

Thomas, C., Wald, L., 2006. A MTF-based distance for the assessment of geometrical quality of fused products. In Proceedings (CD-ROM) of Fusion 2006, 10-13 July 2006, Florence, Italy, paper # 267, IEEE Catalog 06EX1311C.

A MTF-Based Distance for the Assessment of Geometrical Quality of Fused Products

Claire Thomas

Ecole des Mines de Paris,
BP 207, 06904 Sophia Antipolis cedex, France
Claire.thomas@ensmp.fr

Lucien Wald

Ecole des Mines de Paris,
BP 207, 06904 Sophia Antipolis cedex, France
Lucien.wald@ensmp.fr

Abstract - *This paper deals with the assessment of the quality of the products resulting from the application of fusion methods, exploiting the synergy of multimodal images at a low spatial resolution and images at a higher spatial resolution but with a lower spectral content. It concentrates on the assessment of the geometrical quality through the analysis of the quality of selected contours. The first objective of our paper is to present a method for the estimation of the MTF. This method applies to excerpts of images containing long well contrasted linear features. The performance of the method is assessed by applying it to several Ikonos images and comparing the results to published works. The second objective is to demonstrate that the method may be exploited to assess the geometrical quality of the fused product.*

Keywords: image fusion, quality assessment, modulation transfer function, satellite images.

1 Introduction

Several approaches of sensor fusion exist which apply on a data set comprising multimodal images at a low spatial resolution and images at a higher spatial resolution but with a lower spectral content. They aim at constructing synthetic multimodal images having the highest spatial resolution available within the data set [1]. Under concern in this paper are only those methods that claim to provide synthetic images close to reality when enhancing the spatial resolution, and not those that provide only a better visual representation of the data set.

This paper deals with the assessment of the quality of the products resulting from the application of one of these methods, the so-called fused products. It concentrates on the assessment of the geometrical quality through the analysis of the quality of selected contours.

The spatial resolution of a sensor is a characteristic of importance for the analysis of its performance. The spatial resolution remains a difficult concept though many attempts were made to better define it. The resolution characterizes the ability of an imaging system to describe how much details are visible in the image, depending on their size. Image quality is often characterized by well-

known measures and criteria, usually linked to radiometry, like mean, contrast, brightness, noise variance, radiometric resolution or granularity. A few tools are available to characterize the spatial performance of an imaging device. The most often used are Point Spread Function-like functions, gradient or high-pass filters in general, Ground Sample Distance (GSD), Full Width at Half Maximum (FWHM), or visual analyses [2, 3, 4, 5].

Each tool is incomplete since it characterizes only partly the geometrical performance in fusion. There is a need for a function or a criterion that characterizes the overall geometrical quality of the images. The problem of specifying resolution and perceived sharpness was solved with the introduction of the Modulation Transfer Function (MTF), a precise measurement made in frequency domain. By definition, the MTF is the Fourier transform of the response of an instrument to a pulse. It decreases the contrast of the image with the frequency, caused by finite detector size, sensor motion, diffraction, aberrations, atmospheric scattering, turbulence, and electronic effects [6]. Only the mean of the image is perfectly preserved, and the MTF tends to gather all gray values around this mean. If the spatial frequency of an object is higher than the cut-off frequency, its image will be characterized by a uniform gray value.

We propose to assess the MTF of fused images and compare it to that of a reference image to quantify the geometrical quality of the synthesized images. Considering that the only valid reference is made of the original multispectral images, we propose to call upon a change in scales and to operate at a lower resolution, an approach promoted by several authors and discussed later.

To do so, we need to know the MTF and we need also a tool for assessing it from an image. In the field of Earth observation, the MTF is usually not delivered by satellite data providers. Information about the intrinsic quality of imaging device products like the shape of the MTF and more particularly the cut-off frequency is not clearly given or even defined in publicly-available literature. Distortions and limit of sensors are well guarded. For instance in the past, the documentation of SPOT Image remains very vague about the MTF [7]. It gives specifications for the cut-off frequency of SPOT panchromatic image without informing about the way it was calculated. Moreover, no information was found

about multispectral images. Another example is the Quickbird specifications: the only available information is: “high contrast (MTF)” and “61 centimeters GSD for panchromatic and 2.44 meter GSD for multispectral at nadir”. The MTF function is not clearly described by manufacturers. In addition, use of various terms like MTF, resolving power, cut-off frequency, minimum allowable MTF (percentage of attenuation of the original contrast) at the Nyquist frequency, create confusion to readers.

