

MULTISCALE HAAR TRANSFORM FOR BLUR ESTIMATION FROM A SET OF IMAGES

Lâmân Lelégard, Bruno Vallet, Mathieu Brédif

Université Paris Est, IGN, Laboratoire MATIS
4, Av. Pasteur

94165 Saint Mandé Cedex, FRANCE

laman.lelegard@ign.fr (corresponding author), bruno.vallet@ign.fr, mathieu.bredif@ign.fr
<http://recherche.ign.fr/labos/matis>

Working Groups I/2, III/1, III/4, III/5

KEY WORDS: Sharpness, Haar transform, multiscale, calibration

ABSTRACT:

This paper proposes a method to estimate the local sharpness of an optical system through the wavelet-based analysis of a large set of images it acquired. Assuming a space-invariant distribution of image features, such as in the aerial photography context, the proposed approach produces a sharpness map of the imaging device over 16x16 pixels blocks that enables, for instance, the detection of optical defects and the qualification of the mosaicking of multiple sensor images into a larger composite image. The proposed analysis is based on accumulating of the edge maps corresponding to the first levels of the Haar Transform of each image of the dataset, following the intuition that statistically, each pixel will see the same image structures. We propose a calibration method to transform these accumulated edge maps into a sharpness map by approximating the local PSF (Point Spread Function) with a Gaussian blur.

1 INTRODUCTION

Characterizing the spatial resolution of an imaging system is an important field of image processing and is used for assessing its image quality and for restoration purposes.

This characterization can be obtained by shooting some perfectly known objects, preferably periodic patterns such as Foucault resolution targets or Siemens stars (Fleury and Mathieu, 1956) to deduce the smallest periodic detail discernible by the system through the determination of a Modular Transfer Function (MTF) (Becker et al., 2007). This calibration technique is mainly used as a global Point Spread Function (PSF) characterization of the imaging system. However, some imaging systems (mounted with fisheye lenses for instance) show a very space-dependent resolution. In these circumstances, a local study is more suitable and can be done by using a wall of targets, such as Siemens stars (Kedzieriski, 2008). Another calibration method consists in materializing a punctual source by a laser beam in order to calculate the PSF (Du and Voss, 2004). However, these approaches require that the optical device go through a calibration procedure in a controlled environment where the appropriate targets are displayed.

Conversely, some blind estimation methods were recently presented that use the edges present in an image. In the case of airborne imagery, a mission over an urban area provides images with a large amount of edges. Assuming a Gaussian PSF and an equal distribution of edge orientations, (Luxen and Forstner, 2002) estimates the standard deviation of the Gaussian blur. An alternative way of considering the problem is to study the image frequencies by comparing the local spectrum (obtained after integrating the Fourier Transform of a local patch in an image across the polar coordinate theta) to the global image spectrum (Liu et al., 2008). By varying the size of the patch used, one can choose a compromise between the quality of the frequency estimation and of the localization.

In an intermediate approach (Zhang and Bergholm, 1997) local information (like edges) is considered at different scales using

differences of Gaussians. One interesting assessment is the behavior of edges according to the scale at which they are observed. Sharp edges vanish at coarse scale whereas diffuse ones become sharper when looked at coarser scales. An application proposed by (Zhang and Bergholm, 1997) is blur estimation applied to deduce depth from focus. Multi-scale analysis is also an interesting compromise between spatial and frequency accuracy. The use of Haar wavelets in (Tong et al., 2004) is an interesting intermediate solution. This is the framework that we investigate in this work.

The blind estimation methods cited above rely on a single image so the blur caused by the optical system cannot be distinguished from the smoothness of the imaged object itself. Our contribution is twofold: we overcome this limitation by extending Tong's method (designed for a single image) to a large set of images, and we propose a quantitative characterization of sharpness through a blur radius.

2 OVERVIEW

Our method is based on Tong's blur detection method that relies itself on Haar wavelets. We will start by recapitulating his approach, then explain the two contributions of our paper, and finally expose the assumptions that we make on our datasets.

2.1 Haar wavelets

Tong's method comes down to three main steps (Figure 1):

1. Do a 3 levels Haar wavelets transform.
2. Extract multi-scale normal edges maps E_i^{norm} and maximal edge maps E_i^{max}
3. Apply rules to these maps to estimate the sharpness.

