Published in Image Processing On Line on 2018–10–26.
Submitted on 2016–12–01, accepted on 2018–09–28.
ISSN 2105–1232 © 2018 IPOL & the authors CC–BY–NC–SA
This article is available online with supplementary materials,
software, datasets and online demo at
https://doi.org/10.5201/ipol.2018.196

Structural Similarity Metrics for Quality Image Fusion Assessment: Algorithms

Silvina Pistonesi[1,2], Jorge Martinez[1], Silvia María Ojeda[3], Ronny Vallejos[4]

[1] Departamento de Matemática, Universidad Nacional del Sur, Argentina
(lpistone@criba.edu.ar, martinez@uns.edu.ar)
[2] Facultad Regional Bahía Blanca, Universidad Tecnológica Nacional, Argentina
[3] Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, Argentina
(ojeda@mate.uncor.edu)
[4]Departamento de Matemática, Universidad Técnica Federico Santa María, Chile
(ronny.vallejos@usm.cl)

Abstract

The wide use of image fusion techniques in different fields such as medical diagnostics, digital camera vision, military and surveillance applications, among others, has motivated the development of various image quality fusion metrics, in order to evaluate them. In this paper, we study and implement the algorithms of non-reference image structural similarity based metrics for fusion assessment: Piella's metric, Cvejic's metric, Yang's metric, and Codispersion Fusion Quality metric. We conduct the comparative experiment of the selected image fusion metrics over four multiresolution image fusion algorithms, performed on different pairs of images used in different applications.

Source Code

The reviewed source code for this article and documentation for these algorithms are available from the web page of this article[1]. We used a MATLAB code in the implementation of the algorithms.

Keywords: image fusion; image quality metrics; structural similarity; non-reference quality measures

1 Introduction

Image fusion is the process of combining information available from two or more images of a scene into a single composite image which is more informative and suitable for both visual perception and computer processing. The goal of image fusion is to reduce uncertainty and minimize redundancy in the output, while maximizing the relevant information particular to an application of interest.

[1] https://doi.org/10.5201/ipol.2018.196

Quality assessment of different image fusion schemes is traditionally carried out by subjective evaluations. Even though this method [10] is reliable, it is expensive and too slow for real world applications. Therefore, several objective image quality measures that are consistent with human visual perception, have been proposed to predict image fusion quality automatically. They are classified into four groups according to their characteristics: information theory based metrics, image feature based metrics, human perception inspired fusion metrics, and image structural similarity based metrics [4]. In the context of measures based on image structural similarity, Piella's metric [7], Cvejic's metric [1], Yang's metric [13], and Codispersion Fusion Quality metric [8] have been developed. These fusion performance measures are based on the Universal Image Quality index (Q) [11]. They do not require a reference image, namely, they do not need the construction of some kind of ideal fused image for using it as a reference for comparing with the experimental fused results, in order to evaluate the performance of different fusion algorithms [1, 7]. Also, these objective quality assessment measures are based on a sliding window approach, taking into account local measures to estimate how well the relevant information from the input images is present in the fused images, without introducing distortions.

The objective of this work is to provide reproducible implementations of the above mentioned measures, along with a quantitative evaluation of their performance.

Table 1 summarizes the details of the adopted notation in the pseudo-codes. The rest of the paper is organized as follows: in Section 2 we introduce the image structural similarity quality indexes and give an overview of the structural similarity-based metrics for image fusion. Along with this Section, the pseudo-codes which implement the described algorithms are presented. Finally, Section 3 contains experimental results obtained by the implemented algorithms, including examples to compare their performance.

2 Image Quality Measures

This section provides a description of the image structural similarity measures and then presents the structural similarity based metrics for image fusion. Additionally, the pseudo codes, which implement the described measures are exposed.

In what follows, we let $x = \{x_{i,j} | 1 \leq i \leq M, 1 \leq j \leq N\}$ and $y = \{y_{i,j} | 1 \leq i \leq M, 1 \leq j \leq N\}$, with $M, N \in \mathbb{N}$, be the original and test image signals, respectively.

2.1 Image Structural Similarity Measures

Structural information is one of the important components that affect the visual quality of digital images. In this context, the structural information are those attributes that represent the structure of objects in the scene, independent from the average luminance and contrast. The image structural similarity measures presented below are based on the assumption that the human visual system is highly adapted to extract structural information from the field of vision.

2.1.1 Universal Image Quality Index (Q)

The Universal Image Quality index Q was introduced by Wang and Bovik [11] and is defined as follows

$$Q(x, y) = \frac{4 S_{xy} \bar{x} \bar{y}}{\left({S_x}^2 + {S_y}^2\right)\left(\bar{x}^2 + \bar{y}^2\right)}, \tag{1}$$

where $\overline{x}$ and $\overline{y}$ are the sample average values of images x and y, S_x, S_y and S_{xy} are the sample deviations and the sample covariance, respectively.

x, y	Matrix of size $M \times N$, whose elements are the pixel values in 8-bit grayscale.
f	Matrix of size $M \times N$, whose elements are the pixel values in 8-bit grayscale corresponding to the fused image.
$window$	Local window for statistics (by default, a Gaussian kernel of size 11×11 and standard deviation 1.5).
m_x	Matrix whose elements are the local means computed in each sliding window w of the x image.
s_x^2	Matrix whose elements are the local variances computed in each sliding window w of the x image.
s_{xy}	Matrix whose elements are the local covariances computed in each sliding window w between x and y images.
q_map	Matrix whose elements are the local Q index values computed in each sliding window w between two images.
$ssim\_map$	Matrix whose elements are the local $SSIM$ index values computed in each sliding window w between two images.
x_s, x_{s+h}	Matrix whose elements are the pixel values of the set $X' = \{s \in x : s + h \in x\}$ and of the set $\{s' = s + h : s \in X'\}$, respectively, in a specific direction h.
cq_map	Matrix whose elements are the local CQ index values computed in each sliding window w between two images, in a specific direction h.
λ	Matrix whose elements are the values of the local weights λ computed in each sliding window w of the source images.
c	Matrix whose elements are the values of the local weights c computed in each sliding window w of the source images.
C	Matrix whose elements are the values of the numerator of the local weight c computed in each sliding window w of the source images.
x', y' and f'	Matrix whose elements are the Euclidean norm of the horizontal and vertical gradient images.
sim	Matrix whose elements are the values of the local weighting factor sim, computed in each sliding window w of the source images and the fused image.
cq_maps	List where each element is a cq_map in a specific direction h.
cq_map_max	Matrix whose elements are the maximum local CQ index values according the different directions h, computed in each sliding window w between two images.
cq_max_map	Matrix whose elements are the local $CQ_{\max}$ index values, computed in each sliding window w between two images.
q_S_map, q_W_map, $q_{E_1}\_map$, q_C_map, q_Y_map, cq_M_map	Matrix whose elements are the local Q_S, Q_W, Q_{E_1}, Q_C, Q_Y and CQ_M index values, computed in each sliding window w of size w_size, respectively (quality maps of the fused image).

Table 1: Summary of the notation used in the pseudo-codes.

This index can be decomposed in the following three different factors:

$$Q(x,y) = s(x,y)\cdot l(x,y)\cdot c(x,y) = \frac{S_{xy}}{S_x S_y}\cdot\frac{2\bar{x}\bar{y}}{\bar{x}^2+\bar{y}^2}\cdot\frac{2S_xS_y}{{S_x}^2+{S_y}^2}. \tag{2}$$

The first factor in (2) measures the degree of linear correlation between x and y; the second measures the similarity between the luminance of x and y, and the third factor measures the similarity related to the contrast between the images.

Considering that image signals are generally non-stationary and that image distortions may be space-variant, it is more appropriate to measure the image quality index Q over local regions and then combine the different results into a single measure. Therefore, Wang and Bovik [11] propose to use a sliding window approach: starting from the top-left corner of the two images x, y, a sliding window of a fixed size block by block over the entire image until the bottom-right corner is reached. Finally, the overall image quality index Q is determined by averaging all local quality indexes $Q(x,y|w)$ computed in all sliding windows $w \in W$

$$Q = \sum_{w\in W}\frac{Q(x,y|w)}{|W|}, \tag{3}$$

with W the family of all windows and $|W|$ the cardinality of W.

Pseudo-code of the Q Index. A pseudo-code for the implementation of the Q index algorithm is presented in Algorithm 1.

Algorithm 1: Pseudo-code of the Q similarity quality index algorithm.

input : original and test image signals x and y, window size w_size.
output: q index value between two images and q_map, quality map of the test image.

$m_x \leftarrow \text{mean}(x, w\_size)$ — *Computes the local mean.*
$m_y \leftarrow \text{mean}(y, w\_size)$
$s_x^2 \leftarrow \text{variance}(x, w\_size)$ — *Computes the local variance.*
$s_y^2 \leftarrow \text{variance}(y, w\_size)$
$s_{xy} \leftarrow \text{covariance}(x, y, w\_size)$ — *Computes the local covariance.*
$den_c \leftarrow s_x^2 + s_y^2$ — *Computes the contrast factor denominator (2).*
$den_l \leftarrow m_x^2 + m_y^2$ — *Computes the luminance factor denominator (2).*
$q\_map \leftarrow 1$ — *Initialized with ones.*
forall the *sliding window* $w, w \in W$ **do**
  if $den_c(w) = 0$ *and* $den_l(w) \neq 0$ **then**
    $q\_map\,(w) \leftarrow l(x,y|w)$ — *Computes the luminance factor at w window (2).*
  else
    $q\_map\,(w) \leftarrow s(x,y|w))\cdot l(x,y|w)\cdot c(x,y|w)$ — *Computes the Q index at w window (2).*
$q \leftarrow \frac{1}{|W|}\sum_{w\in W} q\_map\,(w)$ — *Compute the Q index (3).*

A Matlab implementation of the Q index is available online at https://ece.uwaterloo.ca/~z70wang/research/quality_index/demo.html.

