

# FUSIÓN MORFOLÓGICA DE IMÁGENES IR Y VISUALES UTILIZANDO EL MODELO LIP

**Oscar Ricardo Delfín Santiesteban**

Cidesi

*oscar.delfin.santiesteban@gmail.com*

**Iván Ramón Teról Villalobos**

Cideteq

*iterol@cideteq.mx*

## **Resumen**

La fusión de imágenes es el proceso de combinar la información de una escena que proviene de dos o más imágenes fuente en una sola con una mejor percepción visual y espacial que puede proporcionar detalles que en su conjunto, no pueden ser observados en las imágenes por separado. En este estudio, se presenta una metodología que permite realizar este procedimiento combinando el modelo de procesamiento logarítmico de imágenes (LIP Model) y las transformaciones morfológicas por reconstrucciones.

**Palabras Claves:** Imagen Visual, imagen IR, modelo LIP, morfología matemática.

## **Abstract**

*Image fusion is the process of combining information from a scene that comes from two or more source images into a single one with better visual and spatial perception that can provide details that as a whole cannot be seen in separate images. In this study, a methodology is presented that allows performing this procedure combining the logarithmic image processing model (LIP Model) and the morphological transformations by reconstructions.*

**Keywords:** LIP Model, IR Image, mathematical morphology, visual Image.

## 1. Introducción

Capturar y registrar nuestro entorno, a través de imágenes o videos, ha sido una de las actividades de las que se ha valido el hombre para resolver algunas problemáticas o tener evidencia de que ciertos eventos ocurrieron. Para llevar a cabo la captura de las escenas de interés, hacemos uso de determinados sensores que se ajustan a las condiciones del entorno de manera tal, que nos provean de información visual suficiente para tomar alguna decisión. El gran avance tecnológico nos permite hoy en día contar con diferentes tipos de sensores de los cuales, podemos obtener imágenes bajo condiciones muy particulares. Por ejemplo, cámaras visuales, cámaras de visión nocturna, cámaras infrarrojas, cámaras de longitud de onda milimétrica, cámaras de rayos X o cámaras pancromáticas [Omar, 2014]. Las imágenes que se generan de estas tecnologías, aportan diferente tipo de información de una misma escena. Por ejemplo, en la figura 1a, se presenta una imagen visual tomada de noche en la que no se puede apreciar que hay una persona cruzando. Esta información es revelada en la toma 1b, que corresponde a la misma escena, pero la imagen es infra-roja. Observe como, una misma escena con dos sensores que registran cosas diferentes, proveen información complementaria.

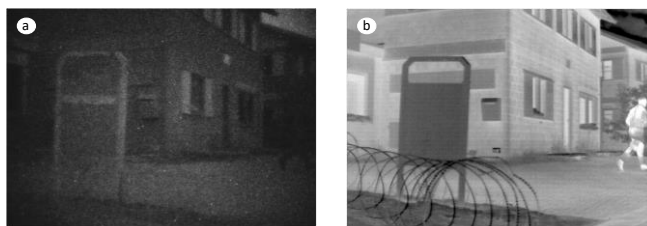


Figura 1 Imagen cortesía de TNO Human Factors Institute, The Netherlands

El tener la posibilidad de unir dos o más fuentes de información provenientes de dos o más tecnologías diferentes y obtener en una sola toma lo mejor de ambas permite observar detalles y características de una misma escena que tal vez de forma separada no es posible percibir. Al proceso de combinar imágenes provenientes de diferentes sensores y generar otra con las mejores características de estas es lo que se conoce como fusión de imágenes. Los propósitos de la

fusión de imágenes son: a) obtener una imagen con una mejor percepción visual y espacial de las estructuras que componen las imágenes por separado, b) minimizar la información redundante c) maximizar la información la relevante.

Para lograr fusionar las imágenes provenientes de varias tecnologías, se han desarrollado una gran cantidad de metodologías que se ajustan a alguna aplicación o problemática en particular. Las aplicaciones que se han valido de estas metodologías son diversas:

- *Geo-ciencia*. El objetivo es la detección, clasificación y seguimiento de fenómenos terrestres. En este tipo de aplicaciones, los sensores ofrecen imágenes de baja y alta resolución que por lo general incluyen dos imágenes: una pancromática de alta resolución espacial (PAN) y una multi-espectral con baja resolución espacial (MS). El fusionar estos dos tipos de imágenes permite que los investigadores obtengan imágenes donde se aprecia una gran extensión de tierra con muchos detalles [Yang, 2012].
- *Diagnóstico médico*. Se tienen diferentes tipos de imágenes como resultado de la aplicación de ciertas técnicas no invasivas para detectar tumores, tejidos blandos, tejidos duros y diversas patologías. Resonancia magnética (MRI), Tomografía Computarizada (TM), Tomografía por emisión de positrones (PET), y ultra sonido son algunas de las imágenes utilizadas en este campo [Li, 2014]. La fusión de este tipo de imágenes permite a los médicos realizar un diagnóstico de mayor calidad y con mayor asertividad.
- *Imágenes multi-foco*. Una misma escena capturada con diferentes sensores presentan zonas en la que aparece el primer plan bien definido en una, pero en la otra se presenta como un segundo plano [Li, 2013].

En este trabajo de investigación, se abordará la fusión de imágenes IR y visuales en escalas de grises. Las imágenes infra-rojas (IR), son ampliamente utilizadas, debido a que, cómo se sabe, cualquier objeto que tenga una temperatura por encima del cero absoluto emite una radiación en la zona infrarroja. Esta radiación es captada por sensores diseñados para recibir estas longitudes de onda que son no perceptibles por el ojo humano y las presenta en un formato que podemos

entender. Las imágenes IR son una herramienta muy versátil que es aplicada en la resolución de innumerables problemas. En aplicaciones petroquímicas es utilizada para la detección de pérdida de aislamiento en los procesos de refinería, inspección de soldaduras, inspección de tuberías, evaluación de la eficiencia y calidad de intercambiadores de calor, entre otros. En aplicaciones aeroespaciales, son utilizadas para el diagnóstico de llantas y frenos, diagnóstico de sistemas de descongelación en alas, detección de desgaste de materiales compuestos, inspección de líneas eléctricas de alta tensión, inspección de tuberías y áreas con grietas por corrosión. En sistemas de seguridad, se han reportado aplicaciones para detectar portación de armas ocultas [Xue, 2003], identificación y rastreo de objetivos, detección de minas y autenticación de personal [Mayet, 1996], reconocimiento de rostros [Piella, 2003].