3-level Haar transform of a grey level image (Lenna 512x512 pixels)

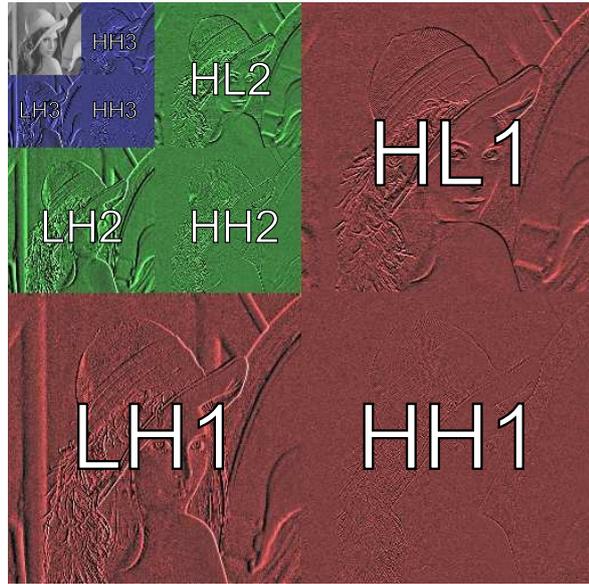


Figure 1: Haar wavelets transform and edge maps

The Haar decomposition of an image I is defined by:

$$\begin{pmatrix} LL_{l+1}(i, j) \\ LH_{l+1}(i, j) \\ HL_{l+1}(i, j) \\ HH_{l+1}(i, j) \end{pmatrix} = M_{Haar} \begin{pmatrix} LL_l(2i, 2j) \\ LL_l(2i, 2j + 1) \\ LL_l(2i + 1, 2j) \\ LL_l(2i + 1, 2j + 1) \end{pmatrix} \quad (1)$$

where L and H stand for Low and High frequencies, $LL_0 = I$ and the Haar matrix is:

$$M_{Haar} = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \quad (2)$$

For each level $l = 1..3$, an edge map is obtained by calculating (for each pixel) the norm:

$$E_l^{norm} = \sqrt{LH_l^2 + HL_l^2 + HH_l^2} \quad (3)$$

The E_l^{norm} do not have the same size, so Tong proposes to define a maximal edge map E_l^{max} of constant size to make possible a

comparison between different levels:

$$E_l^{max}(i, j) = \max_{di, dj=0}^{2^{4-l}-1} E_l^{norm}(2^{4-l}i + di, 2^{4-l}j + dj) \quad (4)$$

thus all E_l^{max} are $2^4 = 16$ times smaller than the input image. The E_l^{max} measure the level of detail of the image at scale 2^l on 16×16 pixels blocks. Tong et al. choose to apply rules based on inequalities on the E_l^{max} in order to characterize qualitatively the image sharpness.

2.2 Our approach

The novelty introduced in this paper compared to Tong's approach is twofold:

1. Compute an average \bar{E}_l^{max} of the E_l^{max} over a large set of images acquired with the same imaging system, such that \bar{E}_l^{max} characterize only the optical quality of the imaging system itself. Obviously, \bar{E}_l^{max} will also depend on the statistical properties of the set of images used.
2. Exploit the \bar{E}_l^{max} to define a quantitative measure of the local sharpness. We chose to quantify local sharpness by assimilating the PSF to a Gaussian blur which radius ($\sigma =$ standard deviation) we will estimate. In other terms, we look for σ as a function:

$$\sigma = \sigma(\bar{E}_1^{max}, \bar{E}_2^{max}, \bar{E}_3^{max}) \quad \sigma : \mathbb{R}^3 \rightarrow \mathbb{R} \quad (5)$$

The main idea developed in this paper is to look for $\sigma(\dots)$ as the composition of two functions:

$$\sigma = c(r(\bar{E}_1^{max}, \bar{E}_2^{max}, \bar{E}_3^{max})) \quad (6)$$

- $r : \mathbb{R}^3 \rightarrow \mathbb{R}$ is a space reduction function, which will reduce our problem from 3 to 1 dimensions. We will explain what properties this function should have and propose a pertinent choice for this function in the next section.
- $c : [0, 1] \rightarrow \mathbb{R}$ is a monotonous calibration function linking an r value to an actual blur radius σ . Because the \bar{E}_l^{max} depend on the actual statistical properties of the dataset used, a calibration function c should be defined for each dataset. Computation of this calibration function as well as its dependence on image statistics is studied in the next section.