2.1.2 Structural Similarity Index ($SSIM$)

In order to measure the structural distortions of two images x and y, Wang et al. proposed a structural similarity ($SSIM$) index [12], defined as follows

$$\begin{aligned} SSIM(x,y) &= [s(x,y)]^{\gamma} \cdot [l(x,y)]^{\alpha} \cdot [c(x,y)]^{\beta} \\ &= \left(\frac{S_{xy}+C_3}{S_xS_y+C_3}\right)^{\gamma} \cdot \left(\frac{2\bar{x}\bar{y}+C_1}{\bar{x}^2+\bar{y}^2+C_1}\right)^{\alpha} \cdot \left(\frac{2S_xS_y+C_2}{{S_x}^2+{S_y}^2+C_2}\right)^{\beta}. \end{aligned} \tag{4}$$

The parameters α, β and γ adjust the relative importance of the three components. The constants C_1, C_2 and C_3 are included to avoid instability when denominators are very close to zero. In order to simplify the expression (4), Wang et al. set $\alpha = \beta = \gamma = 1$ and $C_3 = C_2/2$. This results in a specific form of the $SSIM$ index

$$SSIM(x,y) = \frac{(2\bar{x}\bar{y}+C_1)(2S_{xy}+C_2)}{(\bar{x}^2+\bar{y}^2+C_1)\left({S_x}^2+{S_y}^2+C_2\right)}. \tag{5}$$

In [12], the authors abovementioned considered $C_1 = (k_1L)^2$ and $C_2 = (k_2L)^2$, where L is the dynamic range of the pixel values (255 for 8-bit grayscale images), and k_1 y k_2 are small constants. The $SSIM$ measure uses the parameter setting: $k_1 = 0.01$ and $k_1 = 0.03$. These values are somewhat arbitrary, but Wang et al. found that in their experiments on benchmark databases, the performance of the $SSIM$ index algorithm was fairly insensitive to variations of the values.
The Q index, defined in (1), corresponds to a special case of $SSIM$ index, if $C_1 = C_2 = 0$.

As stated above, it is suitable for image quality assessment to apply the $SSIM$ index locally rather than globally. To evaluate the overall image quality, they used a mean $SSIM$ ($MSSIM$) index, computed by averaging the local indexes $SSIM(x, y|w)$ in $w \in W$,

$$MSSIM = \sum_{w \in W} \frac{SSIM(x,y|w)}{|W|}. \tag{6}$$

With the aim that the $MSSIM$ measure exhibits a better consistency with the qualitative visual appearance, Wang et al. used as default a 11×11 circular symmetric Gaussian weighting function w $= \{w_i | i = 1, \ldots, M \cdot N\}$, with standard deviation 1.5 samples, normalized to unit sum $\left(\sum_{i=1}^{M \cdot N} w_i\right)$.

Pseudo-code of $SSIM$ Index. A pseudo-code for the implementation of the $SSIM$ index algorithm is shown in Algorithm 2.

An online Matlab implementation of $SSIM$ index is available at https://ece.uwaterloo.ca/~z70wang/research/ssim/.

2.1.3 Codispersion Quality Index (CQ)

Ojeda et al. [5] suggested a modification of the Q index based on the codispersion coefficient. This similarity measure, labeled CQ index, captures different levels of spatial similarity between two images by considering different directions, h, in the two-dimensional space. This is not commonly assessed by other measures of similarity between images. The CQ index, in a specific direction h, is defined by the following equation

$$CQ(x,y,h) = \widehat{\rho}\,(h) \cdot l(x,y) \cdot c(x,y) = \frac{\sum\limits_{s,s+h \in D} a_s b_s}{\sqrt{\widehat{V}_x\,(h)\,\widehat{V}_y\,(h)}} \cdot \frac{2\bar{x}\bar{y}}{\bar{x}^2+\bar{y}^2} \cdot \frac{2S_xS_y}{{S_x}^2+{S_y}^2}, \tag{7}$$

where $\widehat{\rho}\,(h)$, is the sample codispersion coefficient in the direction h, with $s = (i, j)$, $h = (h_1, h_2)$, $D \subset Z^d$, D a finite set, $a_s = x_{s+h} - x_s$, $b_s = y_{s+h} - y_s$, $\widehat{V}_x\,(h) = \sum\limits_{s,s+h \in D} {a_s}^2$, and $\widehat{V}_y\,(h) = \sum\limits_{s,s+h \in D} {b_s}^2$.

Algorithm 2: Pseudo-code of the $SSIM$ similarity quality index algorithm.

input : original and test image signals x and y, constants $k = [k_1\ k_2]$, local window for statistics $window$, dynamic range of the images L and window size w_size.
output: $mssim$ index value between two images and $ssim\_map$, quality map of the test image.

$C_1 \leftarrow (k_1 \cdot L)^2$
$C_2 \leftarrow (k_2 \cdot L)^2$
$x \leftarrow \text{convolution}(x, window)$
$y \leftarrow \text{convolution}(y, window)$
$w\_size \leftarrow$ *size of the window.*
$m_x \leftarrow \text{mean}(x, w\_size)$ — *Computes the local mean.*
$m_y \leftarrow \text{mean}(y, w\_size)$
$s_x^2 \leftarrow \text{variance}(x, w\_size)$ — *Computes the local variance.*
$s_y^2 \leftarrow \text{variance}(y, w\_size)$
$s_{xy} \leftarrow \text{covariance}(x, y, w\_size)$ — *Computes the local covariance.*
$den_l \leftarrow m_x^2 + m_y^2 + C_1$ — *Computes the luminance factor denominator (4).*
$den_c \leftarrow s_x^2 + s_y^2 + C_2$ — *Computes the contrast factor denominator (4).*
$ssim\_map \leftarrow 1$ — *Initialized with ones.*
forall the *sliding window* $w, w \in W$ **do**
  if $C_1 > 0$ *and* $C_2 > 0$ **then**
    $ssim\_map\,(w) \leftarrow s(x, y|w) \cdot l(x, y|w) \cdot c(x, y|w)$ — *Computes the SSIM index at w window (5).*
  else
    if $den_l\,(w) \cdot den_c\,(w) > 0$ **then**
      $ssim\_map\,(w) \leftarrow s(x, y|w) \cdot l(x, y|w) \cdot c(x, y|w)$
      Computes the SSIM index at w window (5).
    if $den_l\,(w) \neq 0$ *and* $den_c\,(w) = 0$ **then**
      $ssim\_map\,(w) \leftarrow l(x, y|w)$
      Computes the luminance factor at w window (4).
$mssim \leftarrow \frac{1}{|W|} \sum_{w \in W} ssim\_map\,(w)$ — *Computes the MSSIM index (6).*

It is obvious that $|\widehat{\rho}\,(h)| \leq 1$.
The CQ value can be calculated locally and averaged to get the overall index

$$CQ = \sum_{w \in W} \frac{CQ(x, y, h|w)}{|W|}. \tag{8}$$

Pseudo-code of the CQ Index. A pseudo-code for the implementation of the previously described CQ index algorithm, is presented in Algorithm 3.

2.2 Image Fusion Metrics

This subsection provides an overview of the metrics based on structural similarity for image fusion and their pseudo-codes[2].

[2]We denoted element-wise product of matrices with .*, the division with ./ and the power with .^.

Algorithm 3: Pseudo-code of the CQ similarity quality index algorithm.

input : original and test image signals x and y, direction h, window size w_size.
output: cq index value between two images and cq_map, quality map of the test image in a specific direction h.

$m_x \leftarrow \text{mean}(x, w\_size)$ *Computes the local mean.*
$m_y \leftarrow \text{mean}(y, w\_size)$
$s_x^2 \leftarrow \text{variance}(x, w\_size)$ *Computes the local variance.*
$s_y^2 \leftarrow \text{variance}(y, w\_size)$
$x_s \leftarrow$ subimage of x composed of the pixels of the set $X' = \{s \in x : s + h \in x\}$ *See Figure 1.*
$x_{s+h} \leftarrow$ subimage of x composed of the pixels $\{s' = s + h : s \in X'\}$
$y_s \leftarrow$ subimage of y composed of the pixels of the set $Y' = \{s \in y : s + h \in y\}$
$y_{s+h} \leftarrow$ subimage of y composed of the pixels $\{s' = s + h : s \in Y'\}$
$a_s \leftarrow x_{s+h} - x_s$
$b_s \leftarrow y_{s+h} - y_s$
$\widehat{V}_x \leftarrow \sum_{s,s+h} {a_s}^2$
$\widehat{V}_y \leftarrow \sum_{s,s+h} {b_s}^2$
$den_{\widehat{\rho}} \leftarrow \sqrt{\widehat{V}_x \cdot \widehat{V}_y}$ *Computes the $\widehat{\rho}$ factor denominator (7).*
$den_l \leftarrow m_x^2 + m_y^2$ *Computes the luminance factor denominator (7).*
$den_c \leftarrow s_x^2 + s_y^2$ *Computes the contrast factor denominator (7).*
$den \leftarrow den_{\widehat{\rho}} \cdot den_l \cdot den_c$ *Computes the CQ index denominator (7).*
$cq\_map \leftarrow 1$ *Initialized with ones.*
forall the *sliding window* $w, w \in W$ **do**
  if $den(w) \neq 0$ **then**
    $cq\_map\,(w) \leftarrow \widehat{\rho}\,(w) \cdot l(x, y|w) \cdot c(x, y|w)$ *Computes the CQ index at w window (7).*
  else
    if $den_{\widehat{\rho}}\,(w) \neq 0$ *and* $den_l\,(w) \neq 0$ *and* $den_c\,(w) = 0$ **then**
      $cq\_map\,(w) \leftarrow \widehat{\rho}\,(w) \cdot l(x, y|w)$
    if $den_{\widehat{\rho}}\,(w) \neq 0$ *and* $den_l\,(w) = 0$ *and* $den_c\,(w) \neq 0$ **then**
      $cq\_map\,(w) \leftarrow \widehat{\rho}\,(w) \cdot c(x, y|w)$
    if $den_{\widehat{\rho}}\,(w) \neq 0$ *and* $den_l\,(w) = 0$ *and* $den_c\,(w) = 0$ **then**
      $cq\_map\,(w) \leftarrow \widehat{\rho}\,(w)$
    if $den_{\widehat{\rho}}\,(w) = 0$ *and* $den_l\,(w) \neq 0$ *and* $den_c\,(w) \neq 0$ **then**
      $cq\_map\,(w) \leftarrow l(x, y|w) \cdot c(x, y|w)$
    if $den_{\widehat{\rho}}\,(w) = 0$ *and* $den_l\,(w) \neq 0$ *and* $den_c\,(w) = 0$ **then**
      $cq\_map\,(w) \leftarrow l(x, y|w)$
    if $den_{\widehat{\rho}}\,(w) = 0$ *and* $den_l\,(w) = 0$ *and* $den_c\,(w) \neq 0$ **then**
      $cq\_map\,(w) \leftarrow c(x, y|w)$
$cq \leftarrow \frac{1}{|W|} \sum_{w \in W} cq\_map\,(w)$ *Computes CQ index (8).*

2.2.1 Piella's Metrics (Q_S, Q_W and Q_E)

Three fusion quality metrics based on the Q index were proposed by Piella and Heijmans [7]. In these measures, $Q\,(x, f|w)$ and $Q\,(y, f|w)$ are calculated between the source images (x and y) and the fused image (f) in a sliding window w.

