Paquete de Python para el Procesamiento de Imágenes Usando Varios Modelos no Lineales

Carlos Toledo Silva

Tutor
Damian Valdés Santiago

Universidad de La Habana

5 de diciembre de 2022

Introducción – Surgimiento de los modelos LIP

Puntos claves:

- Bajo ciertas circunstancias el Procesamiento Lineal Clásico de Imágenes (CLIP) demuestra sus limitaciones.
- Alan V. Oppenheim introduce la teoría homomórfica.
- Michel Jourlin y Jean-Charles Pinoli proponen el primer modelo logarítmico para el procesamiento de imágenes definido desde un punto de vista físico.
- Vasile Patrascu y Vasile Buzuloiu proponen otro modelo logarítmico definido desde un punto de vista matemático.
- Constantin Vertan y colaboradores proponen un modelo pseudo-logarítmico.
- Se siguen desarrollando modelos no lineales, muchos inspirados en los anteriores.
- Con estos modelos se han desarrollado diversas aplicaciones: corrección de iluminación, mejora de contraste, mejora de imagen en color, ecualización de histogramas, mejora de rango dinámico, detección de bordes, etc.

Introducción – Objetivos

Objetivos generales:

- Diseño e implementación de un módulo de Python donde se encuentren los modelos no lineales para el procesamiento de imágenes abordados en este trabajo.
- Evaluación de los diferentes modelos en cuanto a las diferentes operaciones, algoritmos, tipos de imágenes y situaciones

Objetivos específicos:

- Revisar la literatura referente a los modelos no lineales para el procesamiento de imágenes.
- Implementación en Python de los diferentes modelos a partir de la bibliografía consultada.
- Creación de *datasets* con diferentes tipos de imágenes para evaluar los modelos.
- Selección de métricas para la evaluación de los modelos.
- Evaluación de los modelos.
- Arribar a conclusiones sobre la eficacia y utilidad de estos modelos.

Estado del Arte – Antecedentes Matemáticos

Definición

Sean $(E, K, \oplus, \otimes)$ *y* $(F, K, +, \cdot)$ *dos estructuras algebraicas,* $K \subseteq \mathbb{R}$*, se denomina homomorfismo a la relación* $\varphi : E \to F$ *que satisface:*

$$\varphi(u \oplus v) = \varphi(u) + \varphi(v), \forall u, v \in E, \tag{1}$$

$$\varphi(\alpha \otimes u) = \alpha \cdot \varphi(u), \forall u \in E, \forall \alpha \in K. \tag{2}$$

La admisión de un inverso y la existencia de biyectividad indica la presencia de un isomorfismo.

En este trabajo:

- $E \subseteq \mathbb{R}$, conjunto de definición de las imágenes.
- $F \subseteq \mathbb{R}$, conjunto de definición de las estructuras vectoriales. Si $u \in E$ entonces $\varphi(u) \in F$
- $I(D, E)$ conjunto de imágenes definidas en $D \subset \mathbb{R}^2$ y valoradas en E.
- $I(D, F)$ conjunto de las estructuras vectoriales definidas en $D \subset \mathbb{R}^2$ y valoradas en F. Si $f \in I(D, E)$ entonces $\varphi(f) \in I(D, F)$.

Ejemplos:

- Si los valores de la imagen son intensidades, como cualquier plano en representación de color RGB, el conjunto E tiene la forma $[0, M)$;
- En el caso del espacio YUV, para los canales de diferencias de color, el conjunto E tiene una forma simétrica, como $(-M/2, +M/2)$
- Una imagen en niveles de gris toma sus valores (niveles de gris) en la escala de grises $[0, M)$. Para imágenes de 8 bits, $M = 256$ y los 256 niveles de gris están en la escala de números enteros $[0, \cdots, 255]$