2.3 Assumptions

Characterizing an optical system based on a set of images that it acquired will only be valid statistically if these images have good statistical properties. In particular, the following assumptions should be made on the image dataset:

- Camera settings are constant for all the images.
- Shot objects must be in focus.
- Images should be shot without motion blur.
- The edge presence probability is uniform on the whole image.
- The number of images should be large enough.

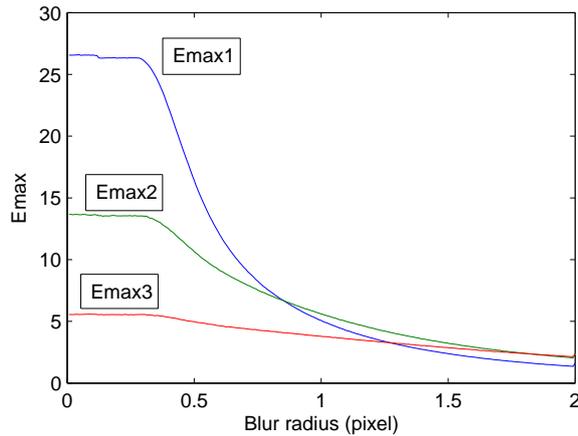


Figure 2: \bar{E}^{max} values for white noise images blurred by a Gaussian of increasing radius.

All these assumptions are usually verified in the case of aerial imagery, provided that motion blur is corrected (by charge transfer for instance), and that the area covered has sufficient texture and edges (forest, city, ...) It is not the case for landscape photographs where the edges are localized at the bottom of the images, and objects at various distances cannot all be in focus simultaneously.

3 BLUR ESTIMATION FROM EDGE MAPS

The aim of this section is to choose the space reduction function r , and to propose a method to compute the calibration function c from a dataset.

3.1 A first experiment

In order to make the exposition clearer, we will start by a simple experiment to exhibit the dependence of the \bar{E}^{max} on blur. The experiment consists in computing the \bar{E}^{max} on a dataset of white noise images (with a given standard deviation), blurred by a Gaussian blur of known radius σ varying from 0 to 2 pixels (Fig. 2). It is easy to see that \bar{E}^{max} depends linearly on the contrast, so their absolute value do not convey a real meaning. Interestingly, the \bar{E}_1^{max} and \bar{E}_2^{max} curves cross between 0.5 and 1, and \bar{E}_2^{max} and \bar{E}_3^{max} cross between 1 and 2. This is quite natural as \bar{E}_1^{max} , \bar{E}_2^{max} and \bar{E}_3^{max} corresponds to scales 1, 2 and 4 pixels respectively, and a Gaussian of standard deviation σ is a structure of size (diameter) 2σ . This is very coherent with the idea that \bar{E}_i^{max} exhibits certain scales in the image. Quite logically, the curves are flat below $\sigma = 0.5$ pixels as structures of size lower than 1 pixel are not representable, such that blurs of radii lower than 0.5 are not discernible.

3.2 Space reduction function

The main interest of the space reduction function r is to make calibration easier without loss of generality. In particular, we can make r invariant to contrast by expressing it:

$$r(\bar{E}_1^{max}, \bar{E}_2^{max}, \bar{E}_3^{max}) = s(E_1, E_2)$$

$$\text{with } E_1 = \frac{\bar{E}_2^{max}}{\bar{E}_1^{max}} \quad E_2 = \frac{\bar{E}_3^{max}}{\bar{E}_2^{max}} \quad (7)$$

The \bar{E}^{max} ratios are displayed in Fig. 3 and show quite similar behavior but at a different scale. Thus we can simply choose

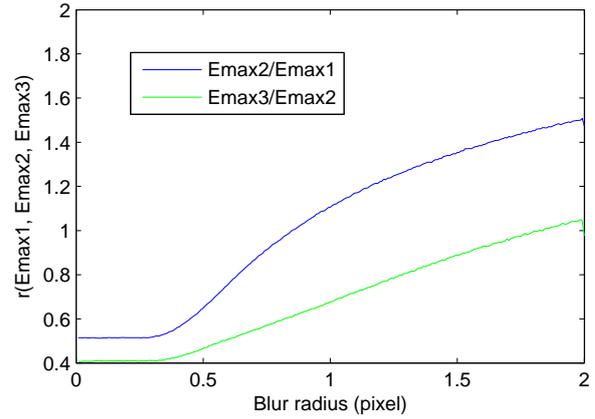


Figure 3: Ratios of \bar{E}^{max} values for white noise images.