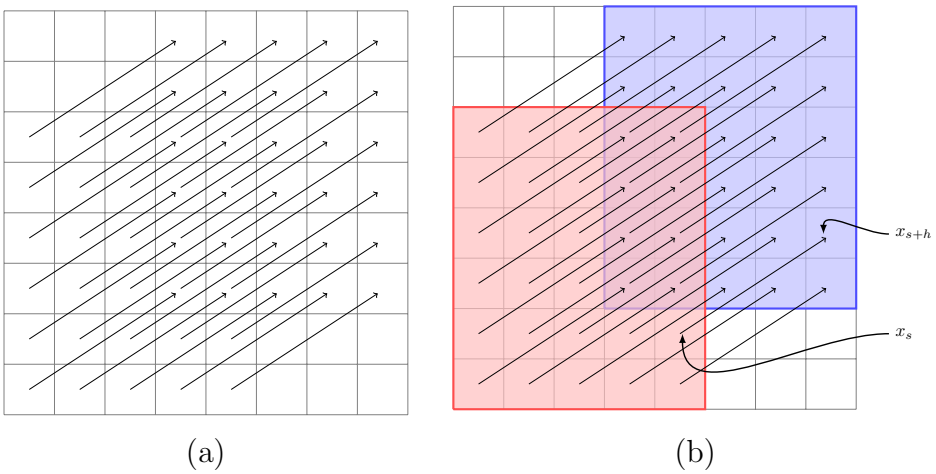


Figure 1: (a) An image x of 8×8 pixels size where the arrows indicate the direction h considered to compute the CQ index. (b) The x_s (subimage of x) composed of the pixels of the set $X' = \{s \in x : s + h \in x\}$ (in red) and the x_{s+h} (subimage of x) composed of the pixels of the set $\{s' = s + h \in x : s \in X'\}$ (in blue).

The first measure is the following

$$Q_S(x, y, f) = \frac{1}{|W|} \sum_{w \in W} \left[\lambda(w)\, Q(x, f|w) + (1 - \lambda(w))\, Q(y, f|w)\right], \tag{9}$$

where the weight $\lambda(w)$ is defined as

$$\lambda(w) = \frac{s(x|w)}{s(x|w) + s(y|w)}. \tag{10}$$

The $s(x|w)$ and $s(y|w)$ are the local saliencies of the two input images x and y within the window w, respectively. They reflect the local relevance of the source image within the window w, and it may depend on contrast, sharpness, or entropy. In Piella's implementation, $s(\cdot|w)$ is the variance of the image within window w.

Pseudo-code of Q_S metric. A pseudo-code for the implementation of the Q_S metric algorithm is shown in Algorithm 5. Algorithm 4 corresponds to the implementation of the local weight $\lambda(w)$, indicating the relative importance of the source images.

The second, *weighted fusion quality index*, gives more weight to those windows where the saliency of the input images is higher. These correspond to areas which are likely to be perceptually important parts of the underlying scene. It is defined as

$$Q_W(x, y, f) = \sum_{w \in W} c(w) \left[\lambda(w)\, Q(x, f|w) + (1 - \lambda(w))\, Q(y, f|w)\right], \tag{11}$$

where $\lambda(w)$ is calculated as in (10) and the coefficient $c(w)$ is computed as

$$c(w) = \frac{C(w)}{\sum_{w' \in W} C(w')}, \tag{12}$$

with

$$C(w) = \max\{s(x|w), s(y|w)\}. \tag{13}$$

Algorithm 4: Pseudo-code of the local weight λ algorithm.

input : source images x and y, window size w_size.
output: weight λ, a map whose elements are the values of λ, computed in each window w of the input images.

$s_x^2 \leftarrow \text{variance}(x, w\_size)$ *Computes the local variance.*
$s_y^2 \leftarrow \text{variance}(y, w\_size)$
forall the *sliding window* $w, w \in W$ **do**
  $den\,(w) \leftarrow s_x^2\,(w) + s_y^2\,(w)$
  if $den\,(w) = 0$ **then**
    $\lambda(w) \leftarrow 0$
  else
    $\lambda(w) \leftarrow \frac{s_x^2(w)}{s_x^2(w)+s_y^2(w)}$ *Computes the weight (10).*

Algorithm 5: Pseudo-code of the Q_S fusion quality metric algorithm.

input : source images x and y, fused image f, window size w_size.
output: q_S, the fusion quality metric value and q_S_map, the quality assessment map of the fusion image.

$q\_map_{xf} \leftarrow Q(x, f, w\_size)$ ▷ *Algorithm 1.*
$q\_map_{yf} \leftarrow Q(y, f, w\_size)$ ▷ *Algorithm 1.*
$\lambda \leftarrow weight\_\lambda\,(x, y, w\_size)$ ▷ *Algorithm 4.*
$q_S\_map \leftarrow \lambda .^* q\_map_{xf} + (1 - \lambda) .^* q\_map_{yf}\,(w)$
$q_S \leftarrow \frac{1}{|W|} \sum_{w \in W} q_S\_map(w)$ *Computes the Q_S metric (9).*

In Piella's implementation, $C(w)$ is defined as in (13). They also suggested another way to calculate it, as $C(w) = s\,(x|w) + s\,(y|w)$.

Remark. If the denominator in (12) is zero, we suggest to set $c = 1/\,|W|$, so in this case the Q_W= Q_S.

Pseudo-code of the Q_W metric. A pseudo-code for the implementation of the Q_W metric algorithm, described above, is presented in Algorithm 7. First, the pseudo-code of the implementation of the coefficient $c\,(w)$, is given (Algorithm 6).

Finally, Piella and Heijmans introduced the *edge-dependent fusion quality index* [7]. This measure takes into account human visual system characteristics, such as sensitivity to edges information. It is obtained as follows

$$Q_{E_1}(x, y, f) = Q_W\,(x, y, f) \cdot Q_W\,(x', y', f')^{\alpha}\,. \tag{14}$$

Another variant of this metric was suggested by Piella in [6], defined as follows

$$Q_{E_2}(x, y, f) = Q_W\,(x, y, f)^{1-\alpha} \cdot Q_W\,(x', y', f')^{\alpha}\,. \tag{15}$$

In (14) and (15), $Q_W\,(x', y', f')$ is the Q_W calculated with the "edge images" x', y' and f' (e.g. the Euclidean norm of the horizontal and vertical gradient images), and α is a parameter that weights the edge contribution information, $\alpha \in [0, 1]$. The closer α is to one, the more important is the edge image.

Algorithm 6: Pseudo-code of weight c algorithm.

input : source images x and y, window size w_size.
output: c coefficient, a map whose elements are the values of c, computed in each window w of the input images.

$s_x^2 \leftarrow \text{variance}(x, w\_size)$ *Computes the local variance.*
$s_y^2 \leftarrow \text{variance}(y, w\_size)$
$c \leftarrow 0$ *Initialized with zeros.*
forall the *sliding window* $w, w \in W$ **do**
  $C(w) \leftarrow \max\left\{s_x^2(w), s_y^2(w)\right\}$ *Computes C at w window (13).*
$den\_c \leftarrow \sum_{w \in W} C(w)$
if $den\_c \neq 0$ **then**
  $c \leftarrow C./\sum_{w \in W} C(w)$ *Computes the c coefficient (12).*
else
  $c \leftarrow 1/|W|$ *See remark in text.*

Algorithm 7: Pseudo-code of the Q_W fusion quality metric algorithm.

input : source images x, y, and fused image f, window size w_size.
output: q_W, the fusion quality metric value and q_W_map, the quality assessment map of the fused image.

$q\_map_{xf} \leftarrow Q(x, f, w\_size)$ ▷ *Algorithm 1.*
$q\_map_{yf} \leftarrow Q(y, f, w\_size)$ ▷ *Algorithm 1.*
$\lambda \leftarrow weight\_\lambda(x, y, w\_size)$ ▷ *Algorithm 4.*
$c \leftarrow weight\_c(x, y, w\_size)$ ▷ *Algorithm 6.*
$q_W\_map \leftarrow c.^*(\lambda.^* q\_map_{xf} + (1 - \lambda).^* q\_map_{yf})$
$q_W \leftarrow \sum_{w \in W} q_W\_map$ *Computes the Q_W metric (11).*

Pseudo-code of the Q_{E_1} metric. A pseudo-code for the implementation of the Q_{E_1} metric algorithm, described above, is presented in Algorithm 8.

Algorithm 8: Pseudo-code of the Q_{E_1} fusion quality metric algorithm.

input : source images x, y and fused image f, window size w_size and α value.
output: q_{E_1}, the fusion quality metric value and $q_{E_1}\_map$, the quality assessment map of the fused image.

$x' \leftarrow \text{norm\_gradient}(x)$ *Compute the norm of image gradient.*
$y' \leftarrow \text{norm\_gradient}(y)$
$f' \leftarrow \text{norm\_gradient}(f)$
$\left[q_{W_{xyf}} \;\; q_W\_map_{xyf}\right] \leftarrow Q_W(x, y, f, w\_size)$ ▷ *Algorithm 7.*
$\left[q_{W_{x'y'f'}} \;\; q_W\_map_{x'y'f'}\right] \leftarrow Q_W(x', y', f', w\_size)$ ▷ *Algorithm 7.*
$q_{E_1}\_map \leftarrow q_W\_map_{xyf}.^*(q_W\_map_{x'y'f'}) \;.\hat{}\; \alpha$
$q_{E_1} \leftarrow q_{W_{xyf}} \cdot q_{W_{x'y'f'}}^{\alpha}$ *Computes the Q_{E_1} metric (14).*

The pseudo-code of the implementation of the Q_{E_2} algorithm is analogous to that of the Q_{E_1} metric.

