

## Mejora Automática de Contraste en Imágenes Médicas usando Fuzzy C-Means

M. en C. Alvaro Anzueto Rios, UPIITA-IPN, Lab. Biomecánica, aanzuetor@ipn.mx.

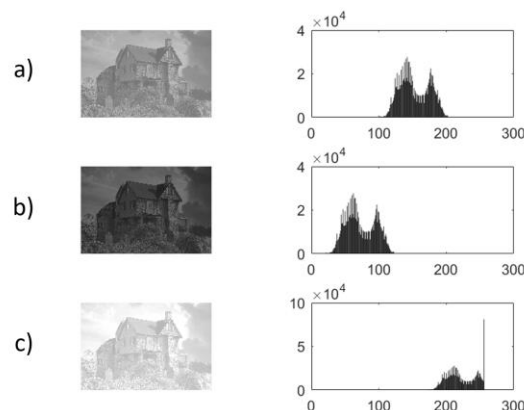
Alumno: Carlos Emiliano Solórzano Espíndola, UPIITA-IPN, Lab. Biomecánica,  
carlosemiliano04@gmail.com

### Resumen

*En el presente trabajo, se muestra una propuesta para restauración contraste para imágenes de ultrasonido, con el objetivo de realzar detalles que permitan identificar características relacionadas a patologías, de manera que sean más fácilmente identificables, la finalidad de esto es ayudar en el diagnóstico médico. El uso del algoritmo por lógica difusa permite en una primera instancia, mediante una etapa de aprendizaje, el ajuste automático a los valores correspondientes de la imagen y en una segunda, crear una nueva imagen usando los valores de pertenencia obtenidos para cada conjunto difuso aprendido y una defusificación propuesta de manera que los conjuntos sean distribuidos uniformemente a lo largo del histograma.*

### Introducción

El contraste se puede definir como una medida de la dispersión de los pixeles en una imagen a lo largo de todos los valores posibles, para el formato de la misma (2). Algunas imágenes contienen una dispersión concentrada en un rango estrecho, siendo así mayormente oscuras, claras o de valores medios; de igual manera pueden existir imágenes con alto contraste que combinan grupos de muy claros y muy oscuros a la vez. Las técnicas de ajuste de contraste tienen como objetivo distribuir de manera uniforme estos valores, mejorando de esta manera la percepción de los objetos en una imagen.



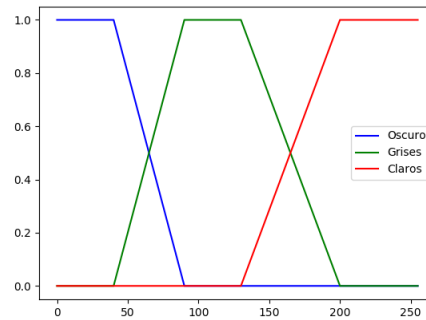
**Figura 1. Imágenes de bajo contraste. a) Valores medios, b) Mayormente oscura, c) Mayormente clara**

El ajuste de contraste, para imágenes médicas, es un campo que mediante años de desarrollo ha logrado ser una herramienta para una mejora en el diagnóstico de enfermedades mediante realce de ciertas características relacionadas a las mismas, así como la investigación para la detección de nuevas características presentes en los estudios y que no son apreciables a simple vista. Estos problemas nacen debido a la relativa baja calidad visual que presentan algunos estudios, al poseer gran parte de la información en una banda relativamente estrecha de valores con una baja varianza, representada en el histograma como un cúmulo. Lo anterior conlleva a una imagen con un bajo contraste, lo cual no solo está presente en imágenes médicas sino también en imágenes de percepción remota, microscopía de electrones e incluso fotografías comunes (2).

A lo largo de los años diversos métodos se han desarrollado a través de representaciones digitales de las mismas para mejorar el contraste de manera automática y realzar los detalles de manera que sean más fáciles de visualizar. Técnicas como la ecualización de histograma abordan el problema estadístico, en el cual se calcula la distribución de probabilidad de la imagen para cada valor dentro de la escala de grises dependiendo del formato de la imagen, una vez obtenida la distribución se calcula una distribución acumulativa correspondiente a una integral discreta de la misma. Se propone entonces una transformación para la cual la nueva distribución de la imagen sea constante para todos los niveles de la escala de grises, proponiendo nuevos valores para cada uno de los píxeles en la imagen original, de manera que su nueva distribución acumulativa sea una recta. Sin embargo las imágenes médicas pueden poseer grandes cúmulos de valores cercanos a 0, lo cual provoca que, al aplicar la ecualización, el valor mínimo de la nueva imagen sea muy alto.