El proceso de fusionar imágenes puede ser clasificada en tres niveles de acuerdo al tipo de procesamiento que se efectúe [Pohl, 1998]:

- A nivel de pixel: se genera una imagen fusionada en la que la información asociada a cada pixel se determina a partir de los pixeles de las imágenes originales.
- A nivel de características: se extraen características que requieran fusionar con mayor interés y después se identifican tamaños, formas, contrastes y texturas previo al paso de fusión.
- A nivel de decisión: la fusión se realiza procesando las imágenes por separado determinando no solo algún objeto de interés, sino también algunos otros parámetros como contraste, ruido, vecindad entre pixeles.

Al escoger cual es el esquema más adecuado de procesamiento dependerá en gran medida de factores como el tipo de datos de entrada, la aplicación y las herramientas con las que se dispone [Piella, 2003]. En particular, las imágenes IR, ofrecen muy baja resolución espacial en comparación con las imágenes visuales, pero ofrecen detalles que no son perceptibles a la vista humana y que, en conjunto permiten entender una escena. Para llevar a cabo la fusión de estos dos tipos de imágenes se han desarrollado innumerables métodos. Por ejemplo,

se han desarrollado métodos de procesamiento multi-resolución, en el que las imágenes de entrada son descompuestas a través de diversas transformaciones wavelet y curvelet, de manera tal, que se van va extrayendo la información sobresaliente para, posteriormente, fusionarlas a través de reglas de fusión [Pohl, 1998], [Pinoli, 1997].

Otros métodos de descomposición multi-escala propuestos en [Toet, 1989], [Toet, 1990], [Toet, 1992], [Toet, 1989] hacen uso de filtros pasa baja que permiten preservar el contraste de las imágenes de entrada. Multi-resolución con filtros morfológicos. Independientemente del método de fusión que se desee emplear, es requisito que las imágenes de entrada estén alineadas para garantizar que la información de cada toma corresponde con las estructuras físicas de la realidad. Una de las grandes aportaciones en el procesamiento de imágenes, ha sido el desarrollado en [Pinoli, 1997], en el que se propone un conjunto de operadores que permiten manipular de forma no lineal los tonos de una imagen formada en escala de grises.

En este estudio, se presenta un método de fusión en el que se incluyen filtros morfológicos en el modelo LIP utilizando un modelo de procesamiento de imágenes conocido como *LIP MODEL*.

## **2. Métodos**

### **Modelo de Procesamiento Logarítmico de Imágenes**

Una imagen está formada por luz reflejada que puede ser representada de manera vectorial. Si se intenta adicionar dos valores de intensidad o bien, multiplicar un valor por un factor constante, el resultado pudiera no estar dentro del intervalo en el que se acotan las imágenes de entrada [Mayet, 1996]. Bajo esta premisa, el modelo LIP, define un conjunto de operadores que permiten manipular las intensidades de grises de una o más imágenes acotando su resultado al espacio de trabajo de cada una de ellas. Este modelo fue introducido por [Pinoli, 1992] en 1980 y se desarrolló para el procesamiento de imágenes hasta 1997 y ha sido utilizado en aplicaciones como remoción de fondos de una imagen, corrección de iluminación, interpolación de imágenes, reconstrucción 3D, restauración de

imágenes, estimación del contraste, segmentación de imágenes, descomposición multi-escala y compresión de imágenes [Mayet, 1996]. Pese a la gran variedad de aplicaciones de este modelo, no se han reportado su aplicación para el problema de fusión. De acuerdo con el modelo LIP, la adición de dos imágenes está definido a través de un operador a un intervalo definido para la cantidad máxima de tonos posibles  $(0, M]$ . En particular, la adición de una imagen  $f(x, y)$  con otra  $g(x, y)$  queda definida como en la ecuación 1.

$$f(x, y) \dagger g(x, y) = f(x, y) + g(x, y) - (f(x, y) * g(x, y) / M) \quad (1)$$

Con  $M = 2^n, \forall x, y \in \mathbb{Z} \rightarrow (0, M]$  donde  $n = 8$ .

Este operador  $\dagger$  llamado adición LIP, permite la adición cerrada en el intervalo  $(0, M]$ . También es posible amplificar el contenido de una imagen a través del operador de producto escalar, cuya definición es mostrada en la ecuación 2.

$$\alpha * f(x, y) = M - M \left( 1 - \frac{f(x, y)}{M} \right)^\alpha \quad (2)$$

Con  $M = 2^n$ . En general, cuando  $\alpha > 1$ , se eleva a tonos claros el contenido de la imagen. Cuando  $\alpha < 1$ , se oscurece la imagen. En  $\alpha = 1$ , el contenido de la imagen queda sin cambios. En la siguiente figura se muestra como son afectadas las tonalidades de unas imágenes para cierto rango de valores de  $\alpha$ .

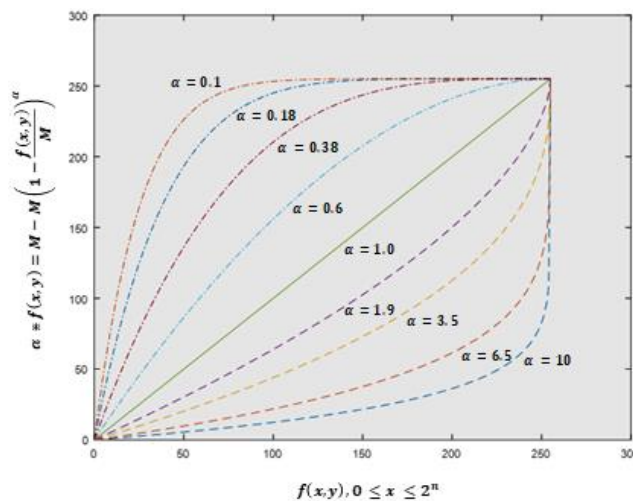


Figura 2 Modificación del tono por el producto escalar LIP.