2.2.2 Cvejic's Metric (Q_C)

Cvejic et al. defined a fusion performance measure [1] that takes into account the local similarity between the source image and the fused image within the same spatial position, as follows

$$Q_C = \frac{1}{|W|} \sum_{w \in W} \text{sim}\left(x, y, f|w\right) \cdot Q\left(x, f|w\right) + \left(1 - \text{sim}\left(x, y, f|w\right)\right) \cdot Q\left(y, f|w\right), \tag{16}$$

with

$$\text{sim}\left(x, y, f|w\right) = \begin{cases} 0, & \text{if} \quad \dfrac{S_{xf}}{S_{xf} + S_{yf}} < 0, \\ \dfrac{S_{xf}}{S_{xf} + S_{yf}}, & \text{if} \quad 0 \leq \dfrac{S_{xf}}{S_{xf} + S_{yf}} \leq 1, \\ 1, & \text{if} \quad \dfrac{S_{xf}}{S_{xf} + S_{yf}} > 1. \end{cases} \tag{17}$$

This weighting factor depends on the similarity in the spatial domain between the source and the fused image. The higher the similarity between the source and fused images, the larger the corresponding weighting factor.

Pseudo-code of the Q_C metric. A pseudo-code for the implementation of the Q_C metric algorithm, is shown in Algorithm 10. In Algorithm 9, the pseudo-code of the implementation of the weighting factor sim is presented.

Algorithm 9: Pseudo-code of the weight sim algorithm.

input : source images x, y and fused image f, window size w_size.
output: sim weighting factor, a map of similarity in spatial domain between the input and the fused image.

$s_{xf} \leftarrow \text{covariance}(x, f, w\_size)$ *Computes the local covariance.*
$s_{yf} \leftarrow \text{covariance}(y, f, w\_size)$
$sim \leftarrow 1$ *Initialized with ones.*
forall the *sliding window* $w, w \in W$ **do**
  if $s_{xf}(w) + s_{yf}(w) = 0$ **then**
    $sim(w) \leftarrow 0$
  else
    $sim(w) \leftarrow s_{xf}(w) / (s_{xf}(w) + s_{yf}(w))$ *Computes the sim weighting factor (17)*
    if $sim(w) > 1$ **then**
      $sim(w) \leftarrow 1$
    if $sim(w) < 0$ **then**
      $sim(w) \leftarrow 0$

2.2.3 Yang's Metric (Q_Y)

Yang et al. proposed another way to use $SSIM$ for fusion assessment [13]. In this metric, the local structural similarity between source images $SSIM\left(x, y|w\right)$ is used as a match measure. For those windows whose $SSIM\left(x, y|w\right)$ is equal to or larger than a given threshold (to distinguish redundant

Algorithm 10: Pseudo-code of the Q_C fusion quality metric algorithm.

input : source images x, y and fused image f, window size w_size.
output: q_C, the fusion quality metric value and the q_C_map, the quality assessment map of the fused image.

$q\_map_{xf} \leftarrow Q(x, f, w\_size)$ ▷ *Algorithm 1.*
$q\_map_{yf} \leftarrow Q(y, f, w\_size)$ ▷ *Algorithm 1.*
$sim \leftarrow weight\_sim(x, y, f, w\_size)$ ▷ *Algorithm 9.*
$q_C\_map \leftarrow sim.^* q\_map_{xf} + (1 - sim).^* q\_map_{yf}$
$q_C \leftarrow \frac{1}{|W|} \sum_{w \in W} q_C\_map(w)$ *Computes the Q_C metric (16).*

information), the weighted average of $SSIM(x, f|w)$ and $SSIM(y, f|w)$ is taken, otherwise, the larger of the two is taken.

$$Q_Y(x, y, f|w) = \begin{cases} \lambda(w) SSIM(x, f|w) + (1 - \lambda(w)) SSIM(y, f|w), \\ \text{if } SSIM(x, y|w) \geq 0.75, \\ \max\{SSIM(x, f|w), SSIM(y, f|w)\}, \\ \text{if } SSIM(x, y|w) < 0.75. \end{cases} \tag{18}$$

The local weight $\lambda(w)$ is defined in (10).

The fusion performance metric, Q_Y, is calculated averaging all values of $Q_Y(x, y, f|w)$ over the whole image,

$$Q_Y(x, y, f) = \sum_{w \in W} \frac{Q_Y(x, y, f|w)}{|W|}. \tag{19}$$

Pseudo-code of the Q_Y metric. A pseudo-code for the implementation of the Q_Y fusion metric algorithm, is given in Algorithm 11.

Algorithm 11: Pseudo-code of the Q_Y fusion quality metric algorithm.

input : source images x, y and fused image f, constants $k = [k_1\ k_2]$, local window for statistics $window$ and dynamic range of the images L.
output: q_Y fusion quality metric value and the q_Y_map, the quality assessment map the fused image.

$ssim\_map_{xf} \leftarrow SSIM(x, f, k, window, L)$ ▷ *Algorithm 2.*
$ssim\_map_{yf} \leftarrow SSIM(y, f, k, window, L)$ ▷ *Algorithm 2.*
$ssim\_map_{xy} \leftarrow SSIM(x, y, k, window, L)$ ▷ *Algorithm 2.*
$\lambda \leftarrow weight\_\lambda(x, y, w\_size)$ ▷ *Algorithm 4.*
$condition \leftarrow ssim\_map_{xy}$
forall the *sliding window* $w, w \in W$ **do**
  if $condition(w) \geq 0.75$ **then**
    $q_Y\_map(w) \leftarrow \lambda(w) \cdot ssim\_map_{xf}(w) + (1 - \lambda(w)) \cdot ssim\_map_{yf}(w)$
  else
    $q_Y\_map(w) \leftarrow \max\{ssim\_map_{xf}(w), ssim\_map_{yf}(w)\}$

Computes the Q_Y metric at the window w (18).
$q_Y \leftarrow \frac{1}{|W|} \sum_{w \in W} q_Y\_map(w)$ *Computes the Q_Y metric (19).*

In order to obtain a image fusion quality measure consistent with human visual evaluations, Yang et al. set the $SSIM$ constants to $k_1 = k_2 = 7.8 \times 10^{-9}$.

2.2.4 Codispersion Fusion Quality Metric (CQ_M)

Following the structure of Piella's metric (11), Pistonesi et al. [8] introduced an objetive measure for image fusion. This fusion measure, labeled CQ_M, is based on a modification of the CQ index, called $CQ_{\max}$ index,

$$CQ_M(x,y,f) = \sum_{w \in W} c\,(w)\,[\lambda\,(w)\,CQ_{\max}\ (x,f|w) + (1-\lambda\,(w))\,CQ_{\max}\ (y,f|w)]\,, \tag{20}$$

with

$$CQ_{\max}(x,y|w) = \max_{\{h \in H\,:\,p(h)\,\geq\,p_0\}} CQ(x,y,h|w), \tag{21}$$

where $p\,(h)$ is the proportion of the pixels in the image corresponding to the direction h in the window w, p_0 is the threshold, denoting the minimum proportion of pixels that is necessary to capture spatial information in different directions and H the family of all directions in a window w (see Figure 8 in Appendix A.1).

In the same sense that the structural similarity indexes were defined in Section 2, the maximum codispersion quality index, $CQ_{\max}$, is determined by averaging all $CQ(x,y,h|w)$ local maximum quality indexes for all the windows $w \in W$,

$$CQ_{\max} = \sum_{w \in W} \frac{CQ_{\max}(x,y|w)}{|W|}. \tag{22}$$

In each evaluated window w, the $CQ_{\max}(x,y|w)$ index, is defined as the maximum value of $CQ(x,y,h|w)$. This implies that $CQ_{\max}$ seeks the direction $h \in H$, that maximizes the CQ in the window w. Note that this direction may not be unique. Its dynamic range is $[-1,1]$.

The CQ_M fusion metric calculates the maximum information transferred from source images into the fused image, by means of the maximum codispersion index.

Pseudo-code of the pixels proportion p, CQ_{max} index and CQ_M metric. Firstly, a pseudo-code for the implementation of the local proportion p of pixels corresponding to the direction h, is presented (Algorithm 12). Then, Algorithm 13 gives a pseudo-code for the implementation of the CQ_{max} index. Finally, Algorithm 14 shows the pseudo-code for the implementation of the CQ_M metric.

Algorithm 12: Pseudo-code of the p proportion algorithm.

input : direction $h = (h_1, h_2)$, window size $w\_size = (w_1, w_2)$.
output: p proportion of the pixels corresponding to the specific direction h in the window w.
$h_1 \leftarrow abs(h_1)$ *Computes the absolute value of the coordinates of the direction h.*
$h_2 \leftarrow abs(h_2)$
if $\left(h_1 > \frac{w_1}{2}\right)$ *or* $\left(h_2 > \frac{w_2}{2}\right)$ **then**
  $p \leftarrow (w_1 - h_1) \cdot (w_2 - h_2) \cdot 2$
else
  $p \leftarrow w_1 \cdot w_2 - 2 \cdot h_1 \cdot h_2$
$p \leftarrow \frac{p}{w_1 \cdot w_2}$ *Computes the proportion of the pixels corresponding to the direction h.*

Note that all these fusion performance measures have a dynamic range of $[-1,1]$. The closer the value to 1, the higher the quality of the composite image.

Algorithm 13: Pseudo-code of CQ_{max} quality index algorithm.

input : original and test image signals x and y, window size w_size and threshold p_0.
output: cq index value between two images and cq_map, quality map of the test image.