The first objective of our paper is to present a method for the estimation of the MTF. This method applies to excerpts of images containing long well contrasted linear features. The performance of the method is assessed by applying it to several Ikonos images and comparing the results to those published by [8, 9].

The second objective is to exploit this tool to assess the geometrical quality of the fused product.

2 Background

The MTF is the amplitude spectrum of the response of a system to a perfect pulse, or Point Spread Function (PSF). For an optical system, the stain of diffusion that stems from a perfect luminous point (called Dirac pulse) corresponds to the PSF of the system. The image of this object consists of a spot of several pixels, bright in the center and progressively darkening away from the center; it is deeply related to sampling rate. It describes how the sharp luminous pulse was spread out by the imaging process. It is known that geometrical distortions and sensor motion can cause asymmetry in the PSF representation, so there might be different spatial resolutions in different directions. Several authors tried to give an idea of the resolution by defining an index based on the PSF function, like the FWHM for Full Width Half Maximum, which is defined as the width within which the PSF drops to half the maximal value. But a single value cannot be representative of such an intricate problem of the overall geometrical quality of a sensor.

The advantage of MTF compared to other indexes is that it provides the spectral behavior of the instrument for each spatial frequency. The more accurate the PSF estimation, the better the MTF one. PSF estimation can be accurately estimated if one processes an image containing a very bright and very narrow spot in a dark and homogeneous scene. Finding or creating such an image is easy for a scanner or a digital camera. But for space-borne sensors, this becomes far more complicated. Even if the luminous pulse can be artificially created using a constructed target like a laser beam, such experiments cannot be performed by all scientists working on remote sensing images.

A solution to this problem is to simulate the pulse from the available information. The usual method for calculating the PSF uses Heavyside-like sharp highly contrasted edges. The image is scanned orthogonal to the edge and a scan provides an Edge Spread Function (ESF). The PSF can be computed by the derivation of this ESF. The edge is not an instantaneous transition of maximum

contrast amplitude but appears to be more gradual with amplitude that does not reach the extremes of the available dynamic range. It gives some information on how the imaging system treats edges. Once the PSF obtained, the MTF is derived from the PSF by applying a Fourier transform to the PSF.

Severe noise problems occur during the derivation phase since derivation acts as a high pass filter on data and prevents from obtaining reliable results. The method we propose focuses on the noise reduction problem.

3 Presentation and validation of the method for MTF estimation

Our method is based on that of [10]. Its principle is that oversampling is a mean to remove noise. It also increases the number of samples, thus better describing the ESF. Artificial oversampling can be performed by the means of interpolation or 0padding created oversampled profiles [11]. But if the edge between the two uniform highly contrasted areas is slightly inclined with respect to the row or column wise direction, one can interleave the successive rows or columns in order to get a naturally oversampled version of the PSF and thus of the MTF in one direction, as shown in Figure 1. Our major improvement to existing methods is to alleviate the use of artificial targets. Our method is able to apply on any image presenting a particular long and tilted edge, as bridges, long roads or even large buildings.

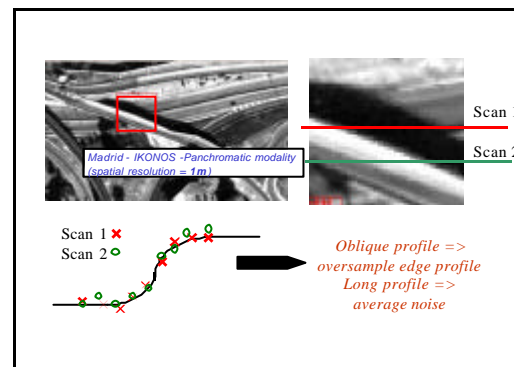


Figure 1. Principle for oversampling

3.1 Step 1: slope parameter estimation

If an edge is inclined away from the vertical direction, each image line corresponds to different versions of the same profile slightly moved because of the tilt. In order to create superimposable profiles, the localization of the inflexion point should be known for each line. This set of points will provide the equation of the straight line approximating the ESF. Given an excerpt of the image exhibiting the edge, such as a bridge, a large road or a building, the location of the maximum gradient is searched on each line crossing the edge. Then, we automatically derive the ESF thanks to the convolution of the image with a Sobel filter. Figure 2 illustrates this point.