Estado del Arte – Antecedentes Matemáticos

Con respecto a la restricción de biyectividad y la existencia de φ^{-1}, las leyes de definición están determinadas por:

$$u \oplus v = \varphi^{-1}(\varphi(u) + \varphi(v)), \tag{3}$$

$$\alpha \otimes u = \varphi^{-1}(\alpha \cdot \varphi(u)). \tag{4}$$

Propiedad de cierre:

$$\forall u, v \in E, z = u \oplus v \Rightarrow z \in E, \tag{5}$$

$$\forall u \in E, \forall \alpha \in K, z = \alpha \otimes v \Rightarrow z \in E. \tag{6}$$

Estas propiedades se mantienen bajo la hipótesis de biyectividad asumida ya que:

$$z = \varphi^{-1}(\varphi(u) + \varphi(v)) \wedge \forall x \in F, \varphi^{-1}(x) \in E \Rightarrow z \in E. \tag{7}$$

$$z = \varphi^{-1}(\alpha \cdot \varphi(u)) \wedge \forall x \in F, \varphi^{-1}(x) \in E \Rightarrow z \in E. \tag{8}$$

Estado del Arte – Antecedentes Matemáticos

Definición

Dadas las dos leyes operativas, $\oplus$, $\otimes$, el conjunto vectorial E y el conjunto escalar K, la definición formal de espacio vectorial implica varias propiedades:

- *La ley de la suma debe ser asociativa, conmutativa, debe tener elemento identidad y elemento inverso*
- *La distributividad debe mantenerse para la multiplicación escalar sobre la suma de vectores y para la multiplicación escalar en el espacio de los escalares.*
- *La multiplicación escalar debe tener elemento identidad y ser compatible con la multiplicación en el espacio de los escalares.*

Las propiedades conmutativas, asociativas y distributivas se pueden comprobar bajo la hipótesis de biyectividad y la admisión de inverso asumidas.

Elemento identidad con respecto a la suma:

$$\forall u \in E, \exists u_0 : u \oplus u_0 = u | \varphi(u_0) = 0. \tag{9}$$

Elemento inverso con respecto a la suma:

$$\forall u \in E, \exists u^- \Rightarrow u \oplus u^- = u_0 | \varphi(u) + \varphi(u^-) = 0 \Rightarrow \varphi(u^-) = -\varphi(u). \tag{10}$$

Elemento identidad de la multiplicación escalar tiene que ser 1:

$$\forall u \in E - \{u_0\}, \exists \alpha_1 : \alpha_1 \otimes u = u | \alpha_1 = 1. \tag{11}$$

Definición

Sea V un espacio vectorial real y K, el conjunto escalar. Un cono C en V es un subconjunto de V que satisface:

1. *si $u_1, u_2 \in C$ y $\alpha, \beta \in K : \alpha, \beta \geq 0$, entonces $\alpha u_1 + \beta u_2 \in C$.*
2. *$C \cap (-C) = 0$.*

Estado del Arte – Modelo LIP

Nombre: Modelo Logarítmico Clásico para el Procesamiento de Imágenes.
Propuesto por: Michel Jourlin y Jean-Charles Pinoli.
Dominio: $E = [0, M) \vee E = (-\infty, M)$.
Funciones de cambio: $v = M - u$, $u = M - v$.
Suma: $v_1 \oplus v_2 = v_1 + v_2 - \frac{v_1 v_2}{M}$.
Multiplicación por un escalar: $\lambda \otimes v = M - M\left(1 - \frac{v}{M}\right)^{\lambda}$.
Opuesto: $\ominus v = -\frac{v}{1-\frac{v}{M}}$.
Resta: $v_1 \ominus v_2 = \frac{v_1 - v_2}{1-\frac{v_2}{M}}$.
Función del isomorfismo: $\varphi(v) = -M \ln\left(1 - \frac{v}{M}\right)$.
Inversa de la función del isomorfismo: $\varphi^{-1}(x) = M\left(1 - e^{-\frac{x}{M}}\right)$.
Propiedades:

- La suma $\oplus$ es asociativa, conmutativa, tiene un elemento identidad: 0 y cada elemento tiene un opuesto.
- La multiplicación escalar $\otimes$ es asociativa, tiene elemento identidad: 1 y la distributividad se mantiene para la multiplicación escalar sobre $\oplus$ y para la multiplicación escalar en el espacio de los escalares.