the ratio corresponding to our scale of interest, or if we are interested in multiple scales, average the corresponding ratios. If scales larger than 4 pixels are of interest, then we can compute more $\bar{E}_{i+1}^{max} / \bar{E}_i^{max}$. In our case, we are interested in characterizing optical systems, that usually have blurs of radii smaller than one pixel, such that we will simply choose:

$$r(\bar{E}_1^{max}, \bar{E}_2^{max}, \bar{E}_3^{max}) = E_1 = \frac{\bar{E}_2^{max}}{\bar{E}_1^{max}} \quad (8)$$

3.3 Calibration function

As the function r is monotonous in σ in our experiment (except for blur radii below 0.4 pixels that are indiscernible because they reach the Shannon limit of the sensor), this relation ship can be inverted to get σ as a function of r . This is exactly the definition of our calibration function c . Consequently, if we had a ground truth (a perfectly sharp image corresponding to each image acquired by the optical system that we want to characterize), then we could compute $r(\dots)$ for these perfectly sharp images blurred with various blur radii σ , and get our calibration function c as the inverse of this function.

In order to compute calibration functions in real cases, we will exploit the idea that the statistics of natural images are relatively insensitive to scaling. More precisely, given a dataset of images acquired with a given optical system, we will build a dataset of perfectly sharp images by subsampling with a factor greater than the largest expected blur (a factor 4 gives this guarantee in most cases). We will assume that this dataset has statistical properties close to those of the perfect dataset at full scale (that we cannot have), and compute our calibration function on that subsampled dataset.

The aim of the next subsection is to estimate the validity of the assumption that we make in our approach that our datasets have scale invariant statistical properties.

3.4 Sensitivity to scale

The statistical properties of natural images are complex and have been widely studied. Some work seem to show that invariance to scale is only true at lower scales (Huang and Mumford, 1999). Moreover, the dataset that we use to estimate the blur of our optical system is composed of aerial images that have specific properties that might differ from those of natural images in general. Thus we chose to test our assumption that r is relatively invariant to scale, by a simple experiment: we evaluated the calibration functions for an input dataset with various subsampling factors.

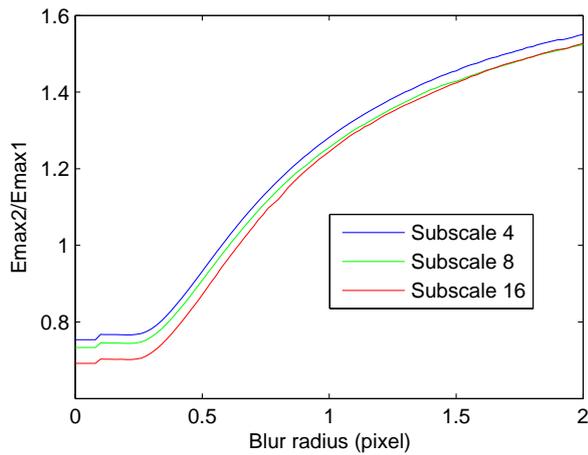


Figure 4: Ratios of \bar{E}^{max} values for subsampled aerial images.

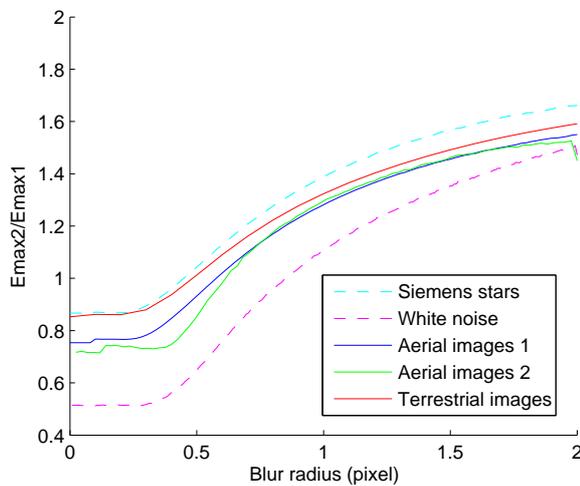


Figure 5: Ratios of \bar{E}^{max} values for synthetic (dashed lines) and real (solid lines) images.