Una manera de conocer el valor de alfa [Michoud, 1997] y que permite maximizar la dinámica de la imagen está dada por ecuación 13.

$$\alpha(f) = \frac{\ln\left(\frac{\ln\left(\frac{M - f_{max}}{M}\right)}{\ln\left(\frac{M - f_{min}}{M}\right)}\right)}{\ln\left(\frac{M - f_{min}}{M - f_{max}}\right)} \quad (3)$$

Donde  $f_{min}$  y  $f_{max}$  representan respectivamente los niveles de gris mínimo y máximo de una imagen.

### Morfología Matemática

La Morfología Matemática descansa en la teoría de conjuntos desarrollada por el matemático ruso Hermann Minkoski y el alemán Hugo Hadwiger en los primeros años del siglo XX. La reformulación de ésta teoría por los franceses Jean Serra y Georges Matheron dieron como consecuencia una teoría para el análisis y procesamiento de estructuras geométricas cuya aplicación más popular ha sido en el análisis de imágenes. En general, la morfología matemática permite realizar transformaciones no lineales sobre estructuras reticulares en espacios bidimensionales mediante el uso de un conjunto de forma conocida llamado elemento estructural cuyo tamaño y forma se escoge, *a priori*, de acuerdo con la morfología del conjunto sobre el que va a operar y en función de lo que se desea obtener. Las operaciones básicas de la teoría morfológica son la erosión y la dilatación y son denotadas en las ecuaciones 4 y 5.

$$\delta_{\lambda\beta}(f)(x) = (f \oplus \lambda\hat{\beta})(x) = \vee \{f(y) : y \in \lambda\hat{\beta}_x\} \quad (4)$$

$$\varepsilon_{\lambda\beta}(f)(x) = (f \ominus \lambda\hat{\beta})(x) = \wedge \{f(y) : y \in \lambda\hat{\beta}_x\} \quad (5)$$

Estas operaciones no admiten inversa, por tanto, no se puede determinar el origen  $x$  desde las imágenes  $\delta_{\lambda\beta}(f)(x)$  o  $\varepsilon_{\lambda\beta}(f)(x)$ . Sin embargo, dada la dualidad de estas transformaciones, es posible aproximarse al elemento original realizando una dilatación dada la erosión o bien, erosionar dada una dilatación. Estas operaciones son llamadas Apertura y Cerradura morfológica y son denotados en ecuaciones 6 y 7.

$$\gamma_{\lambda\beta}(x) = (x \ominus \lambda\beta) \oplus \lambda\beta = \delta_{\lambda\tilde{\beta}}(\varepsilon_{\lambda\beta}(x)) \quad (6)$$

$$\varphi_{\lambda\beta}(x) = (x \oplus \lambda\beta) \ominus \lambda\beta = \varepsilon_{\lambda\tilde{\beta}}(\delta_{\lambda\beta}(x)) \quad (7)$$

La apertura morfológica es de utilidad para eliminar detalles luminosos o claros en relación con elemento estructural quedando el resto de la imagen relativamente sin modificaciones. El cierre, en cambio, elimina detalles oscuros en relación con el elemento estructural. Con estas definiciones formamos una transformación que permite descubrir la información que la apertura y el cierre remueven: la transformada Top-Hat, que se define como la diferencia entre la imagen original y la apertura, llamada Top-Hat sobre blancos y, la diferencia entre el cierre y la imagen original, llamado Top-Hat sobre negros. Su notación es mostrada en las ecuaciones 8 y 9.

$$ThW_{\lambda\beta}(X) = X - \gamma_{\lambda\beta}(x) \quad (8)$$

$$ThB_{\lambda\beta}(X) = \varphi_{\lambda\beta}(x) - X \quad (9)$$

La principal desventaja de la aplicación de las operaciones morfológicas es la distorsión producida por el elemento estructural sobre las estructuras originales de la imagen. A diferencia de las transformadas morfológicas, en las que, a través de aperturas y cerraduras se eliminan ciertas regiones para resaltar algunos otros detalles, las transformadas por reconstrucción corresponden a un grupo de transformadas que permiten preservar la información de origen permitiendo eliminar ciertas zonas que previamente han sido marcadas de acuerdo algún propósito específico. En general, las transformaciones por reconstrucción se forman a partir de una imagen de referencia  $X$  y una imagen  $Y$  llamada “marcador” que crece al interior de la primera. Para construir estas transformaciones se hace uso del concepto de erosión y dilatación geodésica.

La dilatación geodésica de tamaño 1, es denotada en ecuación 10.

$$\delta_X^1(Y) = X \cap \delta_B(Y) \quad (10)$$

Cuando esta dilatación es iterada hasta la estabilidad, obtenemos la transformación por reconstrucción dada por ecuación 11.

$$\rho_X(Y) = R(X, Y) = \lim_{n \rightarrow \infty} \delta_X^n(Y) = \delta_X^1 \delta_X^1 \dots \delta_X^1(Y) \quad (11)$$



Dado que las operaciones morfológicas son duales, entonces, para el caso de la erosión geodésica, tenemos la notación para tamaño 1 en la ecuación 12.

$$\varepsilon_X^1(Y) = X \cap \varepsilon_\beta(Y) \quad (12)$$

Iterando hasta la estabilidad, obtenemos la transformación dual por reconstrucción dada por la ecuación 13.

$$\rho_X^*(Y) = R^*(X, Y) = \lim_{n \rightarrow \infty} \varepsilon_X^n(Y) = \varepsilon_X^1 \varepsilon_X^1 \dots \varepsilon_X^1(Y) \quad (13)$$

Cuando el conjunto marcador Y es igual a la erosión X, obtenemos la apertura y cierre por reconstrucción, ecuaciones 14 y 15.

$$\tilde{Y}_{\lambda B}(X) = \lim_{n \rightarrow \infty} \delta_X^n(\varepsilon_{\lambda B}(X)) = R(X, \varepsilon_{\lambda B}(X)) \quad (14)$$

$$\tilde{\varphi}_{\lambda B}(X) = \lim_{n \rightarrow \infty} \varepsilon_X^n(\delta_{\lambda B}(X)) = R(X, \delta_{\lambda B}(X)) \quad (15)$$

De la misma forma en que fueron establecidas las operaciones Top-Hat morfológica sobre blancos y Top-Hat sobre oscuros, para el caso de la reconstrucción tenemos las ecuaciones 16 y 17.

$$ThW(X) = X - \tilde{Y}_{\lambda B}(X) \quad (16)$$

$$ThB(X) = \tilde{\varphi}_{\lambda B}(X) - X \quad (17)$$

Con los elementos mostrados, a continuación, describiremos como es utilizada esta última transformada para poder intercambiar la información de dos imágenes.