Si $E = [0, M)$ entonces $(E, \mathbb{R}^+, \oplus, \otimes)$ es un cono positivo.
Si $E = (-\infty, M)$ entonces $(E, \mathbb{R}, \oplus, \otimes)$ es un espacio vectorial.

Estado del Arte – Modelo HLIP

Nombre: Modelo Logarítmico Homomórfico para el Procesamiento de Imágenes.
Propuesto por: Vasile Patrascu y Vasile Buzuloiu.
Dominio: $E = (-1, 1)$.
Funciones de cambio: $v = \frac{2}{M}\left(u - \frac{M}{2}\right)$, $u = \frac{M}{2}(v+1)$.
Suma: $v_1 \oplus v_2 = \frac{v_1+v_2}{1+v_1v_2}$.
Multiplicación por un escalar: $\lambda \otimes v = \frac{(1+v)^\lambda-(1-v)^\lambda}{(1+v)^\lambda+(1-v)^\lambda}$, $\lambda \in \mathbb{R}$.
Opuesto: $\ominus v = -v$.
Resta: $v_1 \ominus v_2 = \frac{v_1-v_2}{1-v_1v_2}$.
Función del isomorfismo: $\varphi(v) = \frac{1}{2}\ln\left(\frac{1+v}{1-v}\right)$.
Inversa de la función del isomorfismo: $\varphi^{-1}(x) = \frac{e^{2x}-1}{e^{2x}+1}$.
Propiedades:

- La suma $\oplus$ es asociativa, conmutativa, tiene un elemento identidad: 0 y cada elemento tiene un opuesto.
- La multiplicación escalar $\otimes$ es asociativa, tiene elemento identidad: 1 y la distributividad se mantiene para la multiplicación escalar sobre $\oplus$ y para la multiplicación escalar en el espacio de los escalares.

La estructura $(E, \mathbb{R}, \oplus, \otimes)$ es un espacio vectorial.

Estado del Arte – Modelo PSLIP

Nombre: Modelo Pseudo-logarítmico para el Procesamiento de Imágenes.
Propuesto por: Constantin Vertan y colaboradores.
Dominio: $E = [0, 1)$.
Funciones de cambio: $v = \frac{u}{M}$, $u = M \cdot v$.
Suma: $v_1 \oplus v_2 = \frac{v_1 + v_2 - 2v_1 v_2}{1 - v_1 v_2}$.
Multiplicación por un escalar: $\lambda \otimes v = \frac{\lambda v}{1+(\lambda-1)v}$, $\lambda \in \mathbb{R}^+$.
Resta: $v_1 \ominus v_2 = \frac{v_1 - v_2}{1 + v_1 v_2 - 2v_2} : v_1 \geq v_2$.
Función del isomorfismo: $\varphi(v) = \frac{v}{1-v}$.
Inversa de la función del isomorfismo: $\varphi^{-1}(x) = \frac{x}{1+x}$.
Propiedades:

- La suma $\oplus$ es asociativa, conmutativa y tiene un elemento identidad: 0.
- La multiplicación escalar $\otimes$ es asociativa, tiene elemento identidad: 1 y la distributividad se mantiene para la multiplicación escalar sobre $\oplus$ y para la multiplicación escalar en el espacio de los escalares.

La estructura $(E, \mathbb{R}^+, \oplus, \otimes)$ es un cono positivo. La extensión a un espacio vectorial se puede lograr mediante el uso de la función generadora definida en $(-1, 1) \to (-\infty, +\infty) : \varphi(v) = \frac{v}{1-|v|}$.