The line obtained is often affected by noise due to other surrounding edges. We apply direct Hough transform to estimate the parameters of this line: the slope a and intercept b . An adapted threshold selecting the highest Hough coefficients is followed by the reverse Hough transform; this results in a reconstructed edge of better quality (Figure 3). In order to complete the removal of the noise, a linear regression is performed on the pixels exhibiting 1% of the highest gray levels. Thus, the parameters a and b are obtained.

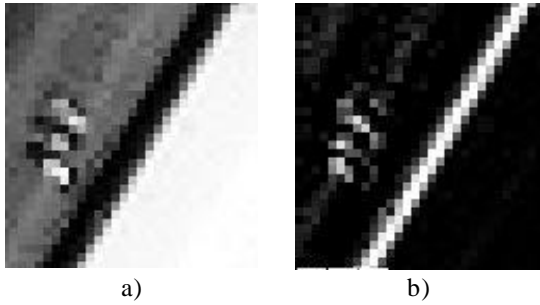


Figure 2. a) original image of the edge, b) numerical derivation (Sobel filter)

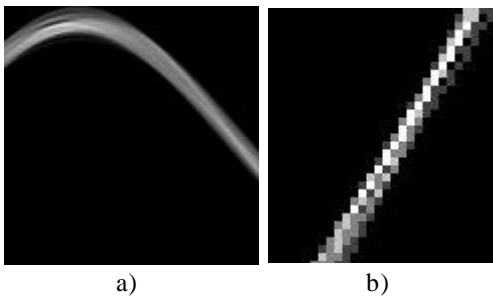


Figure 3. a) Hough coefficients after thresholding, b) reconstructed edge

For this example, we have adjusted by hand a straight line and computed the parameters. They are compared to those resulting from the application of the algorithm (table 1). The values are very close to each other and confirm the relevance of the proposed algorithm.

$a = -1.69$	$a = -1.71$
$b = 45.85$	$b = 46.60$
by algorithm	by hand

Table 1. Edge parameters computed by two means.

Parameters computed	by hand	by the algorithm	Mean absolute error (MAE)
Edge1	$a = -1.83$ $b = 40.45$	$a = -1.78$ $b = 40.02$	MAE on $a = 0.05$
Edge2	$a = -1.73$ $b = 25.81$	$a = -1.78$ $b = 26.12$	
Edge3	$a = -1.83$ $b = 44.28$	$a = -1.80$ $b = 43.20$	
Edge4	$a = -1.79$ $b = 28.50$	$a = -1.70$ $b = 27.44$	
Edge5	$a = -1.71$ $b = 46.60$	$a = -1.69$ $b = 45.85$	

Table 2. Edge parameters computed by two means, and mean absolute error between both estimates, for the five edges.

The same approach was applied to five other edges (table 2). The error made on the slope a is very small compared to that on b ; note that the error on b is smaller than 1 pixel. The slope a is the most important parameter for a good superimposability of the different profiles. From these results, we conclude that this algorithm is a reliable tool for the computation of the edge parameters.

3.2 Step 2: MTF estimation

The second phase of the method is estimation of the MTF. A sigmoid function is adjusted onto the final oversampled profile in order to discard the rest of the noise. Two norms are computed to check the success in adjustment: Chi2 and L2. The smaller the norm, the better the adjustment. There are several advantages of handling an analytical function instead of a numerical one:

- the derivative is easier to compute, since it is also an analytical function;
- the number of samples is perfectly controlled, especially with respect to the Fast Fourier Transform that needs a number of samples of a power of two;
- the inflexion point can be theoretically computed by solving the second derivative of the sigmoid equal to zero. One of the FFT samples should have the value of the inflexion point to avoid an underestimation of the MTF.

The validation of the algorithm is performed by comparing our results to the MTF curve estimations obtained by [8, 9] for panchromatic (Pan) and blue imagery of the satellite Ikonos. We proceed as follows:

- ten edges are selected (five in Pan images, and five in blue ones);
- the algorithm is applied to each edge. A value of the MTF at Nyquist frequency is thus obtained for each edge;
- from these series of values, the mean and the interval of values are computed for each type of imagery. They are compared to those given in [8, 9].

3.3 Results for Pan imagery

The reference values are reported in table 3 [8, 9].