### Modelo LIP como Operador Primitivo de Fusión

Una de las premisas para llevar a cabo la fusión de imágenes, es encontrar un operador que, a nivel de pixel, adicione sus valores y cuyo resultado este acotado a un rango que, para efectos de imágenes codificadas en 8 bits, sea 255. Dado que el operador de adición LIP cumple con esta premisa, se verá a continuación que es lo que ocurre cuando aplicamos este operador sobre 2 imágenes de entrada. En la figura 3, se presenta una imagen visual y una de tipo IR de la misma escena. El resultado de aplicar el operador de adición LIP sobre estas entradas se puede apreciar en figura 4.



Figura 3 Imágenes de entrada.



Figura 4 Resultado de la adición LIP al conjunto de imágenes de entrada.

Como se aprecia en la figura 5, si bien la imagen resultante muestra el contenido de las dos de entrada, aún no es posible apreciar con detalle la información complementaria que la imagen IR aporta. Con la intención de amplificar algunos detalles, a través de la ecuación 3, determinamos el valor de  $\alpha = 0.8956$  para la imagen resultante y aplicamos el producto LIP por este escalar de la ecuación 2. El resultado se muestra en la figura 5.

Observe cómo, efectivamente algunas zonas se ven más claras en comparación con el resultado obtenido en la figura 4, sin embargo, ésta imagen y al igual en el primer resultado, no es posible apreciar la información complementaria que aporta la imagen IR a toda la escena. Es necesario, entonces, encontrar un mecanismo

que, a partir de la adición y multiplicación LIP sobre imágenes de entrada, incorpore los detalles complementarios para terminar el proceso de fusión.



Figura 5 Producto LIP por escalar de la adición de dos imágenes de entrada.

### Modelo Fusión Morfológica usando LIP Model (MLiF)

Como se expuso en la sección anterior, el operador de adición LIP en conjunto con el operador de multiplicación LIP por un escalar, nos ofrece una primera aproximación a un modelo de fusión en el que la información contenida en la imagen fusionada represente en mayor medida la información de las imágenes de entrada.

Para incluir en el resultado final de la fusión la información complementaria de la imagen IR, se propone entonces la extracción de características de la siguiente manera. Sea  $F(x, y)$  y  $G(x, y) \forall (x, y) \in \mathbb{N}$  dos imágenes definidas en  $2^n$  niveles de grises con  $n = 8$ . Aplicando las ecuaciones 16, 17 a las imágenes de entrada, obtenemos las ecuaciones 18, 19, 20 y 21.

$$thW(F) = F - \tilde{\gamma}_{\lambda\beta}(F) \quad (18)$$

$$thB(F) = \tilde{\varphi}_{\lambda\beta}(F) - F \quad (19)$$

$$thW(G) = F - \tilde{\gamma}_{\lambda\beta}(G) \quad (20)$$

$$thB(G) = \tilde{\varphi}_{\lambda\beta}(G) - G \quad (21)$$

Donde  $\beta$  es un elemento estructural plano de  $3 \times 3$  y  $\lambda \in \mathbb{N}$ . Con estas transformaciones por reconstrucción, obtenemos dos nuevos conjuntos que los denotamos en las ecuaciones 22 y 23.

$$Dn = V(thW(F), thB(F)) \tag{22}$$

$$Dm = V(thW(G), thB(G)) \tag{23}$$

Y representan las zonas más claras que sobresalieron en función del tamaño del elemento estructural. A continuación, se determinan las estructuras brillantes de ambas imágenes de entrada a través de una función de composición denotada en las ecuaciones 24 y 25.

$$C_{m,n} = Dn \circ Dm = \begin{cases} 2^n & \text{si } Dn > Dm \\ 0 & \text{en otro caso} \end{cases} \tag{24}$$

$$\bar{C}_{m,n} = C_{m,n} \tag{25}$$

Esta función de composición permite unir las zonas brillantes de ambas imágenes de entrada que fueron obtenidas a través de las transformaciones morfológicas Top – Hat. Esta función de composición es filtrada para homogenizar las zonas brillantes. Finalmente, la imagen fusionada se obtiene mediante la ecuación 26.

$$Fm(F, G) = \begin{cases} \alpha * (F + G) + Dn & \text{si } \bar{C}_{m,n} = 2^n \\ \alpha * (F + G) & \end{cases} \tag{26}$$

En la figura 6, se presenta de manera gráfica una representación del método que acabamos de describir y, en la figura 7, se muestra el pseudocódigo de este.

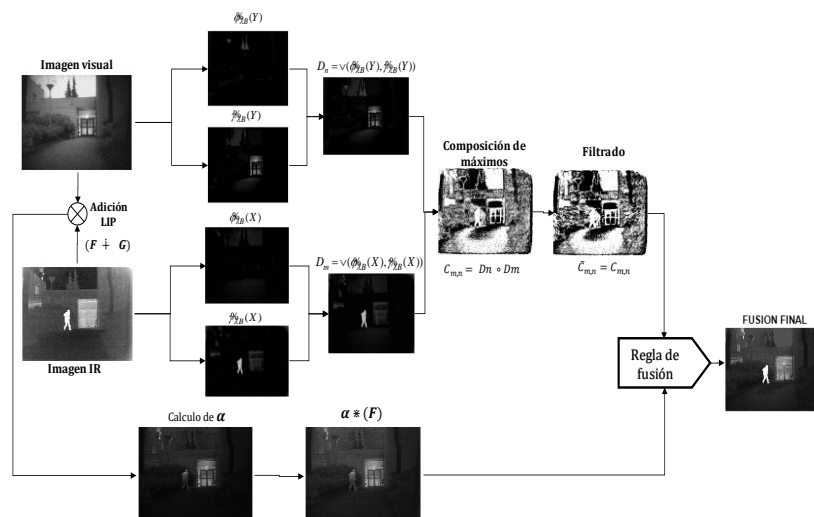


Figura 6 Diagrama de bloques del modelo MLiF.