Estado del Arte – Modelo PLIP

Nombre: Modelo Logarítmico Parametrizado para el Procesamiento de Imágenes.
Propuesto por: Karen Panetta y colaboradores.
Dominio: $E = [0, M) \vee E = (-\infty, M)$.
Funciones de cambio: $v = \mu(M) - u$, $u = \mu(M) - v$.
Suma: $v_1 \oplus v_2 = v_1 + v_2 - \frac{v_1 v_2}{\gamma(M)}$.
Multiplicación por un escalar: $c \otimes v = \gamma(M) - \gamma(M)(1 - \frac{v}{\gamma(M)})^c$.
Opuesto: $\ominus v = -\frac{v}{1 - \frac{v}{\gamma(M)}}$.
Resta: $v_1 \ominus v_2 = \frac{v_1 - v_2}{1 - \frac{v_2}{k(M)}}$.
Función del isomorfismo: $\varphi(v) = -\lambda(M) \ln^{\beta}(1 - \frac{v}{\lambda(M)})$.
Inversa de la función del isomorfismo: $\varphi^{-1}(x) = \lambda(M)(1 - (e^{-\frac{x}{\lambda(M)}})^{\frac{1}{\beta}})$.

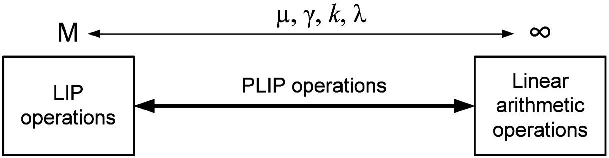

Figura: Operaciones en el modelo PLIP

Propiedades del modelo PLIP:

- En el modelo PLIP, las operaciones aritméticas lineales se reemplazan con nuevas operaciones de la misma manera que en el modelo LIP tradicional.
- Las operaciones en el modelo PLIP son iguales a las del modelo tradicional LIP cuando $\mu(M) = \gamma(M) = k(M) = \lambda(M) = M$ y $\beta = 1$.
- Las operaciones en el modelo PLIP se asemejan a las operaciones aritméticas lineales a medida que $\gamma(M)$, $k(M)$ y $\lambda(M)$ se aproximan al infinito.
- Las operaciones PLIP pueden generar más casos entre los dos casos extremos del LIP y las operaciones aritméticas lineales cuando los parámetros μ, γ, k y λ cambian dentro de $[M, +\infty)$.
- Las operaciones PLIP cumplen las leyes de asociatividad, conmutatividad, elemento identidad, elemento opuesto para la suma y las propiedades distributivas.

Estado del Arte – Modelo SLIP

Nombre: Modelo Logarítmico Simétrico para el Procesamiento de Imágenes.
Propuesto por: Laurent Navarro y colaboradores.
Dominio: $E = (-M, M)$.
Suma: $v_1 \oplus v_2 = Msgn(v_1 + v_2)\left[1 - \left(1 - \frac{|v_1|}{M}\right)^{\gamma_1}\left(1 - \frac{|v_2|}{M}\right)^{\gamma_2}\right]$,
$\gamma_1 = \frac{sgn(v_1)}{sgn(v_1+v_2)}$, $\gamma_2 = \frac{sgn(v_2)}{sgn(v_1+v_2)}$.
Multiplicación por un escalar: $\lambda \otimes v = Msgn(\lambda v) = \left[1 - \left(1 - \frac{|v|}{M}\right)^{|\lambda|}\right]$.
Opuesto: $\ominus v = (-1) \otimes v$.
Resta: $v_1 \ominus v_2 = v_1 \oplus (-1) \otimes v_2$.
Función del isomorfismo: $\varphi(v) = -Msgn(v)\ln\left(1 - \frac{|v|}{M}\right)$.
Inversa de la función del isomorfismo: $\varphi^{-1}(x) = Msgn(x)\left(1 - e^{-\frac{|x|}{M}}\right)$.