The result of this experiment is displayed in Fig. 4. We see on those curves that the error made by using the calibration function at scale 4 instead of 16 will lead to an error around 0.1 pixel on the blur radius estimation in our zone of interest (0.4 to 1 pixel). This means that we can expect this precision when using the calibration curve at scale 4 to approximate the (unknown) calibration curve at scale 1. This is clearly a limitation to our approach as 0.1 pixel is a rather high error for blurs ranging from 0.4 to 1 pixel. However, we can notice that except in the area below 0.4 pixels where blur cannot be distinguished, the blur radii between different scales are proportional. This means that even if the absolute precision is poor, the relative precision is good, such that our approach is pertinent to locate flaws in the optical system as areas where the blur radius increases. In other terms, the limitation on the precision of the estimation will not impair the interpretation of the result.

A last point of interest is to understand the influence of the image dataset statistics on the calibration curve.

3.5 Sensitivity to image statistics

We have already built the calibration function for noise images and for subsampled aerial images, which have rather different statistical properties. In particular, aerial images present structures

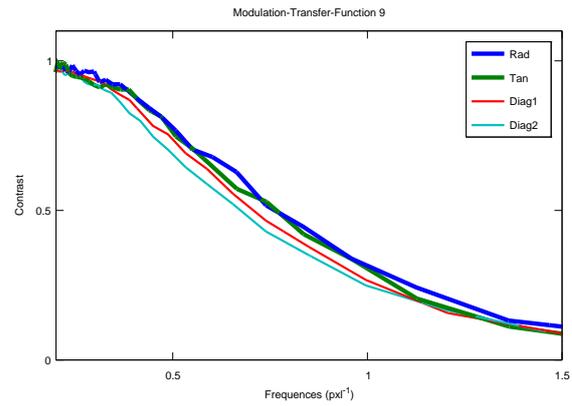


Figure 6: Point spread function at one of the 9 points estimated using a Siemens star.

at various different sizes whereas noise images have mostly structures of sizes close to a pixel. To complete the comparison, we also built calibration functions for a second aerial image dataset, another dataset coming from terrestrial mobile mapping, and a set of synthetic images of packed Siemens stars. The results are displayed in Fig. 5.

We first note that our two synthetic datasets (white noise and Siemens stars) have the most extreme values, one displaying the most irregular structures (white noise), and the other the most regular (Siemens stars). The real datasets are between those extremes: terrestrial imagery in urban areas usually displays large structures, so it is closer to the structured Siemens stars calibration function. Aerial image dataset are quite intermediate, and show similar behavior except around 0.5 pixels. This probably comes from the fact that the second dataset contains more forests which brings more details at very low scales.

In conclusion, the calibration curves are close enough on real images to make visualization of the flaws of the imaging system quite independent of the curve used. For this application, an "average" calibration curve can be used, which saves a lot of computation time. Computing a calibration curve requires to subsample each image of the dataset then compute the E^{max} over each subsampled image, which roughly takes one hour for one thousand images.

3.6 Comparison with blur estimation using Siemens stars

A classical approach to estimate the quality of an optical system is to compute its Modulation Transfer Function (MTF). This can be done by acquiring an image of a Siemens star, then estimating the contrast at various distances of the center (corresponding to a spatial frequency) and in various directions (usually 4). The first aerial dataset that we used contains such a Siemens star that is visible in 9 images, so we applied this procedure to estimate the MTF at 9 different points of the imaging system. The result for one of the points is displayed in Figure 6. We notice that the curves in the various directions are very close, showing isotropy. We estimated a blur radius from these curves for the nine points by computing the contrast as a function of the blur radius and inverting the relationship. The results are displayed in Fig 7. As expected, the error is below 0.1 pixel, and the blur radius is slightly exaggerated by our approach.