$k \leftarrow 0$
forall the $h \in H$ **do**
  See Appendix A.1.
  if $proportion\,(h, w\_size) \geq p_0$ **then**
    Computes the proportion of the pixels corresponding to the direction h in the window w.
    $k \leftarrow k + 1$
    $cq\_maps(k) \leftarrow CQ\,(x, y, h, w\_size)$ ▷ *Algorithm 3.*

$cq\_map\_max \leftarrow max\_similarity\,(cq\_maps)$ *Computes for each window the maximum value of all cq_maps.*
$cq\_max \leftarrow \frac{1}{|W|} \sum_{w \in W} cq\_map\_max\,(w)$ *Computes the CQ_{max} index (22).*

Algorithm 14: Pseudo-code of CQ_M fusion quality metric algorithm.

input : source images x, y and fused image f, window size w_size and threshold p_0.
output: cq_M, the fusion quality metric value and cq_M_map, the quality assessment map of the fused image.

$\lambda \leftarrow weight\_\lambda\,(x, y, w\_size)$ ▷ *Algorithm 4.*
$c \leftarrow weight\_c\,(x, y, w\_size)$ ▷ *Algorithm 6.*
$cq\_max\_map_{xf} \leftarrow CQ_{max}(x, f, w\_size)$ ▷ *Algorithm 13.*
$cq\_max\_map_{yf} \leftarrow CQ_{max}(y, f, w\_size)$ ▷ *Algorithm 13.*
$cq_M\_map \leftarrow c.^* \left[\lambda.^* cq\_max\_map_{xf} + (1 - \lambda)\,.^* cq\_max\_map_{yf}\right]$
$cq_M \leftarrow \sum_{w \in W} cq_M\_map\,(w)$ *Computes the CQ_M metric (20).*

3 Examples

In this section, we present examples of the implemented fusion metrics applied to three different pairs of images used in distinct applications: infra-red (IR) and visual (V), computed tomography (CT) and magnetic resonance imaging (MRI) and multifocus images. In all of them, the following image fusion algorithms were evaluated: Laplacian Pyramid (LP) [14], Ratio Pyramid (RP) [14], Discrete Wavelet Transform (DWT)[3], and Shift Invariant DWT (SIDWT) [9]. The performances of these algorithms have been subjectively tested and accepted in the literature. For the simulation of these methods, the "Image Fusion Toolbox", provided by Rockinger[3], is used.

Parameter Setting. For the four image fusion methods, the approximation coefficients of the two input images averaged and the larger absolute values of the high subbands are selected. In the first example, we performed a 3-level decomposition and in the second and third, a 4-level decomposition was used. In order to ensure fairness of comparison, in the algorithms of the fusion quality metrics we used the same setting that appears in [7],[1],[13] and [8]. For Piella's, Cvejic's and Codispersion Fusion Quality metrics we used the same window size, 8×8 pixels. Also, for the two versions of the edge-dependent fusion quality Q_{E_1} and Q_{E_2} measures, we took as contribution parameter of the edge images: $\alpha = 1$ and $\alpha = 0.5$, respectively. For the proportion parameter of the Codispersion Fusion Quality metric, we set $p_0 = 0.75$ and for Yang's metric, the constants $C_1 = C_2 = 2 \times 10^{-16}$, the dynamic range of the images, $L = 255$ and the w window size used was 7×7 pixels.

[3] http://www.metapix.de/toolbox.htm/

First Example. The "TNO UN Camp" database used as source images consists of 32 sets of infrared (IR) and visual images (V) (see Figure 2).

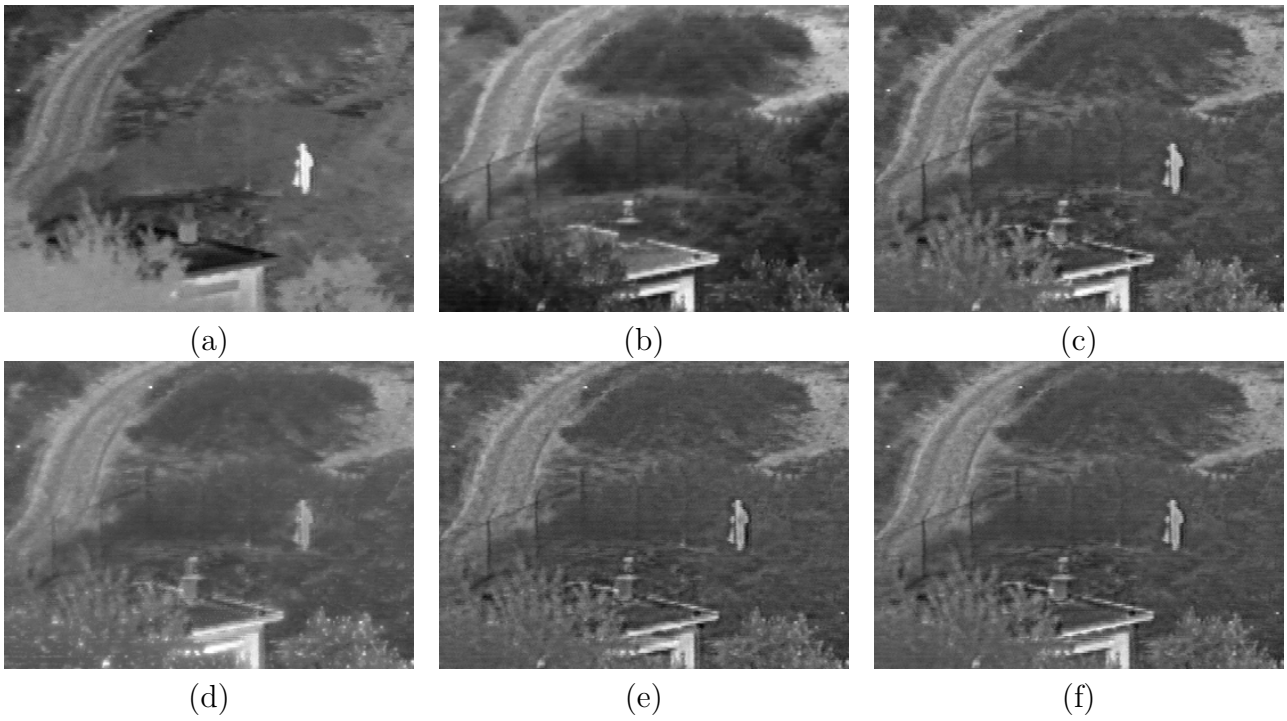


Figure 2: An image of the "TNO UN Camp" database: (a) IR image (b) V image, and (c)-(f) fused image obtained by LP, RP, DWT and SIDWT methods.

Second Example. For the following example, a magnetic resonance imaging (MRI) and computed tomography (CT) image are used (see Figure 3). A visual representation of the fusion performance metrics is represented in Figure 4. These maps allow to visualize the local information about the spatially varying quality of the fused image. In these maps, the brightness indicates the magnitude of the local fusion metric.

Third Example. In this example, a pair of multi-focus images are used (see Figure 5). The quality assessment maps are shown in Figure 6.

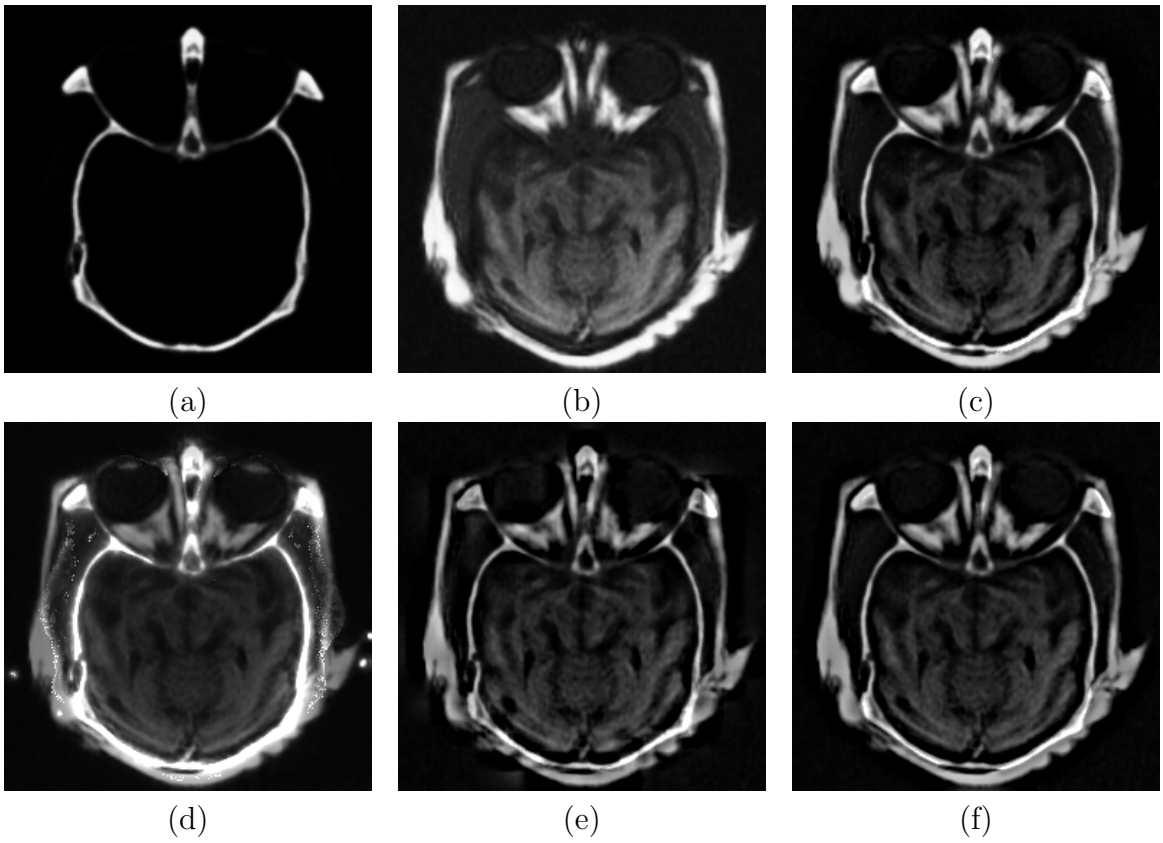


Figure 3: Medical images. (a) a CT source image (b) a MRI source image, and (c)-(f) the fused image obtained by different image fusion methods: LP, RP, DWT and SIDWT, respectively.