---

**Algoritmo propuesto**

---

**inputs** : Imagen visual ( $I_{vis}$ ), Imagen IR ( $I_{ir}$ )  
**output** : imagen fusionada ( $ImFus$ )

$lipAdd \leftarrow$  adición LIP( $I_{vis}, I_{ir}$ )  
 $alphaLIP \leftarrow$  cálculo del parámetro  $\alpha$   
 $preFusLIP \leftarrow$  producto escalar LIP( $lipAdd, alphaLIP$ )

$I_{vis_{THW}} \leftarrow$  TopHat( $I_{vis}$ )  
 $I_{vis_{THB}} \leftarrow$  BottonHat( $I_{vis}$ )  
 $D_n \leftarrow$  max( $I_{vis_{THW}}, I_{vis_{THB}}$ )

$I_{ir_{THW}} \leftarrow$  TopHat( $I_{ir}$ )  
 $I_{ir_{THB}} \leftarrow$  BottonHat( $I_{ir}$ )  
 $D_m \leftarrow$  max( $I_{ir_{THW}}, I_{ir_{THB}}$ )

Para todo pixel ( $i, j$ ) de  $D_n$  hacer  
    Si  $D_n > D_m$  entonces  
         $R_n(i, j) \leftarrow 2^n$   
    sino  
         $R_n(i, j) \leftarrow 0$   
fin para todo

$medianR_n \leftarrow$  filtro de mediana( $R_n$ )

Para todo pixel( $i, j$ ) de  $R_n$  hacer  
    Si  $medianR_n(i, j) == 2^n$   
         $ImFus(i, j) \leftarrow preFus(i, j) + D_n(i, j)$   
    sino  
         $ImFus(i, j) \leftarrow preFus(i, j)$   
fin para todo

---

Figura 7 Pseudocódigo del método propuesto.

### 3. Resultados

#### Conjuntos de Datos Experimentales

Para comprobar la eficiencia del método propuesto, se utilizó un conjunto de 20 pares de imágenes como entrada, previamente registradas y que fueron tomadas de TNO Humans Factors<sup>1</sup>, que contienen tomas militares en diferentes escenarios y que fueron tomadas con diferentes tipos de cámaras. En la figura 8 se muestran los 20 pares de imágenes de entrada que se utilizaron para validar nuestro modelo de fusión; a la izquierda de cada par tenemos la imagen IR y a la derecha se muestra la misma escena, pero en imagen visual. La eficiencia de un método puede valorarse desde dos puntos de vista: a partir de la aplicación que lo requiere a través del observador que hará uso de la fusión y, a través de métricas que cuantifiquen el grado en que la información de entrada fue intercambiada para obtener un resultado final. En cuanto a la primera, pudiera resultar un poco complicado determinar qué método de fusión fue el mejor ya que, visualmente la diferencia entre un resultado y otro pudiera no ser claramente perceptible.

Quizá para alguna aplicación en particular, para un observador, un determinado método es mejor frente a otro. Bajo esta idea, la valoración de eficiencia se dará en términos cualitativos y será subjetiva ya que estará sujeta a la apreciación del observador y estará en función de lo que pudiera buscar o esperar.

<sup>1</sup> El conjunto de datos de entrada está disponible en [http://figshare.com/articles/TNO\\_Image\\_Fusion\\_Dataset/1008029](http://figshare.com/articles/TNO_Image_Fusion_Dataset/1008029)

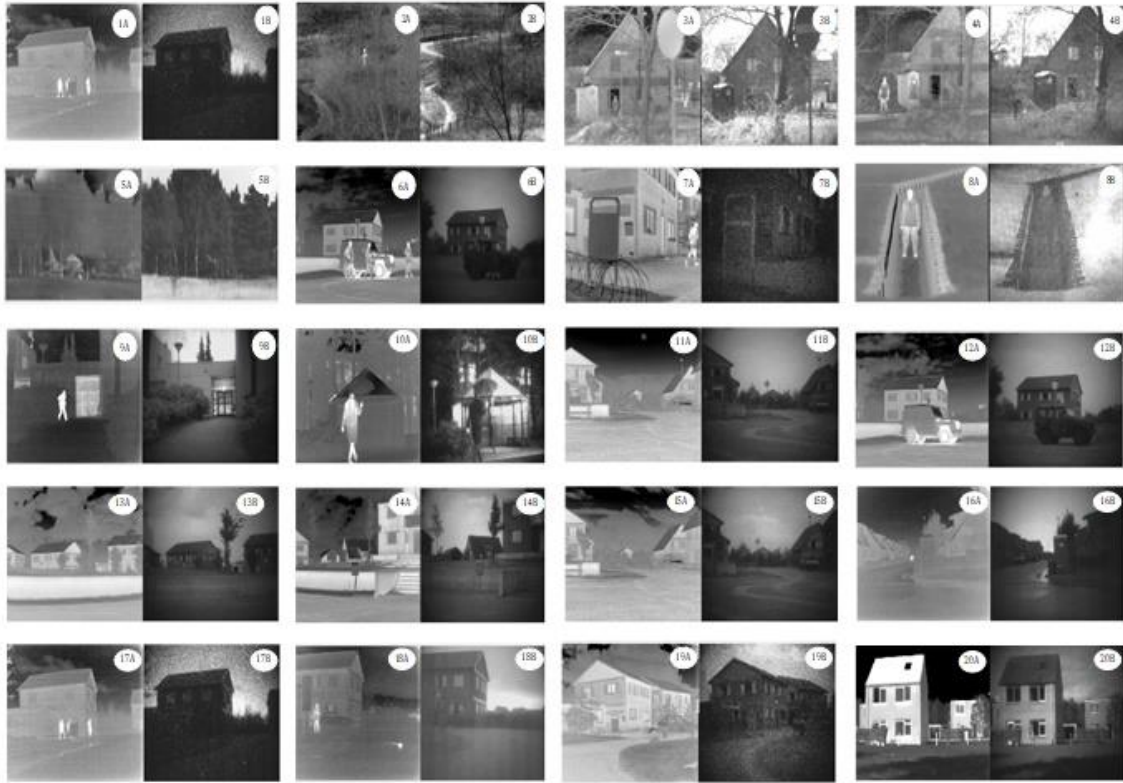


Figura 8 Conjunto de datos experimentales.