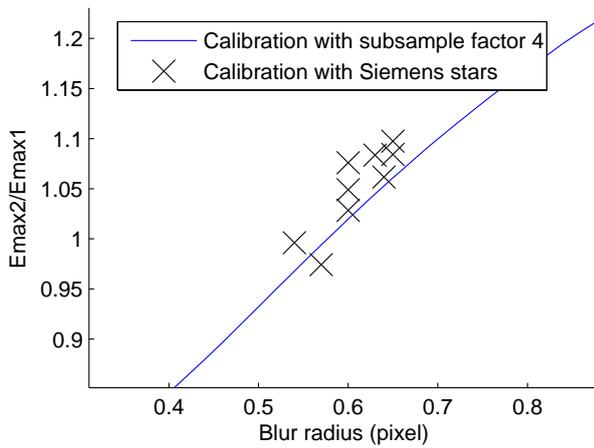


Figure 7: Comparison with blur estimation using Siemens stars: black crosses are at coordinate $(\sigma_{Siemens}, r)$ where $\sigma_{Siemens}$ is the blur estimation with Siemens stars, and $r = \bar{E}_2^{max} / \bar{E}_1^{max}$ is read at the Siemens star center using our approach. The difference between blur estimation using the two approaches is given by horizontal distance from crosses to the calibration curve (blue).

4 RESULTS AND DISCUSSION

Based on the proposed methodology, we are now able to build blur radius maps, with the limitations quoted above, for any optical system for which we have an image dataset satisfying the assumptions of Section 2.3. We will now display the results of this methodology applied to aerial and terrestrial imagery.

4.1 Experiment on aerial imagery

A first set of image is provided from an airborne photogrammetric mission with a DMC. The results of our sharpness estimation for these images are displayed in Figure 8. The small number of images (157) for this mission is compensated by the large resolution of the images (7680x13824) which is a mosaic of four individual panchromatic images. On these images, the top and the bottom areas of the imaging system have a lesser resolution than the center: this could be interpreted as an effect of the deformation (projection) applied to the four images during mosaicking. The second visible artifact is the vertical line in the middle of the figure: the decrease in sharpness in this part of the imaging system may come from the seam between the left and right images. We also notice that some small area (such as the one at the top left) seem more blurry than the average: this might be interpreted as flaw in (or a dust on) the lens or the sensor. Yet it should be noticed that the blur radius is always less than a pixel.

4.2 Experiment on streetside imagery

A second experiment was led on streetside urban images obtained by a camera mounted on a mobile vehicle. The distortion of the camera was corrected prior to applying our method, so the sharpness of the entire pipeline is estimated (Figure 9). One can notice that the grid used for distortion correction is perfectly visible. The sharper area at the bottom of the image is probably due to the fact that our assumption on the homogeneity of edge distribution is not verified as this part of the image always sees the road.

In conclusion, our approach allows us not only to evaluate the quality of an optical system, but also to detect if the image underwent alterations such as interpolations.

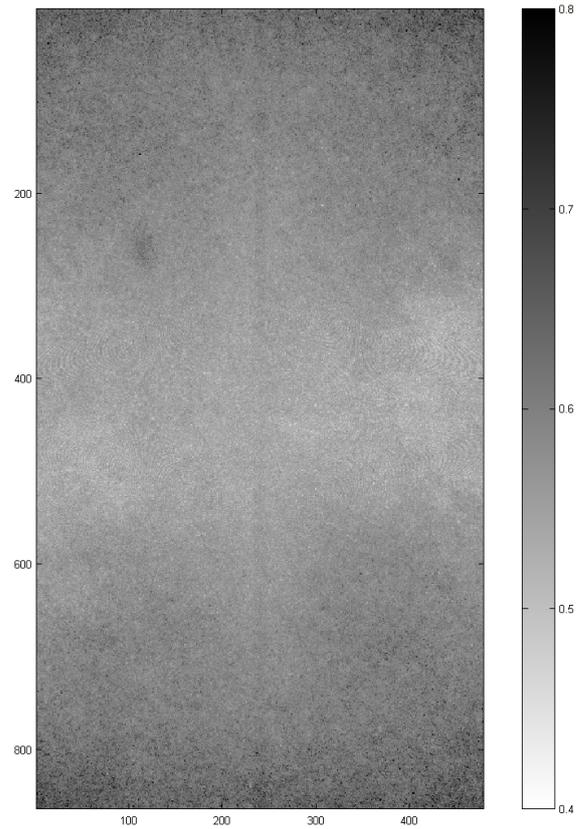


Figure 8: Sharpness image obtained through calibration for the airborne experiment