Estado del Arte – Modelo SLIP

Propiedades:

- La suma $\oplus$ es asociativa, conmutativa, tiene un elemento identidad: 0 y cada elemento tiene un opuesto.
- La multiplicación escalar $\otimes$ es asociativa, tiene elemento identidad: 1 y la distributividad se mantiene para la multiplicación escalar sobre $\oplus$ y para la multiplicación escalar en el espacio de los escalares.

La estructura $(E, \mathbb{R}, \oplus, \otimes)$ es un espacio vectorial.
Relación Matemática entre los Modelos LIP y SLIP:

- Desde un punto de vista matemático, cuando $v \in [0, M)$, las dos funciones generadoras $\varphi_{LIP}(v)$ y $\varphi_{SLIP}(v)$ son exactamente las mismas (sin tener en cuenta la inversión de la escala de grises).
- El modelo SLIP($\odot$) es el mismo que el modelo LIP($\boxdot$) para el rango de valores de píxeles $[0, M)$
- Difieren entre sí en la forma en que tratan los valores negativos. Por ejemplo, cuando $\lambda < 0$ y $v > 0$, tenemos $\lambda \boxtimes v \in (-\infty, 0)$ y $\lambda \otimes v \in (-M, 0)$. Del mismo modo, cuando $-\infty < v_1 < v_2 < M$ se tiene $v_1 \boxminus v_2 \in (-\infty, 0)$ y cuando $-M < v_1 < v_2 < M$, se tiene $v_1 \ominus v_2 \in (-M, 0)$.

Propuesta – Modelo PPSLIP

Nombre: Modelo Pseudo-logarítmico Parametrizado para el Procesamiento de Imágenes.

Versión inicial: Parametrizar las funciones de cambio

Se parametrizan solo las funciones de cambio de la siguiente forma: $v = \frac{u}{\delta(M)}$ y $u = \delta(M)v$, tal que $\delta(M) \geq M$ y mantener el dominio y el resto de operaciones del modelo PSLIP.

Ventaja: Se parametriza el modelo para evitar la pérdida de información.

Desventaja: Las operaciones se parametrizan con el mismo parámetro.

Versión final: Eliminar las funciones de cambio y hacerlos directamente

Dominio: $E = [0, M)$.

Suma: $u_1 \oplus u_2 = \frac{u_1 + u_2 - \frac{2u_1u_2}{\gamma(M)}}{1 - \frac{u_1u_2}{\gamma(M)^2}}$.

Multiplicación por un escalar: $c \otimes u = \frac{cu}{1 + \frac{(c-1)u}{\gamma(M)}}$.

Resta: $u_1 \ominus u_2 = \frac{u_1 - u_2}{1 + \frac{u_1u_2}{k(M)^2} - \frac{2u_2}{k(M)}} : u_1 \geq u_2$.

Función del isomorfismo: $\varphi(u) = \frac{u}{\lambda(M) - u}$.

Inversa de la función del isomorfismo: $\varphi^{-1}(x) = \lambda(M)\frac{x}{1+x}$.

Propuesta – Modelo PPSLIP

Propiedades del modelo PPSLIP:

- En el modelo PPSLIP, las operaciones aritméticas lineales se reemplazan con nuevas operaciones de la misma manera que en el modelo PSLIP tradicional.
- Las operaciones en el modelo PPSLIP son iguales a las del modelo tradicional PSLIP cuando $\gamma(M) = k(M) = \lambda(M) = M$.
- Las operaciones en el modelo PLIP se asemejan a las operaciones aritméticas lineales a medida que $\gamma(M),\ k(M)$ y $\lambda(M)$ se aproximan al infinito.
- Las operaciones PPSLIP pueden generar más casos entre los dos casos extremos del PSLIP y las operaciones aritméticas lineales cuando los parámetros γ, k y λ cambian dentro de $[M, +\infty)$.
- Las operaciones PPSLIP cumplen las leyes de asociatividad, conmutatividad, elemento identidad y las propiedades distributivas.