Specification of the MTF	0.17
Interval of values (6 edges processed)	[0.12, 0.20]

Table 3. Reference values for Pan Ikonos imagery (MTF value at Nyquist frequency)

The resolution of Ikonos Pan image is 1 m, which enables the selection of buildings as edges. The shadows of buildings offer the opportunity to obtain very good step amplitude in the selected edge and fairly homogenous areas. The main drawback of buildings is that the edge is often short, since few downtown buildings are larger than 30 meters. The edges and their characteristics are summarized in table 4.

As expected, edges are not very long and do not exceed 30 pixels. Ikonos data are coded on 11 bits, so edges exhibit large amplitudes in the transition. Chi2 and L2 norms deliver similar results except for the first edge. The mean value and the interval of MTF value at Nyquist frequency are a little overestimated compared to literature. However, we note that the norms for the first edge are very large; this shows that the adjustment of the sigmoid is not accurate at all. If we reject this edge, then the mean value becomes 0.19 and the interval [0.13, 0.31]. The mean value given by our method fits in the interval of reference and is close to the reference value of 0.17 (Table 3).

	Edge1	Edge2	Edge3	Edge4	Edge5
Length (m)	26	22	29	21	36
Mean difference in radiometry	1500	1700	1550	1500	1500
Chi2-norm	0.498	0.039	0.008	0.136	0.318
L2-norm	0.025	0.007	0.001	0.006	0.009
MTF value at Nyquist frequency	0.318	0.309	0.138	0.190	0.133
Mean MTF and interval	0.22		[0.13, 0.32]		

Table 4. Edges and results for pan, with all edges taken into account.

3.4 Results for blue imagery

The reference values are reported in table 5 [8, 9]. The selection of edges is difficult and is limited by the resolution. At 4 m, the buildings are not long enough for MTF estimation. The only available edges are bridges or roads. Roads are not appropriate because the gradient in radiometry is too small and blurred with noise, since sides of straight roads are usually fields. The results are presented in table 6.

Specification of the MTF	0.26
Interval of values (6 edges processed)	[0.22, 0.32]

Table 5. Reference values for blue Ikonos imagery (MTF value at Nyquist frequency)

	Edge1	Edge2	Edge3	Edge4	Edge5
Length (m)	212	170	67	162	218
Mean difference in radiometry	800	700	1300	100	200
Chi2-norm	0.020	0.002	0.025	0.007	0.018
L2-norm	0.001	0.012	0.003	0.001	0.003
MTF value at Nyquist frequency	0.358	0.586	0.322	0.138	0.297
Mean MTF and interval	0.28		[0.14, 0.36]		

Table 6. Edges and results for blue, with all edges taken into account.

The second edge is an outlier: the tilt of the slope does not permit an oversampling of the edge. This is confirmed by the overestimated value of the MTF that proves aliasing. We note that edges 4 and 5 have poor dynamics. The mean MTF value at Nyquist frequency is close to the reference in table 5.

Besides confirming that the method permits to retrieve the MTF, this example demonstrates the difficulty of finding appropriate edges for an accurate assessment of the MTF.

4 Evaluation of the geometrical quality of fused products

The proposed method is a good estimator of the MTF of a sensor if several edges are selected in different images in order to increase the probability to get a correct value by averaging. Beside this first functionality, this method is also a mean to assess the quality of an edge. If the edge is blurred, noisy, not perfectly straight or offers a weak gradient, the curve of the MTF will be located below the actual one. This property may be applied to assess the geometrical quality of a fused product.

Given a fused product and the reference image to which the fused product can be compared, any edge present in the reference image should appear in the fused image with similar MTF values. Consequently, if we apply our method to both images, we can evaluate the discrepancies in MTF values and further, evaluate the quality of the synthesis of the geometrical features by the fusion method. The lower the discrepancy, the better the product.

Usually, the reference image is not available. Considering that the only valid reference is made of the original multispectral images, we propose to perform a change in scales and to operate at a lower resolution, an approach promoted by [12, 13, 14]. The two original sets

composed by Pan and multispectral (MS) modalities are downsampled to reach lower resolutions.

The validation is performed on:

- the blue modality of Quickbird imagery, whose edge was selected on side roads around the city of Fredrickton, Canada;
- the blue and the red modalities of Ikonos images, acquired on a building close to the airport of Toulouse, in France.

