeniería en Comunicaciones y Electrónica, y el de Ingeniería en Computación.