Propuesta – Módulo Implementado – Estructuras

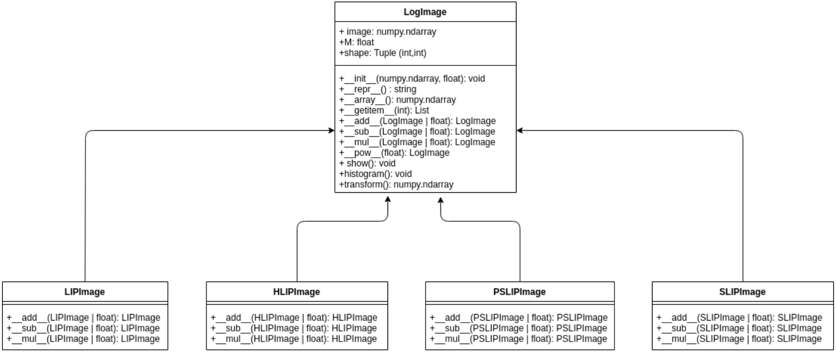

Figura: Diagrama de clases de las estructuras no lineales.

Propuesta – Módulo Implementado – Espacios

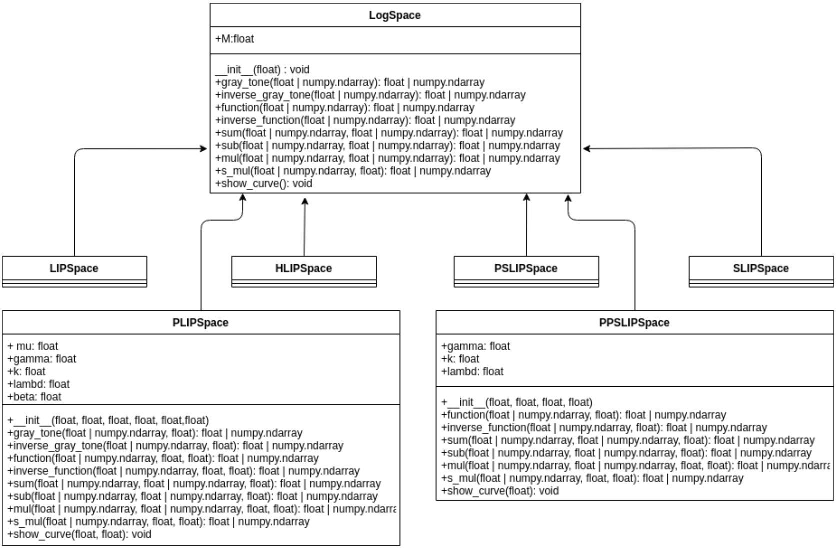

Figura: Diagrama de clases de los espacios no lineales.

Propuesta – Módulo Implementado – Métricas – EMEE

La Medida de Mejora por Entropía se calcula dividiendo una imagen I en $k_1 \times k_2$ bloques, obteniendo el máximo local $I_{maxk,l}$ y el mínimo $I_{mink,l}$ dentro de cada bloque individualmente, y luego procesándolos usando la siguiente ecuación:

$$EMEE_{\alpha,k_1,k_2=\frac{1}{k_1 k_2}} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \alpha \left(\frac{I_{maxk,l}}{I_{mink,l}} \right)^{\alpha} \ln \left(\frac{I_{maxk,l}}{I_{mink,l}} \right), \qquad (12)$$

donde α es una constante que puede ayudar a seleccionar los parámetros. Se eligió $\alpha = 1$ y el tamaño de bloque 4×4, 4×5, 5×4 y 5×5, según las dimensiones de la imagen.

El mejor parámetro (óptimo) se obtiene si se cumple la siguiente condición:

$$EMEE_{optimal} = max_{local}(EMEE(\alpha, \mu, \gamma, k, \lambda, \beta)) \qquad (13)$$

Esta medida se utilizó para obtener las mejores combinaciones de parámetros en el modelo PLIP.