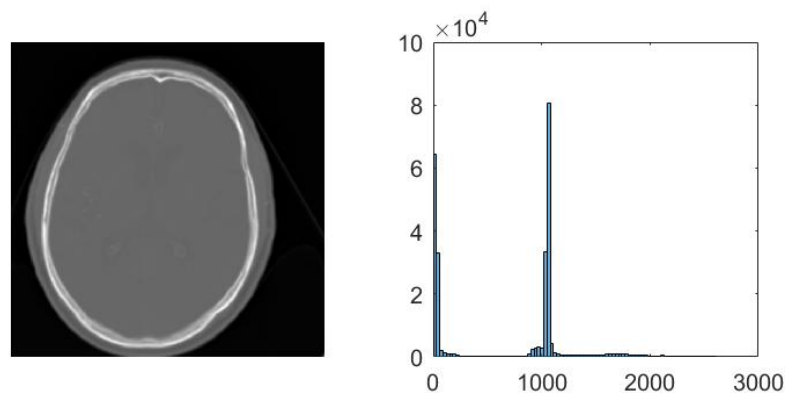
## Metodología

La lógica difusa está basada en la fusificación de variables de manera que posean un grado de pertenencia a diferentes variables lingüísticas, en vez de pertenencia booleana (0 o 1), que conforman conjuntos con diversas formas, siendo posible que contenga valores de pertenencia para más de un conjunto (3); estas pueden ser propuestas directamente por el usuario como se muestra en la figura 2. Mediante lo anterior se pueden establecer reglas que relacionan los conjuntos de las variables de entrada a la de salida, la cual se obtiene mediante diferentes métodos de defusificación, estas reglas son establecidas basadas en la experiencia del usuario, con la finalidad de describir un sistema complejo con operaciones matemáticas básicas.



**Figura 2. Grupos difusos relacionados a variables lingüísticas para valores de escala de grises**

Para la mejora de contraste en este caso se utiliza el algoritmo fuzzy C-means para la obtención de los conjuntos difusos de manera automática. Fuzzy C-means calcula iterativamente, dado un número de conjuntos o clases propuestos, los centros (*means*) que representan el conjunto de datos, de manera que sean agrupables considerando su relación o distancia difusa a cada uno de estos, el resultado son conjuntos que representan la imagen basado en su similitud dados los datos que la representan. Esta técnica ha sido usada previamente para la segmentación de tejidos en estudios dado que cada uno de estos puede ser ubicado en el histograma por grandes cúmulos de píxeles con valores similares en escala de grises debido a que el estudio devuelve valores similares para cada tejido (1, 4), como se puede ver en la figura 3.



**Figura 3. Imagen TAC y su respectivo histograma con cúmulos localizados cerca de 0, 800, 1000 y 1700**

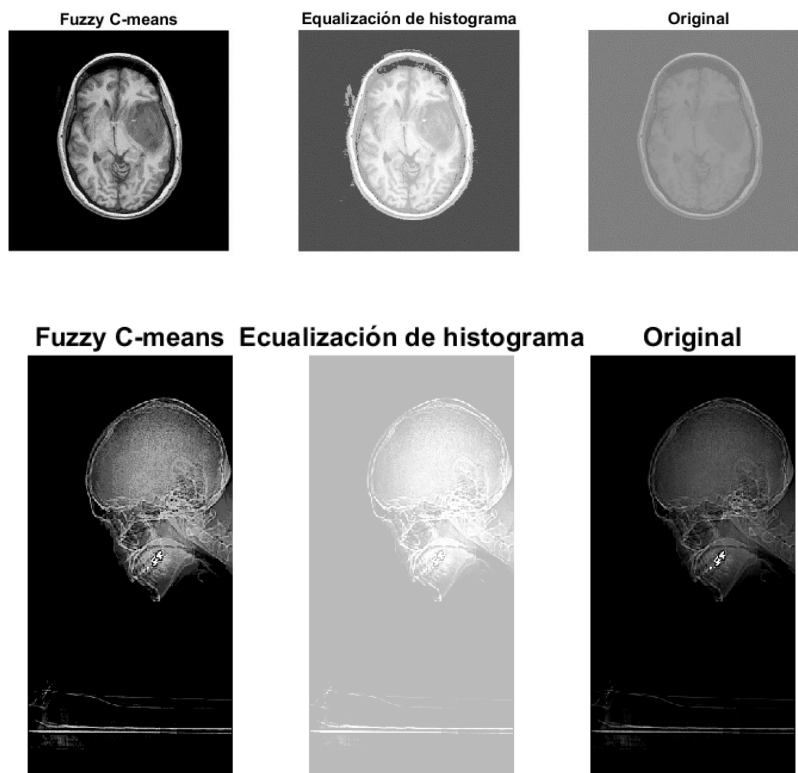
Sin embargo, estos cúmulos pueden estar relativamente cerca, lo cual conlleva a una dificultad para distinguir entre ambos grupos en la visualización de la misma. Ante esto se propone un método que, una vez que los grupos han sido identificados por el algoritmo, sean distribuidos a lo largo del histograma de manera uniforme, es decir, que la diferencia de nivel de gris entre píxeles.