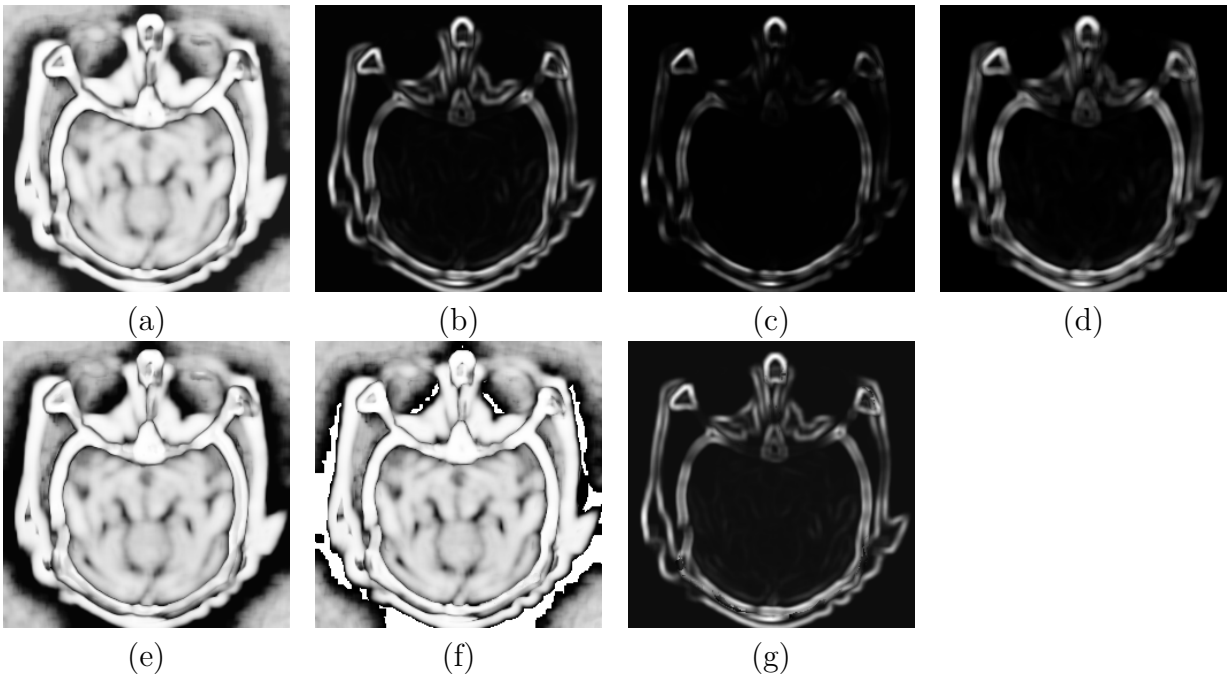


Figure 4: Fusion metrics maps of the fused image using LP method: (a) Q_S map (b) Q_W map (c) Q_{E_1} map (d) Q_{E_2} map (e) Q_C map (f) Q_Y map (g) CQ_M map, respectively. Brightness indicate the magnitude of the local fusion metric.

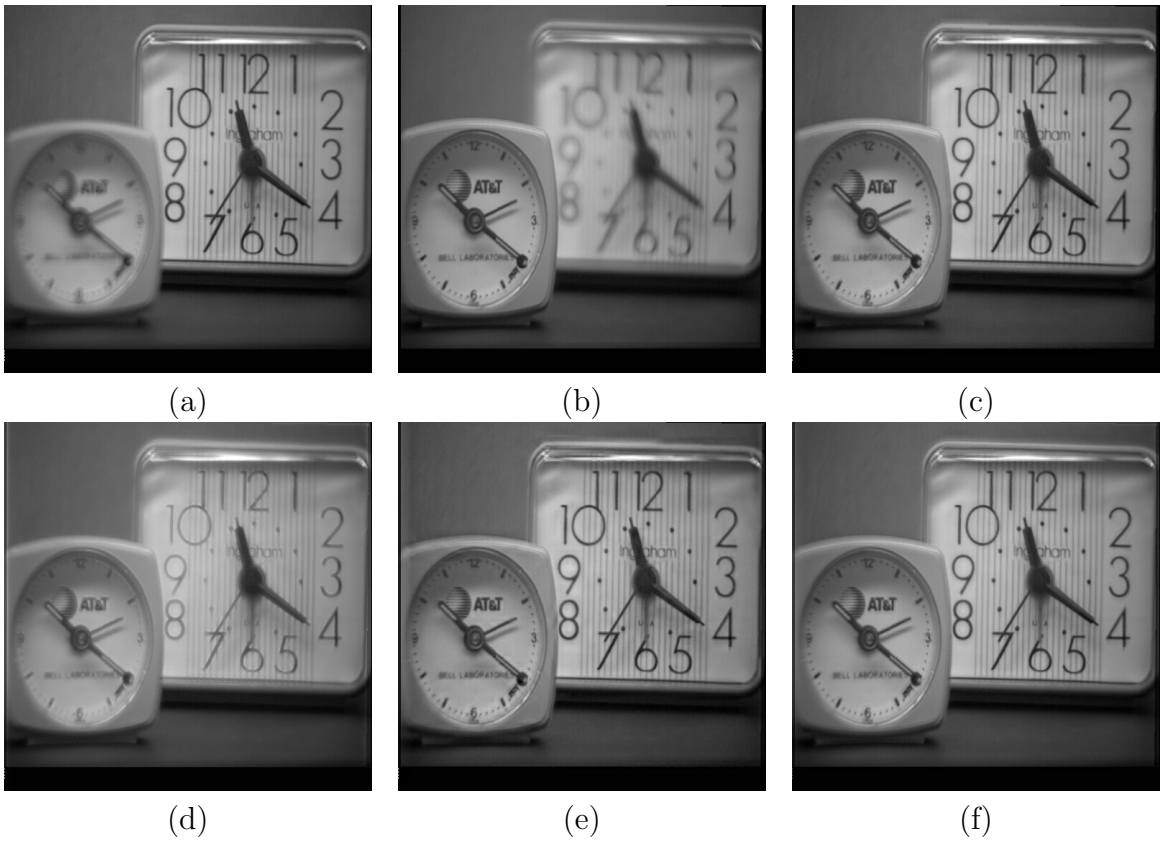


Figure 5: Clock images. (a) a focus on right source clock image (b) a focus on left source clock image, and (c)-(f) the fused image obtained by: LP, RP, DWT and SIDWT methods, respectively.

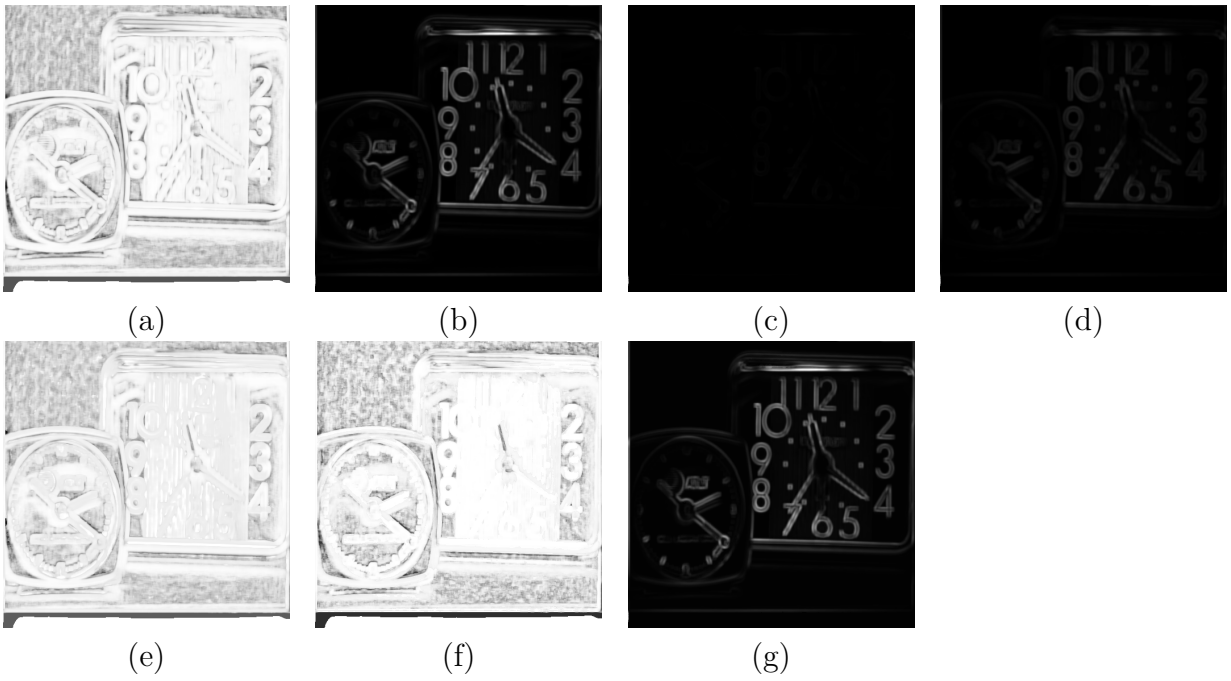


Figure 6: Fusion metrics maps of the fused image using LP method (a) Q_S map (b) Q_W map (c) Q_{E_1} map (d) Q_{E_2} map (e) Q_C map (f) Q_Y map (g) CQ_M map, respectively. Brightness indicate the magnitude of the local fusion metric.

3.1 Results Analysis

In the three different applications, the fusion metrics present similar behaviours, assigning the highest values to the LP and SIDWT methods, followed by DWT and the worst values correspond to the RP method (see Tables 2, 4 and 5). They have a coherent performance with the perceptual evaluations. Resulting images in Figures 2, 3 and 5 clearly certify this.

As seen in the first example, the CQ_M has the lowest standard deviation for LP, SIDWT and DWT fusion methods (see Table 2). This fact demonstrates the good stability of the CQ_M metric for these specific fusion methods [4].