Por otro lado, para valorar de manera cuantitativa la posible ventaja de un modelo sobre otro, se utilizaron algunos indicadores que comparan la similitud, los contornos, el contenido de información, el contraste, entre otras, de las imágenes de entrada con la imagen fusionada. Esto no quiere decir, a priori, que un indicador sea mejor que otro, simplemente miden cosas diferentes. Para este estudio, fueron utilizados cuatro indicadores, i.e., Correlación Cruzada (Corr2D), Error Cuadrático Medio (RMSE), el Índice de la Calidad de la Fusión (IoQ) y el Índice de Similitud Estructural (SSIM). La definición de estos cuatro indicadores se presenta a continuación:

- Correlación Cruzada (Corr2D): Es una medida estadística de similitud entre dos variables aleatorias. Para determinar la correlación de la imagen fusionada contra las originales, decimos que nuestro valor de correlación es el producto de la correlación de las originales contra la fusionada. Es decir, si  $r_{AF} = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(F_{mn} - \bar{F})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2)(\sum_m \sum_n (F_{mn} - \bar{F})^2)}}$  es la correlación de la imagen A

(una del conjunto de entrada) con la fusionada (F) y  $r_{BF} = \frac{\sum_m \sum_n (B_{mn} - B)(F_{mn} - F)}{\sqrt{(\sum_m \sum_n (B_{mn} - B)^2)(\sum_m \sum_n (F_{mn} - F)^2)}}$  es la correlación de la imagen B (la segunda del conjunto de entrada) con la fusionada, entonces nuestro indicador de fusión será:  $r_{AF} * r_{BF}$ .

- Error Cuadrático Medio (RMSE), permite cuantifica el error medio entre cada una de las imágenes de entrada contra la imagen fusionada.
- El Índice de Calidad de la Fusión (IoQ), es una métrica que combina el coeficiente de correlación, la distorsión lumínica y la distorsión del contraste e indica el grado con el que se integraron las imágenes de entrada en la fusionada. El máximo de este valor es 1 y es alcanzado cuando la totalidad de las entradas están integradas en la fusionada.
- El Índice de Similitud Estructural (SSIM) es un indicador que permite valorar calidad de la fusión con base a la degradación de estructuras de información contenidas en la imagen fusionada contra otra de referencia comparando patrones locales de intensidad que han sido normalizados en luminancia y contraste.
- Para contrastar los resultados y poder valorar si el método propuesto es eficiente, se utilizó el mismo conjunto de imágenes con dos algoritmos de fusión que también hacen uso de transformaciones morfológicas para lograr la fusión y que han sido reportados en [Toet, 1989] [Mukhopadhyay, 2001]. Todos los experimentos se realizaron en una laptop con Intel core i5 a 3.3 GHz y 8 GB de memoria y los algoritmos fueron codificados de Matlab.

## Resultados Obtenidos

Los resultados de nuestra experimentación son mostrados en la figura 9. En esta gráfica se presenta el comparativo de la fusión realizada por 3 métodos diferentes. En la imagen de la izquierda, se aprecia el resultado del método propuesto, al centro se presenta el método reportado [Toet, 1989] y a la derecha la fusión realizada por el método reportado en [Mukhopadhyay, 2001]. Visualmente se puede apreciar que los tres métodos de fusión cumplen el objetivo de integrar las imágenes de entrada.