Downsampling is performed by the mean of a generalized Laplacian pyramid (GLP). Thus, we create the images at lower resolutions. For Quickbird images, the blue modality is degraded down to 11.2 m and 2.8 m for the Pan. For Ikonos, blue and red modalities are transformed to reach 16 m and 4 m for the Pan. These downsampled images are input to the fusion process. The fusion process creates fused products at the original resolution of MS images, respectively 2.8 m and 4 m.

The demonstration that our tool permits the assessment of the geometrical quality is made by applying the tool to two fused products resulting from two fusion methods, one of them being known for its poor geometrical quality. Both methods belong to the ARSIS concept [15, 16]. This concept applies multiscale models on input images in order to estimate a model for the transformation of Pan details into the MS images. Some of these multiscale models are known to distort drastically the geometrical quality of fused products. The hierarchical decomposition model of Mallat using a wavelet from Daubechies is such a case [15]. This wavelet is non symmetric, and therefore introduces a shift of half a pixel at each scale of decomposition: Pan and MS images become not superimposable at MS original resolution. This has a drastic effect on fused edges. The multiscale model used in the second method is identical to that used for degradation (GLP). The transformation of Pan details into MS details is a global method called M3 by its authors [15, 16], and is common to both fusion methods.

Figure 4a exhibits the edge of reference selected in the blue Quickbird modality. This edge is a little noisy and do not cover the whole range of dynamics of the image. Nevertheless, it possesses its own MTF signature. Figure 4b is the fused product resulting from the method "Mallat+M3", and figure 4c the result of the method "GLP+M3". As previously explained, the edge in "Mallat" image offers a very poor visual quality compared to that displayed in the "GLP" image.

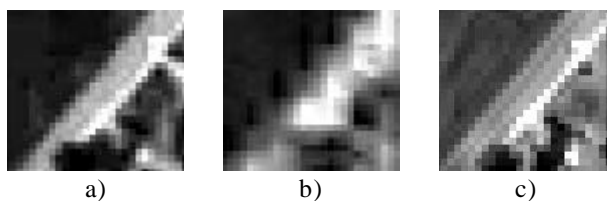


Figure 4. Quickbird Fredrickton – side road, blue modality. a) original, b) Mallat+M3 fused product, c) GLP+M3 fused product.

The MTF curves are displayed in figure 5. The ordinates correspond to the MTF values normalized to the maximum of MTF at the origin. The abscissas are the

spatial frequencies normalized to the sampling frequency. The reference FTM curve is drawn in solid line, in dashed line for the GLP image, and the dotted one for the Mallat image.

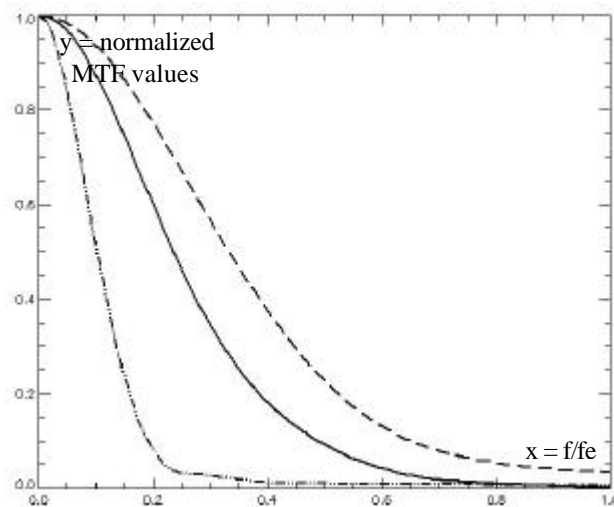


Figure 5. MTF curves: solid line for reference, dashed one for GLP, and dotted one for Mallat.

The MTF of the GLP image is located above that of reference: the GLP edge is more acute than the original one (Figure 4). Nevertheless, it is close to the reference. On the contrary, the Mallat MTF is located far below the reference; the fusion method was not able to transmit the high frequencies to rebuild correctly the edge. The L2 and Chi2 norms quantify the distance between original and synthesized MTF curves. The results are summarized in table 6. They confirm what is seen in figure 5. For a given norm, the difference between the two cases is very large, and large enough to offer a clear discrimination between good and bad quality.

Mallat	GLP
L2 = 3.406	L2 = 0.982
Chi2 = 510.946	Chi2 = 31.862

Table 7. Proximity between reference and the fused products. Quickbird case.

Figure 6 displays an edge as in Figure 4 but for the Ikonos case and blue modality. The pictures in the red modality are very similar.