	Methods			
Metrics	LP	RP	DWT	SIDWT
Q_S	0.6730 ± 0.0072	0.5925 ± 0.0108	0.6374 ± 0.0084	0.6851 ± 0.0076
Q_W	0.7576 ± 0.0062	0.6289 ± 0.0047	0.7103 ± 0.0053	0.7511 ± 0.0050
Q_{E_1}	0.5338 ± 0.0130	0.3213 ± 0.0064	0.4716 ± 0.0104	0.5366 ± 0.0105
Q_{E_2}	0.7305 ± 0.0090	0.5668 ± 0.0056	0.6867 ± 0.0076	0.7325 ± 0.0072
Q_C	0.7172 ± 0.0039	0.6488 ± 0.0061	0.6810 ± 0.0050	0.7228 ± 0.0044
Q_Y	0.7806 ± 0.0048	0.7044 ± 0.0071	0.7452 ± 0.0060	0.7877 ± 0.0052
CQ_M	0.8332 ± 0.0033	0.7354 ± 0.0072	0.8108 ± 0.0034	0.8417 ± 0.0033

Table 2: Objective evaluations of different image fusion metrics for the fused images in "TNO UN Camp" database (Mean ± Standard Deviation).

The evaluation results of "TNO UN Camp" images, in Table 3, reveal that the four largest values of Kendall τ rank correlation coefficient [2] correspond to the correlation of Q_Y and Q_C, Q_Y and Q_S, Q_C and Q_S and Q_W with Q_{E_1} and Q_{E_2}, respectively. These outcomes are consistent with those obtained by Lui et al. [4]. Also, CQ_M shows reasonable agreement with all of them.

Kendall Correlation	Q_S	Q_W	Q_{E_1}	Q_{E_2}	Q_C	Q_Y	CQ_M
Q_S	1	0.594	0.649	0.649	**0.937**	**0.954**	0.786
Q_W		1	**0.847**	**0.847**	0.605	0.594	0.706
Q_{E_1}			1	1	0.658	0.642	0.781
Q_{E_2}				1	0.658	0.642	0.781
Q_C					1	**0.969**	0.771
Q_Y						1	0.770
CQ_M							1

Table 3: The correlation matrix of image fusion metrics for the fused images in "TNO UN Camp" database. The best Kendall τ correlation coefficient values are highlighted in bold typeface.

The corresponding fusion evaluation results given in Tables 4 and 5 show that the LP method has higher values of Q_{E_1} and Q_{E_2} (metrics that take into account edge information) than the SIDWT method. This is due to its high-pass characteristic and to its known ability to preserve edges and reduce the ringing artifacts around them [7]. There is a stronger separation between the good results (LP, SIDWT and DWT) and the bad results (RP).

From the examples we can appreciate that the fusion metrics provide a relative assessment on the quality of the fused image, depending on the goals and operational requirements of the application of interest. For a general discussion about the fusion metrics, see Lui et. al [4].

	Methods			
Metrics	LP	RP	DWT	SIDWT
Q_S	0.6168	0.6020	0.6090	0.6506
Q_W	0.8089	0.6319	0.7314	0.7780
Q_{E_1}	0.6565	0.3217	0.5598	0.6429
Q_{E_2}	0.8102	0.5672	0.7482	0.8018
Q_C	0.6247	0.6053	0.6190	0.6587
Q_Y	0.6874	0.6182	0.6368	0.6692
CQ_M	0.8391	0.6903	0.7718	0.8169

Table 4: Objective evaluations of different image fusion metrics for a fused medical image.

	Methods			
Metrics	LP	RP	DWT	SIDWT
Q_S	0.8220	0.7440	0.7824	0.8366
Q_W	0.9272	0.7878	0.9139	0.9217
Q_{E_1}	0.8188	0.4774	0.8025	0.8027
Q_{E_2}	0.9048	0.6910	0.8958	0.8959
Q_C	0.8284	0.7564	0.7919	0.8368
Q_Y	0.8816	0.7879	0.8461	0.8853
CQ_M	0.9451	0.8257	0.9362	0.9413

Table 5: Objective evaluations of different image fusion metrics for a fused clock image.

A Appendix

A.1 Implementation Details

In this subsection, we present some details about directions h within a window w, considered in the CQ and CQ_{max} implementation algorithms.

We assume that a window w is a two-dimensional rectangular grid consisting of $m \times n$ pixels, $m, n \in \mathbb{N}$, evenly spaced. The pixel location is denoted by $s = (i, j), i = 1, \ldots, m; j = 1, \ldots, n$. For our implementation, the window can be characterized by a matrix structure, where the row is the first pixel coordinate and the column, the second (see Figure 7).

Let H be the set of all directions of a window w of an image. In order to facilitate the algorithmic implementation, we considere H as the union of two disjoint sets H_1 ($H_1{}^-$) and H_2 ($H_2{}^-$),

$$H_1 = \{h = (h_1, h_2) \in H | 0 \leq h_1 \leq (m-1), 1 \leq h_2 \leq (n-1), h_1, h_2 \in \mathbb{Z}\},$$

$$\left(H_1{}^- = \{h = (h_1, h_2) \in H | -(m-1) \leq h_1 \leq 0, -(n-1) \leq h_2 \leq -1, h_1, h_2 \in \mathbb{Z}\}\right)$$

and

$$H_2 = \{h = (h_1, h_2) \in H | 1 \leq h_1 \leq (m-1), -(n-1) \leq h_2 \leq 0, h_1, h_2 \in \mathbb{Z}\}.$$

$$\left(H_2{}^- = \{h = (h_1, h_2) \in H | -(m-1) \leq h_1 \leq -1, 0 \leq h_2 \leq (n-1), h_1, h_2 \in \mathbb{Z}\}\right)$$

Figure 8 illustrates the H_1 and H_2 sets considered in a 8×8 window size.

x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}
x_{21}	x_{22}	x_{23}	x_{24}	x_{25}	x_{26}	x_{27}	x_{28}
x_{31}	x_{32}	x_{33}	x_{34}	x_{35}	x_{36}	x_{37}	x_{38}
x_{41}	x_{42}	x_{43}	x_{44}	x_{45}	x_{46}	x_{47}	x_{48}
x_{51}	x_{52}	x_{53}	x_{54}	x_{55}	x_{56}	x_{57}	x_{58}
x_{61}	x_{62}	x_{63}	x_{64}	x_{65}	x_{66}	x_{67}	x_{68}
x_{71}	x_{72}	x_{73}	x_{74}	x_{75}	x_{76}	x_{77}	x_{78}
x_{81}	x_{82}	x_{83}	x_{84}	x_{85}	x_{86}	x_{87}	x_{88}

Figure 7: Matrix structure of a window w (size 8×8 pixels) for the algorithm implementation.

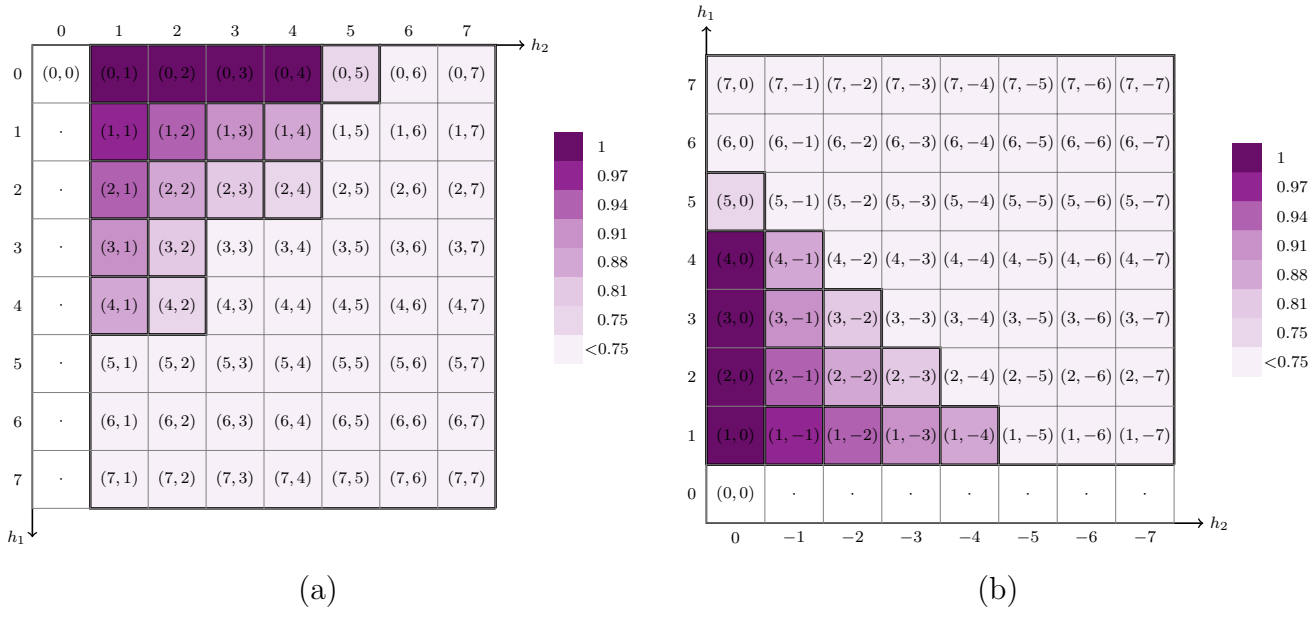


Figure 8: Directions h of a window w for the algorithm implementation. (a) H_1 directions set (b) H_2 directions set. Proportion of the pixels in the window corresponding to the direction h. Darkness indicates more pixels proportion.