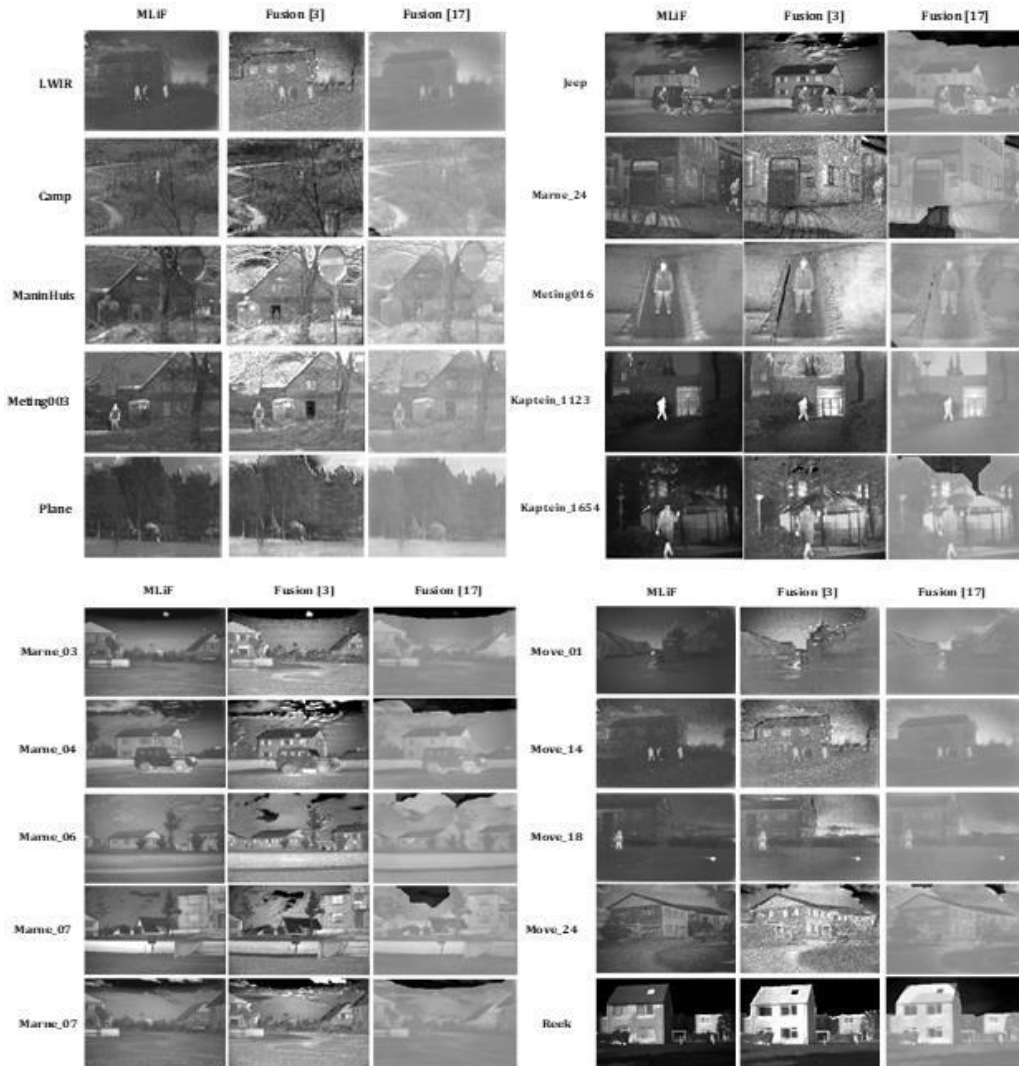


Figura 9 Resultados de la fusión de imágenes de entrada y su comparación con dos métodos de fusión.

Sin embargo, también es posible observar que, en algunos pares fusionados por los métodos reportados en [Toet, 1989], [Mukhopadhyay, 2001] no fusiona de forma eficiente. Se aprecian zonas que no fueron fusionadas o bien, aparecieron algunos contornos que ninguna de las originales tiene y esto es debido a la cantidad de escalamientos del elemento estructural que fueron usados para llevar a cabo la fusión. La ventaja del modelo propuesto es que el elemento estructural no es escalado, su tamaño permanece fijo y solo es usado para extraer la información complementaria de la toma IR e integrarla a la adición LIP previamente realizada.



En la tabla 1 se presentan los resultados obtenidos de aplicar el modelo de fusión propuesto y la fusión realizada con los dos modelos mencionados. Las celdas marcadas indican el mejor método de fusión de cada una de las imágenes de entrada. Se puede apreciar, que el indicador IoQ resultó el mejor método en la gran mayoría. Sin embargo, para el indicador SSIM el método propuesto fue el mejor en 13 de los 20 pares de entrada. En cuanto a la correlación cruzada 9 de los 20 pares de entrada y para el error cuadrático medio 10 de los 20 pares el método propuesto resulto mejor.

Tabla 1 Tabla comparativa de modelos de fusión con diferentes indicadores.

	IoQu			PEARSON			RMSE			SSIM		
	MLiF	Fusion[3]	Fusion[17]	MLiF	Fusion[3]	Fusion[17]	MLiF	Fusion[3]	Fusion[17]	MLiF	Fusion[3]	Fusion[17]
Lwir	-0.0041	0.1408	0.0972	-0.0041	0.0075	-0.0849	8.2542	10.5531	22.5769	0.3873	0.1611	0.2648
Camp	0.0472	0.2148	0.0608	0.0472	0.1137	-0.0577	0.7724	0.9917	55.9463	0.3760	0.3550	0.2407
MainHuis	0.1122	0.2513	0.1171	0.1122	0.1816	-0.0129	25.8455	7.4859	9.8593	0.3140	0.3654	0.4332
Meting003	0.0204	0.2339	0.1136	0.0204	0.0105	-0.0546	17.8241	5.6631	8.4309	0.4068	0.4194	0.5034
Plane	0.2032	0.2194	0.1017	0.2032	0.1998	-0.0064	3.4090	2.0385	19.3434	0.5934	0.6223	0.5894
Jeep	-0.0727	0.2218	0.1325	-0.0727	-0.0307	-0.1971	6.3579	9.3890	35.1170	0.4139	0.3039	0.3473
Marne_24	0.1563	0.1932	0.1641	0.1563	0.2156	0.0451	8.1696	7.7910	15.4476	0.2873	0.1441	0.2261
Meting016	0.1879	0.1735	0.0748	0.1879	0.0483	0.0671	2.9577	2.4209	3.7885	0.5958	0.4812	0.5388
Kaptein_1123	0.3097	0.2960	0.0829	0.3097	0.3129	0.2435	8.5048	4.5340	50.7185	0.3869	0.4018	0.3377
Kaptein_1654	0.2459	0.1383	0.0614	0.2459	0.2171	0.1065	16.8994	4.4315	30.0477	0.2775	0.3146	0.3455
Marne_03	-0.0646	0.1916	0.1236	-0.0646	-0.1132	-0.1467	14.2445	18.0274	25.3263	0.4042	0.2421	0.3723
Marne_04	-0.0691	0.2293	0.1303	-0.0691	-0.0159	-0.1776	8.3654	15.6860	24.6565	0.4104	0.2612	0.3759
Marne_06	0.0516	0.2025	0.1410	0.0516	-0.0664	0.0128	10.9190	6.2471	15.6696	0.4471	0.3662	0.4611
Marne_07	-0.0160	0.2594	0.1385	-0.0160	0.0340	-0.1257	10.9692	12.7138	28.6898	0.4341	0.3228	0.3845
Marne_11	0.0956	0.2024	0.1377	0.0956	-0.0324	-0.0954	10.8614	15.4343	17.8831	0.4354	0.2619	0.4046
Move_01	-0.0848	0.2054	0.0942	-0.0848	0.0357	-0.0552	7.9024	9.2538	17.6585	0.5392	0.3592	0.4753
Move_14	-0.0584	0.1350	0.0911	-0.0584	0.0724	-0.1953	10.3856	15.6774	24.6844	0.3756	0.1427	0.2516
Move_18	0.0823	0.3532	0.0904	0.0823	-0.0381	0.0144	5.1085	1.9798	5.9776	0.5601	0.5142	0.5580
Move_24	0.0388	0.1432	0.1295	0.0388	0.1314	0.0452	11.9498	16.8183	15.1465	0.3192	0.1206	0.2626
Reek	0.3668	0.2306	0.1199	0.3668	0.3519	0.0596	18.5196	4.0143	25.8226	0.2614	0.3782	0.2314

#### 4. Discusión

Por los resultados obtenidos, se concluye qué si bien el uso de indicadores permite cuantificar el grado de fusión de un método sobre otro, el resultado también dependerá de qué es lo que mida un indicador en particular y cómo lo hace. Mientras un indicador pudiera concluir que un método de fusión resultó

mejor que otro, para otro indicador que mida algo distinto, tal vez otro método de fusión resulte más eficiente. Por tanto, no habrá un indicador a través del cual podamos concluir que un método es mejor que otro y por esta razón los métodos de fusión deberán de estar en función de la aplicación para la que se quiera y de lo que el observador este buscando y/o esperando encontrar.