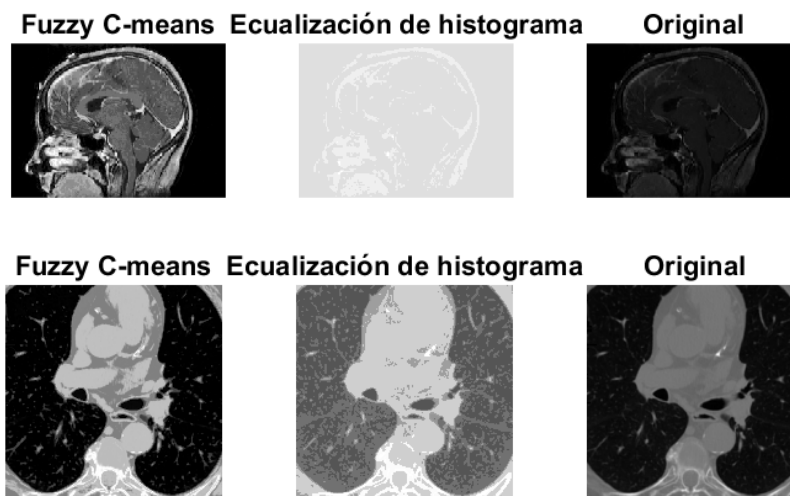
El algoritmo entonces sigue los siguientes pasos:

- Proposición de número de conjuntos, estos serán el mismo número de entrada que de salida.
- Estimación de los conjuntos por fuzzy c-means
- Proposición de los conjuntos de salida de manera que
- Defusificación usando los conjuntos de salida, en este caso se usaron simplemente singletones y el método del centroide.
- Generación de nueva imagen

## Resultados

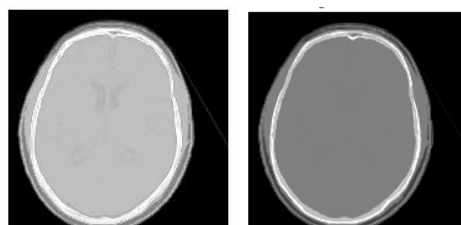
Para la evaluación del algoritmo se ha probado con diversas imágenes de estudios médicos obtenidas de la base de datos de *David Clunie's Medical Image Format Site (6)* la cual alberga diversos estudios de TAC y resonancia magnética (MRI) guardados en formato DICOM de 16 bits. Con objetivo de comparación se muestran junto al resultado de ecualización de histograma y la imagen original. Ecualización de histograma fue elegida ya que es uno de los métodos más usados en la actualidad debido a su fácil implementación y bajo tiempo de ejecución.





**Figura 4. Resultados comparando el resultado del algoritmo propuesto a ecuación de histograma y la imagen original**

También se encontró que en ocasiones una diferente inicialización de parámetros conlleva a una diferente imagen final para el mismo número de clases, esto se puede deber a mínimos locales que en posteriores iteraciones no permiten llegar al resultado óptimo.



**Figura 5. Resultado del algoritmo usando 5 clases con diferente inicialización**

El algoritmo era ejecutado iterativamente hasta que la diferencia total de los centros con los de la iteración anterior fuera menor a cierto valor umbral elegido por el usuario. Por tanto el tiempo de ejecución variaba en gran medida dado el número de centros a encontrar y una inicialización cerca o distante del resultado, esto se puede observar en la figura 6.



**Figura 6. Tiempo de ejecución para encontrar diferente número de centros**

## Conclusiones

El algoritmo propuesto cumple su propósito de mejorar la visualización de las imágenes mediante el realce de contraste y en general logra una buena dispersión a lo largo del histograma identificando los cúmulos que representan los tejidos y distribuirlos de manera uniforme. Al usar el método de defusificación del centroide combinado con la proposición automática de singletons de manera que aseguren la mayor dispersión a lo largo del histograma permiten la generación de una nueva imagen con contraste para tejidos con valores cercanos en la imagen original. Como desventajas del algoritmo se tiene que es susceptible al ruido por lo que requiere de técnicas de pre-procesamiento para limpiar la imagen, así como un tiempo de ejecución relativamente alto para caso en que se busquen un gran número de centros.

## Referencias

1. Adhikari, S. K., et. al. (2015). A Spatial Fuzzy C-means Algorithm with Application to MRI Image Segmentation: IEEE
2. González, R., Woods, R. (2001). Digital Image Processing. New Jersey, USA: Prentice Hall
3. Rich, E. & Knight, K. (2010). Inteligencia Artificial (2da ed.). Madrid: McGraw-Hill
4. T. Chaira. A novel intuitionistic fuzzy c-means clustering algorithm and its application to medical images (2011). Applied Soft Computing, no. 11, pp. 1711-1717, Elsevier.
5. T. Qi and Y. Lu. Color image segmentation –an innovative approach (2002). Pattern Recognition, no. 35, pp.395-405, Elsevier.
6. Clunie, D. (2016, December, 19). CT and MRI Multi-frame test DICOM images at NEMA. Recuperado el 29 de marzo de 2017 de <ftp://medical.nema.org/medical/dicom/Multiframe>