Propuesta – Módulo Implementado – Métricas – C_p

En bibliografía consultada, denotando $p_i = (x_i, y_i)$, los pares de coordenadas que definen la posición espacial de un píxel en una imagen, el contraste absoluto entre dos píxeles distintos $p_1, p_2 \in D$, para una imagen, $f \in I(D, E)$ se define por la relación:

$$C_A(p_1, p_2) = \frac{1}{d(p_1, p_2)} \cdot \frac{|f(p_1) - f(p_2)|}{1 - \frac{f(p_1) \cdot f(p_2)}{M^2}}, \tag{14}$$

Luego de la realización de una serie de experimentos, se decidió modificar la ecuación anterior, cambiando la definición a:

$$C_A(p_1, p_2) = \frac{|f(p_1) - f(p_2)|}{d(p_1, p_2)} \cdot \frac{256}{M} \tag{15}$$

El contraste para un píxel arbitrario $p \in D$, para una imagen $f \in I(D, E)$, se define por la media del contraste absoluto entre el píxel p y los píxeles $(p_i)_{i=1,\cdots,n}$ que pertenecen a una vecindad V:

$$C(p) = \frac{1}{n} \sum_{i=1}^{n} C_A(p, p_i). \tag{16}$$

Luego el valor de C_p de una imagen es el promedio de los valores de $C(p)$ de los pixeles que la conforman.

Propuesta – Módulo Implementado – Algoritmos

Detección de bordes

Para esto se pueden utilizar diferentes filtros como Sobel, Prewitt, Scharr, etc. En particular los tres mencionados anteriormente aparecen implementados en la librería `skimage.filters` de Python.

Unsharp masking

Unsharp Masking es una técnica para mejorar la calidad de una imagen. Esta técnica consiste en determinar los bordes de una imagen y luego fusionar la imagen original con la imagen de bordes de dicha imagen.

Propuesta – Módulo Implementado – Algoritmos

Transformación Afín

Considérense la transformación afín en el conjunto de imágenes $I(D,E)$, definida a continuación: $\psi : I(D,E) \to I(D,E)$

$$\psi(f) = \lambda \otimes (f \oplus \tau), \tag{17}$$

donde $\lambda \in \mathbb{R}, \tau \in E$ y $f \in I(D,E)$.
Los parámetros (λ, τ) se eligen de tal manera que se obtenga una nueva imagen muy cercana a una imagen con una distribución uniforme de los niveles de gris en el conjunto $E = (a,b)$.
Bajo este criterio la imagen mejorada debe tener su media $\mu_u = \frac{a+b}{2}$ y su varianza $\sigma_u^2 = \frac{(b-a)^2}{12}$.
En estas condiciones para cualquier imagen f con la media μ_f y la varianza σ_f^2, la transformación afín se convierte en:

$$\psi(f) = \frac{\sigma_u}{\sigma_f} \otimes (f \ominus \mu_f). \tag{18}$$

Importante: Este algoritmo solo se puede utilizar en espacios acotados.

Detalles de implementación – Modelos

Las imágenes en niveles de gris se definen en el intervalo $[0, M)$.
Problema con los modelos LIP y HLIP pues, si u es un nivel de gris, v su tono de gris correspondiente y $u = 0$ entonces:

- LIP: $v = M - u = M - 0 = M$. Contradicción $E = [0, M)$.
- HLIP: $v = \frac{2}{M}(u - \frac{M}{2}) = \frac{2}{M}(0 - \frac{M}{2}) = \frac{2}{M}(-\frac{M}{2}) = -1$. Contradicción $E = (-1, 1)$

Solución: Antes de cambiar la imagen de niveles de gris a tonos de gris, esta primero se procesa de tal forma que el menor valor admisible sea un número cercano a cero, pero mayor que este. En esta implementación se decidió utilizar el número 0,0001.
Esto también se hizo para el modelo PLIP ya que este constituye una generalización del modelo LIP.
Todos los modelos implementados, tanto en forma de estructura como de espacio, se encuentran en el submódulo "models". Cada uno con su archivo correspondiente

Detalles de implementación – Métricas – EMEE