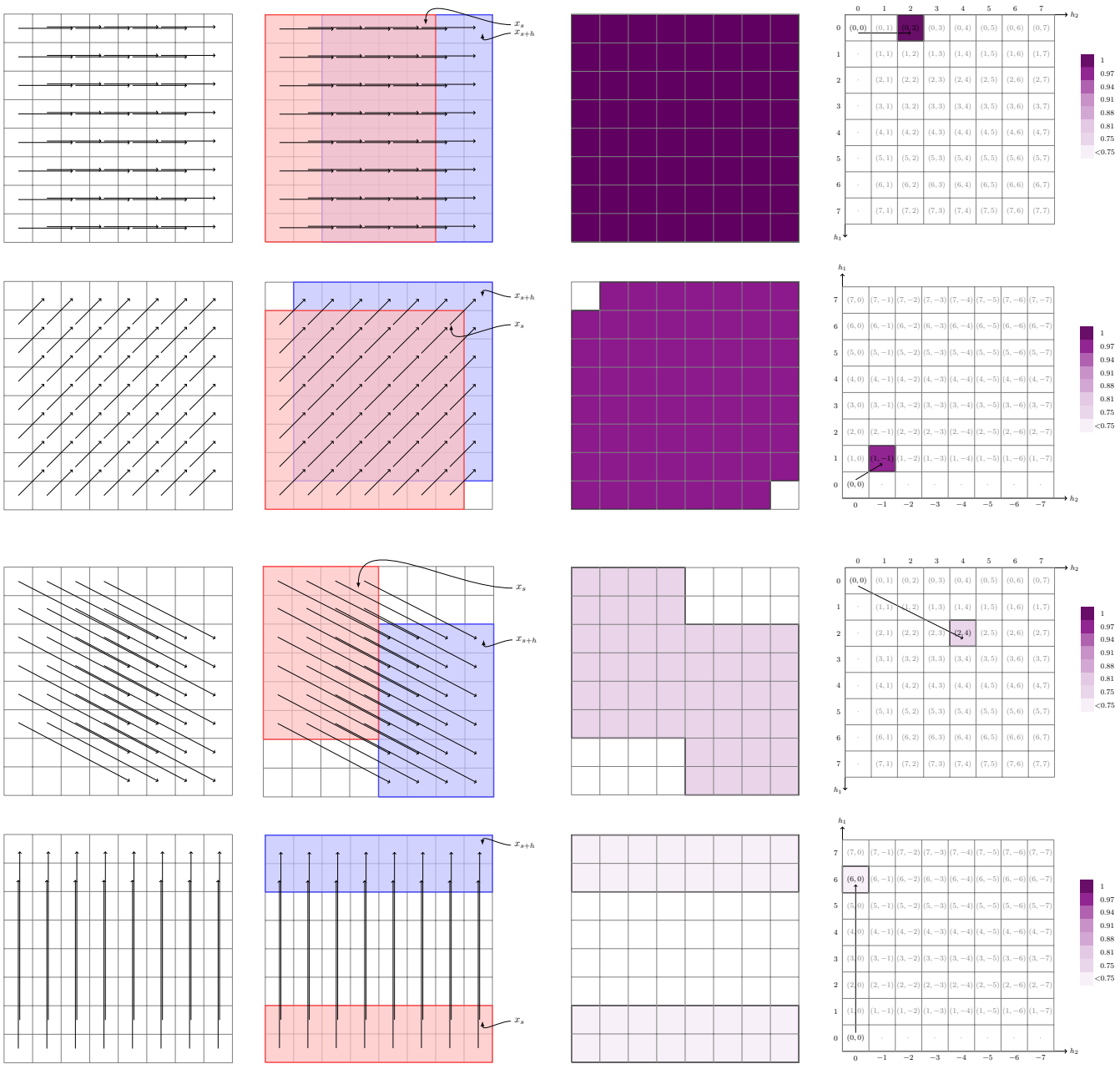
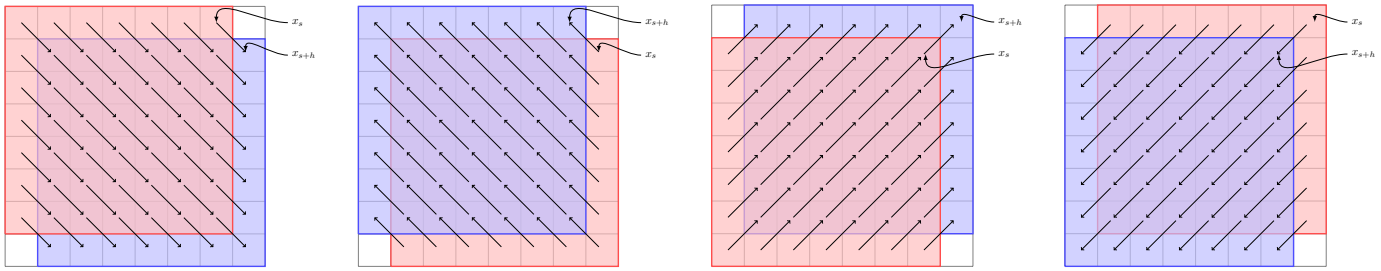


Figure 9: Example of four different directions h. First column, a horizontal direction (—), a diagonal direction (/), another diagonal direction (\) and a vertical direction (|). Second column, x_s, coloured in red and x_{s+h}, in blue. Third column, the pixels proportion corresponding to these directions h: 1, 0.97, 0.75 and 0.5, respectively. Darkness indicates more pixels proportion. Fourth column, the $h = (0, 2)$, $h = (1, -1)$, $h = (2, 4)$ and $h = (6, 0)$ directions, respectively.

An example of four different directions h is represented in Figure 9.

It can be proved that if $h \in H_1$, $\exists\, h' \in {H_1}^-$ such that $\hat{\rho}(h) = \hat{\rho}(h')$ (for example, let $h = (1,1) \in H_1$ and $h' = (-1,-1) \in {H_1}^-$). Similarly, if $h \in H_2$, $\exists\, h' \in {H_2}^-$ such that $\hat{\rho}(h) = \hat{\rho}(h')$ (for instance, let $h = (1,-1) \in H_2$ and $h' = (-1,1) \in {H_2}^-$). Figure 10 demonstrates these results visually. For these reasons, we take into account the sets H_1 and H_2.



(a) (b)

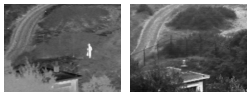
Figure 10: Example of equivalent directions h. (a) $h = (1,1) \in H_1$ and $h' = (-1,-1) \in {H_1}^-$ (b) $h = (1,-1) \in H_2$ and $h' = (-1,1) \in {H_2}^-$.

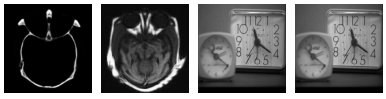
Acknowledgments

S. Pistonesi and J. Martinez would like to thank SGCyT-UNS for financial assistance, S. M. Ojeda supported by SeCyT-UNC, and R. Vallejos by AC3E under grant No FB-0008. Part of this article was written while S. Pistonesi and J. Martinez were visiting the Centre de Mathématiques et de Leurs Applications (CMLA) of the École Normale Supérieure de Cachan, Cachan, France. S. Pistonesi and J. Martinez are also grateful to Miguel Colom, Enric Meinhardt-Llopis and Jean-Michel Morel for their hospitality and kindness. The authors would like to thank the IPOL editors and anonymous reviewers for their valuable insights and suggestions that helped to improve and clarify the manuscript.

Image Credits

All images by the authors except:

 from the TNO UN Camp database[4]

 provided by Rockinger[5]

[4] http://www.deakin.edu.au/~mhossny/fusion/
[5] http://www.metapix.de/toolbox.htm/

References

[1] N. Cvejic, A. Loza, D. Bull, and N. Canagarajah, *A similarity metric for assessment of image fusion algorithms*, International Journal of Signal Processing, 2 (2005), pp. 178–182.

[2] M. G. Kendall, *A New Measure of Rank Correlation*, Biometrika, 30 (1938), pp. 81–93. http://dx.doi.org/10.2307/2332226.

[3] H. Li, B. S. Manjunath, and S. K. Mitra, *Multisensor image fusion using the wavelet transform*, Graph. Models Image Process., 57 (1995), pp. 235–245. http://dx.doi.org/10.1006/gmip.1995.1022.

[4] Z. Liu, Z. Blasch, E. Xue, J. Zhao, R. Laganiere, and W. Wu, *Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 34 (2012), pp. 94–109. http://doi.ieeecomputersociety.org/10.1109/TPAMI.2011.109.

[5] S. M. Ojeda, R.O. Vallejos, and P.W. Lamberti, *Measure of similarity between images based on the codispersion coefficient*, Journal of Electronic Imaging, 21 (2012), pp. 023019–1–023019–5. http://dx.doi.org/10.1117/1.JEI.21.2.023019.

[6] G. Piella, *Adaptative Wavelets and their Aplications to Image Fusion and Compression*, PhD thesis, Faculteit der Natuurwetenschappen, Wiskunde en Informatica. Universiteit van Amsterdam, 2003.

[7] G. Piella and H. Heijmans, *A new quality metric for image fusion*, in Proceedings of the International Conference on Image Processing, 2003, pp. 173–176. http://dx.doi.org/10.1109/ICIP.2003.1247209.

[8] S. Pistonesi, J. Martinez, S. M. Ojeda, and R. O. Vallejos, *A novel quality image fusion assessment based on maximum codispersion*, in Proceedings of the XX Iberoamerican Congress on Pattern Recognition, Alvaro Pardo and Josef Kittler, eds., vol. 9423 of Lecture Notes in Computer Science, Springer, 2015, pp. 383–390. http://dx.doi.org/10.1007/978-3-319-25751-8_46.

[9] O. Rockinger, *Image sequence fusion using a shift-invariant wavelet transform*, in Proceedings of the International Conference on Image Processing (3), 1997, pp. 288–291. http://doi.ieeecomputersociety.org/10.1109/ICIP.1997.632093.

[10] A. Toet and E. M. Franken, *Perceptual evaluation of different image fusion schemes*, Displays, 24 (2003), pp. 25–37. http://dx.doi.org/10.1016/S0141-9382(02)00069-0.

[11] Z. Wang and A.C. Bovik, *A universal image quality index*, IEEE Signal Processing Letters, 9 (2002), pp. 81–84. http://dx.doi.org/10.1109/97.995823.

[12] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, *Image quality assessment: From error visibility to structural similarity*, IEEE Transactions on Image Processing, 13 (2004), pp. 600–612. http://dx.doi.org/10.1109/TIP.2003.819861.

[13] C. Yang, J. Zhang, X. Wang, and X. Liu, *A novel similarity based quality metric for image fusion*, Information Fusion, 9 (2008), pp. 156–160. http://dx.doi.org/10.1016/j.inffus.2006.09.001.

[14] Z. Zhang and R. S. Blum, *A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application*, in Proceedings of the IEEE, vol. 87, 1999, pp. 1315–1326. http://dx.doi.org/10.1109/5.775414.