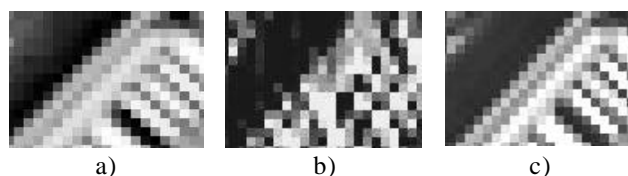


Figure 6. Ikonos Toulouse – building edge, blue modality. a) original, b) Mallat+M3 fused product, c) GLP+M3 fused product.

The MTF curves are displayed in figure 7 for the red modality and figure 8 for the blue one. In both cases, the geometrical quality of the edge in the Mallat image is very poor. The minimization between the Mallat edge and the

sigmoid fails to converge, which explains the absurd overestimation of the MTF displayed in figures 7 and 8. On the contrary, for fusion with GLP decomposition model, the estimation of the MTF is visually very close to the reference in both cases, slightly below for red case and slightly above for blue one.

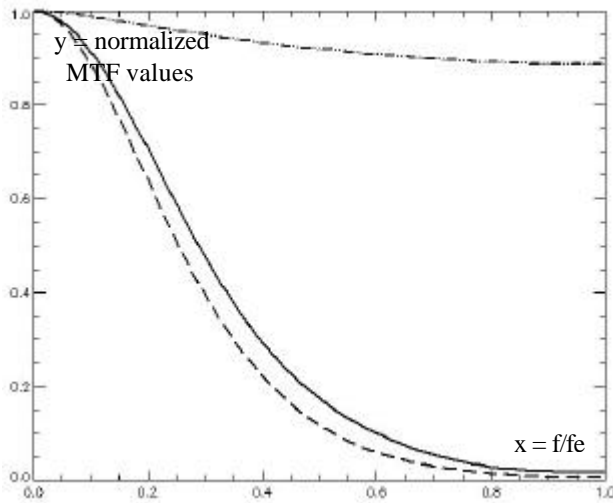


Figure 7. As Figure 5, but for Ikonos and red modality

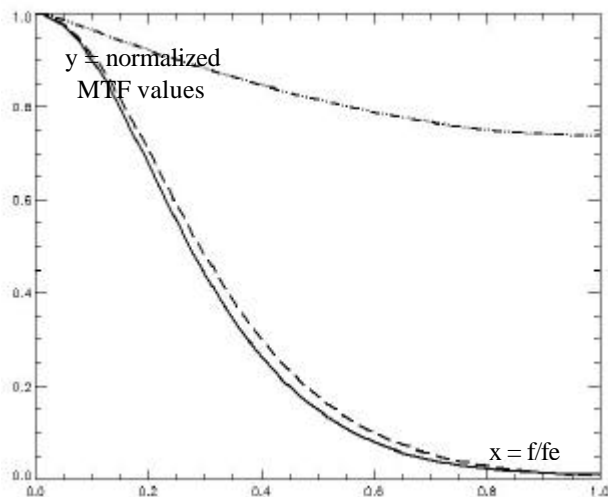


Figure 8. As Figure 5, but for Ikonos and blue modality

These visual observations are confirmed by the L2 and Chi2 values in table 8. The norms are smaller for the GLP image than for the Mallat image and by far. The norm L2 for GLP is smaller in this case than in the previous case. This corresponds to the closer proximity observed visually in figures 7 and 8 compared to figure 5. This may signify that this distance is a tool to quantify the geometrical quality of a fused product.

One should not pay attention to Chi2 underestimated values compared to the first experiment because, when the distance is too large between two series of numbers as in this case, Chi2 also failed to converge and delivers values that are not representative.

	Mallat	GLP
Red	L2 = 29.196 Chi2 = 322.77	L2 = 0.143 Chi2 = 9.76
Blue	L2 = 21.967	L2 = 0.035

Chi2 = 280 Chi2 = 1.346

Table 8. Proximity between reference and the fused products. Ikonos case.

More edges should be tested to determine an accurate and reliable threshold for each of the two distances. The drawback of these distances is that they are not able to distinguish between over- and underestimation of the MTF since values are always positive. A L1 norm (sum of the differences between the two series of values instead of the sum of the squares) may be necessary in that purpose.

5 Conclusions and perspectives

The work presented here demonstrates that the method proposed allows an accurate estimation of the MTF of a sensor provided enough edges of high geometrical quality can be found in several different images from the same sensor.