## 5. Conclusiones

En este estudio se mostró el uso de los operadores básicos del modelo LIP para fusionar imágenes visuales e imágenes IR. Obtenemos una imagen “pre-fusionada”, tomando como base la adición LIP de las imágenes de entrada. Posteriormente, aplicando el operador de producto LIP por un escalar  $\alpha$  de esta “pre-fusión” logramos acentuar las zonas más brillantes. Sin embargo, también se mencionó que la sola aplicación de éstas operación no es suficiente para lograr una fusión de calidad, ya que estos operadores operan a nivel de pixel y no permite extraer la información complementaria que ofrecen las tomas IR de una escena en particular. Se usaron las transformaciones por reconstrucción Top Hat para extraer dicha información y desarrollamos un modelo de fusión que permite unir el resultado de este procesamiento al resultado que ofrecen los operadores de adición y multiplicación por un escalar LIP.

Dado que el método propuesto hace uso de transformaciones morfológicas, hemos comparado nuestra propuesta contra dos métodos que se han reportado en [Toet, 1989] [Mukhopadhyay, 2001] que también hace uso de estas transformaciones y que, además, hacen uso del procesamiento multi-escala, haciendo que el proceso de fusión resulte más complejo. Los resultados revelan que, en términos visuales, nuestra propuesta resultó mejor ya que solo empleamos las transformaciones morfológicas para extraer la información complementaria y la agregamos a la pre-fusión realizada a través del modelo LIP, evitando de esta manera, la aparición de contornos no deseados como fue el caso con los otros dos modelos.

Se ha mostrado, cómo cuatro indicadores que actúan de diferente manera reportan que, de acuerdo a como realiza la medición, un método de fusión resulta

efectivo para ciertas imágenes mientras que para otras el mismo indicador pudiera ser diferente. Esto nos habla de la complejidad del proceso de fusión de señales y que su efectividad estará acompañada por dos factores: de la aplicación para la cual será utilizada y de lo que un observador espera encontrar.

## **6. Bibliografía y Referencias**

- [1] Gyaourova A., G. Bebis, I., Fusion of infrared and visible images for face recognition. *Computer Vision–ECCV*, 2004.
- [2] Lewis J. J., R.J. O’Callaghan, S.G. Nikolov, D.R. Bull. Pixel–and region-based image fusion with complex wavelets. *Information Fusion*, Volume 8, Issue 2, pp. 119–130, 2007.
- [3] Li S., X. Kang, J. Hu, B. Yang. Image matting for fusion of multifocus in dynamic scenes. *Information Fusion* 14, pp. 147–162, 2013.
- [4] Li T., Y. Wang, C. Chang, N. Hu, Y. Zheng, Color-Appearance based fusion of gray and pseudo-color images for medical applications. *Information Fusion* 19, pp. 103–114, 2014.
- [5] Ma J., C. Chen, C. Li, J. Huang, Infrared visible image fusion via gradient transfer and total variation minimization. *Information Fusion* 31, 2016.
- [6] Matsopoulos G. K., S. Marshall, Application of morphological pyramids: fusion of MR and CT phantoms. *Journal of Visual Communications and Image Representation* 6 (2), PP. 167 –207, 1995.
- [7] Mayet F., J.C. Pinoli, M. Jourlin, Physical Justification and applications of the LIP Model for the Processing of transmitted light images. *Traitement du Signal*, Volume 13 No. 3, 1996.
- [8] Michoud P., J.C. Pinoli, M. Jourlin, Les applications industrielle et biomédicales du modèle LIP, *Seizème Colloque Gretsi*, pp. 15–19, Septembre 1997.
- [9] Mukhopadhyay S., B. Chanda, Fusion of 2D grayscale images using multiscale morphology. *Pattern Recognition* 34, pp. 1939–1949, 2001.
- [10] Omar A., T. Stathaki. Image Fusion: an overview. *Fifth International Conference on Intelligent System, modelling and Simulation*, 2014.

- [11] Piella G., A general framework for multiresolution image fusion: from pixels to regions. *Information Fusion* 4, pp. 259-280, 2003.
- [12] Pinoli J. C., The Logarithmic Image Processing Model: Connection with Human Brightness Perception and Contrast Estimator. *Journal of Mathematical Imaging and Vision* 7, pp. 341–158, 1997.
- [13] Pohl C., J.L. Genderen, Multisensor image fusion in remote sensing. *International Journal of Remote Sensing*. 19 (5), pp. 823–854, 1998.
- [14] Singh R., A. Khare, Fusion of multimodal medical using Daubechies complex wavelet transform–A multiresoution approach. *Information Fusion* 19, pp. 49–60, 2014.
- [15] Toet A., Image Fusion by ratio of low–pass pyramind. *Pattern Recognition Letters* 9 (4), pp. 245–253, 1989.
- [16] Toet A., Hierarchical Image Fusion, *Machine Vision and Applications*, 1990.
- [17] Toet A., Multiscale contrast enhancement with applications to image fusion. *Optical Engineering* 31(5), pp. 1026-1031, may 1992.
- [18] Toet A., Merging thermal and visual images by a constrast pyramids. *Optical Engineering* 28(7), pp. 789 -792, July 1989.
- [19] Xue Z., R.S.Blum, Concealed Weapon Detection Using Color Image Fusion. Electrical and Computer Engineering Deparment. Lehigh University, 2003.
- [20] Yang S., M. Wang, L. Jiao. Fusion of multispectral and panchromatic images based on support value transform and adaptive principal component analysis. *Information Fusion* 13, pp. 177–184, 2012.