5 CONCLUSION AND PERSPECTIVES

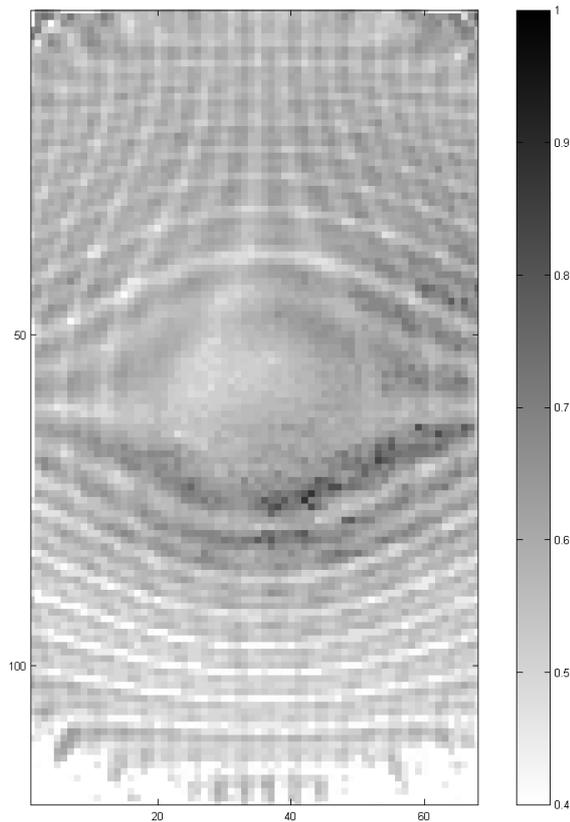
Our work aims at quantifying the amount of blur induced by an optical system from a large set of images that it has acquired. This is extremely useful in an operational context as it avoids immobilizing an expensive resource (an aerial camera) in a lab to perform its evaluation. It can also be used as a complement to lab calibration as manipulations of the imaging system to transfer it from the lab to the plane might slightly alter its characteristics. Finally, we can think of another application of our method that would be to evaluate the stability of the optical quality in time during its utilization, enabling online certification of the imaging system.

We have shown that our approach allows to build a sharpness map of the imaging system, such that our sharpness estimation is much denser than sparse approaches based on Siemens stars for instance. We have shown that our estimation has an absolute accuracy around 0.1 pixel (in blur radius) which is close to what can be achieved based on MTF estimation. But more importantly, we have a very good relative quality which allows for an easy visual inspection of the localization of possible quality artifacts in the image.

The method developed is targeted at a very limited blur radius range of interest, but can be easily extended by using other ratios for dimension reduction, and using more than 3 Haar levels.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Didier Boldo for the original idea that gave birth to this work.



Zhang, W. and Bergholm, F., 1997. Multi-scale blur estimation and edge type classification for scene analysis. In: *International Journal of Computer Vision*, Vol. 24, pp. 219–250.

Figure 9: The sharpness estimate for streetside experiment exhibits the grid used for the correction of the distortion

REFERENCES

- Becker, S., Haala, N., Honkavaara, E. and Markelin, L., 2007. Image restoration for resolution improvement of digital aerial images : A comparison of large format digital cameras. In: SFPT (ed.), *ISPRS Commission Technique I. Symposium*, Marne-la-Vallée, France, pp. 5–10.
- Du, H. and Voss, K., 2004. Effects of point-spread function on calibration and radiometric accuracy of ccd camera. In: *Applied Optics*, Vol. 43number 3, pp. 665–670.
- Fleury, P. and Mathieu, J., 1956. Chapitre: 18 - photographie projection. In: Eyrolle (ed.), *Image Optique*, Paris, France, pp. 413–432.
- Huang, J. and Mumford, D., 1999. Statistics of natural images and models. In: *IEE Conf. on Computer Vision and Pattern Recognition*, pp. 541–547.
- Kedzierski, M., 2008. Precise determination of fisheye lens resolution. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 37, Beijing, China, pp. 761–764.
- Liu, R., Li, Z. and Jia, J., 2008. Image partial blur detection and classification. In: *IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, Alaska, pp. 1–8.
- Luxen, M. and Forstner, W., 2002. Characterizing image quality: Blind estimation of the point spread function from a single image. In: *Photogrammetric Computer Vision*, Graz, Austria, pp. 211–217.
- Tong, H., Li, M., Zhang, H. and Zhang, C., 2004. Blur detection for digital images using wavelet transform. In: *IEEE International Conference on Multimedia and Expo*, Vol. 1, Taipei, Taiwan, pp. 17–20.