If enough edges of high quality can be found in a single image, then the MTF of the image may be characterized.

Furthermore, this method may be exploited to qualify the geometrical quality of edges in a fused product.

This work is preliminary. More work on more images is necessary to establish a firm basis for the method. One criticism to our method is that the obtained curve does not represent the actual MTF of a sensor. The reason is that the MTF of a sensor is the convolution of several effects; one of them is the MTF of the optical system, whose profile is the Airy spot function. This function is known in 1-D as the Gibbs effect and creates halos around objects. Airy function is the product of a sinus cardinal and a decreasing exponential whose particularity is to not present a zero tangent at the origin. As it can be seen in the several MTF plots, our model of the MTF proposes a non-null tangent at the origin. One possible improvement consists in the adjustment of the raw MTF approximation by a function able to better model the optics of the sensor.

As for the assessment of quality in image fusion, tests must be multiplied. More images must be processed and more fused products must be tested. More fusion methods should be employed. It would be interesting to exploit results of campaigns performed by image analysts on fused products to relate them to the results of our method.

References

- [1] L. Wald, *Data Fusion. Definitions and Architectures - Fusion of Images of Different Spatial Resolutions*, Presses de l'Ecole, Ecole des Mines de Paris, Paris, France, ISBN 2-911762-38-X, 200 p., 2002.
- [2] J. Li, Spatial quality evaluation of fusion of different resolution images. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Amsterdam, Pays-Bas, vol. 33, 2000.

- [3] J. Zhou, D.L. Civco, J.A. Silvander, A wavelet transform method to merge Landsat TM and SPOT panchromatic data. *International Journal of Remote Sensing*, vol. 19, 4, 743-757, 1998.
- [4] J. Vrabel, Multispectral imagery advanced band sharpening study. *Photogrammetric Engineering and Remote Sensing*, vol. 66, 1, 73-79, 2000.
- [5] IRARS, Civil NIIRS Reference Guide, Imagery Resolution Assessment and Reporting Standards Committee, U.S. Government, Washington, D.C., 4 p., 3 annexes, 1996.
- [6] G.C. Holst, Electro-optical imaging system performance, Bellingham, WA: SPIE Optical Engineering Press : 146-147, 1995.
- [7] Anonymous, Guide des utilisateurs des données SPOT, Editeurs CNES et SPOT Image, Toulouse, France, 3 volumes, 1986, revised in January 1991.
- [8] R. Ryan, B. Baldrige, R.A. Schowengerdt, T. Choi, D.L. Helder, S. Blonski, Ikonos spatial resolution and image interpretability characterization. *Remote Sensing of Environment*, vol. 88, 37-52, 2003.
- [9] G. Dial, H. Bowen, F. Gerlach, J. Grodecki, R. Oleszczuk, Ikonos satellite, imagery, and products. *Remote Sensing of Environment*, vol. 88 : 23-36, 2003.
- [10] D. Léger, F. Viallefont, E. Hillairet, A. Meygret, In-flight refocusing and MTF assessment of SPOT 5 HRG and HRS. In SPIE International Symposium on Optical and Technology, Agia Pelagia, Crete, 22-27 September 2002.
- [11] Luengo, 2000
- [12] C. K. Munechika, J. S. Warnick, C. Salvaggio and J. R. Schott, Resolution enhancement of multispectral image data to improve classification accuracy. *Photogrammetric Engineering & Remote Sensing*, 59, 1, 67-72, 1993.
- [13] L. Wald, T. Ranchin and M. Mangolini, Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images. *Photogrammetric Engineering & Remote Sensing*, 63, 6, 691-699, 1997.
- [14] C. Thomas and L. Wald, Assessment of the quality of fused products. In Proceedings of the 24th EARSeL Symposium "New Strategies for European Remote Sensing", 25-27 May 2004, Dubrovnik, Croatia, Oluic (ed.), Millpress, Rotterdam, ISBN 90 5966 003 X, pp. 317-325, 2005.
- [15] T. Ranchin, L. Wald, Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation. *Photogrammetric Engineering and Remote Sensing*, 66(1), 49-61, 2000.
- [16] T. Ranchin, B. Aiazzi, L. Alparone, S. Baronti, L. Wald, Image fusion. The ARSIS concept and some successful implementation schemes. *ISPRS Journal of Photogrammetry & Remote Sensing*, 58, 4-18, 2003.