OBJECTIVE ANALYSIS OF ILLUMINATION NOISE IN FLUORESCENCE MICROSCOPY

Miguel Coimbra[1], João Paulo Silva Cunha[1], Ralf Bausinger[2], Christoph Bräuchle[2], Andreas Zumbusch[2]

[1]IEETA, Universidade de Aveiro
Departamento Electrónica e Telecomunicações
Portugal.

[2]Ludwig-Maximilians Universität,
Department Chemie und Biochemie,
Munich, Germany.

ABSTRACT

Single particle fluorescence microscopy poses specific computer vision problems that need to be addressed in detail. Illumination noise, resulting from the light fluorescence characteristics of observed samples, hinders the performance of traditional segmentation techniques such as brightness thresholding and block matching shape analysis. This paper studies the characteristics of illumination noise and its effect on these segmentation techniques.

1. INTRODUCTION

A recent breakthrough in biomedical science is the capacity to track individual particles such as viruses using fluorescence microscopy. This technique provides unprecedented detailed insight into biological problems that is impossible to obtain with traditional ensemble averaging approaches. It consists in labeling a single particle with one or very few fluorescent dye molecules, enabling individual particle tracking with high spatial and temporal resolution using fluorescence microscopy. This technique has been used by Seisenberger [1] to study infection pathways of adeno-associated viruses, Babcock [2] to study nuclear trafficking of viral genes and by Lakadamyali [3] for visualization of the infection of influenza particles.

The new single particle tracking techniques are generating large quantities of video data that need to be processed. Current systems either depend on full manual segmentation or manual verification and correction of automatic segmentation and tracking results. These semi-automatic systems rely on traditional image processing techniques (e.g. detection of local maximum followed by shape analysis [2][3]) that are not necessarily adequate for this imaging system. There are previous examples in literature (Gravel [4]) where the importance of studying the noise characteristics of medical image systems, instead of just assuming them, is shown. Results on this paper will show that single particle tracking using fluorescence microscopy has specific characteristics that automatic video processing applications should be aware of.

Illumination noise is analyzed in Section 2 and target shape in Section 3. Segmentation results are detailed in Section 4 and final conclusions drawn in Section 5.

2. ILLUMINATION NOISE

A major source of noise in fluorescence microscopy is that, although specific targets are labeled with highly fluorescent markers, the whole sample is typically lightly fluorescent, creating a global illumination pattern that we call *illumination noise*. This noise can reduce our target-background contrast significantly making it difficult to detect and segment individual particles in some situations.

Illumination noise is not spatially homogenous. In fact, more noise seems to be present in the center of the image than in the edges. We measured each pixel's mean gray value μ_p (1) and variance Vp (2) where $f(x,y)$ is the pixel value for frame t from a total of N frames.

$$\mu_p(x,y) = \frac{1}{N}\sum_t f(x,y) \quad (1)$$

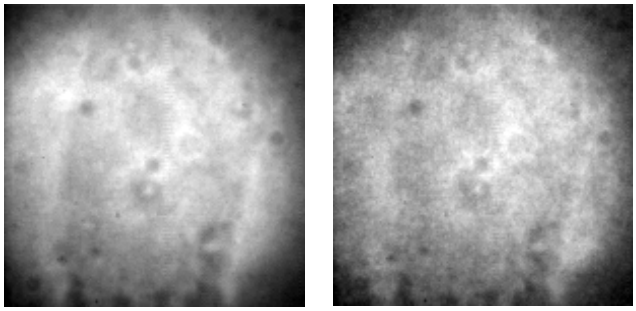
$$V_p(x,y) = \frac{1}{N}\sum_t \left[f(x,y) - \mu_p(x,y)\right]^2 \quad (2)$$

All experiments were done using a Pentium 3.2 Ghz and software created with Visual C++. A sample of 400 images from different sequences was selected from a larger set. The criterion for this choice was to choose sequences with either one or no particles per image (up to 50 particles per image can be obtained) so that a high percentage of screen area is the noise we want to measure. Images are 8-bit gray level with 151x151 resolution. They correspond to temporal resolutions from 35 f/sec to 1 f/min and spatial resolutions of 85x85 nm2 per pixel. All particles were manually annotated in the images, and a small area (15x15) around them was not used in the noise studies.

Each pixel's mean gray value μ_p (1) and variance Vp (2) were obtained using all images and results can be seen

on Figure 1 (images were contrast stretched for visibility). It is noticeable that both the mean and the variance of the noise seem to be stronger near the center of the images. This is confirmed by analysis of Figure 2, where the average values of μ_p are plotted against their distance to the center.

We can also analyze the temporal homogeneity of illumination noise by observing the variation of the mean gray level of each image over time. These variations are quite drastic (approximately between gray level 20 and 150) and its distribution is plotted in Figure 3.



a) μ_p b) Vp

Figure 1 – Spatial distribution of illumination noise.

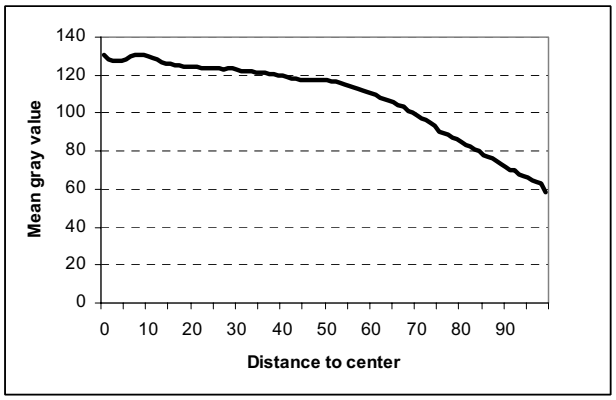


Figure 2 – Influence of pixel position on illum. noise.

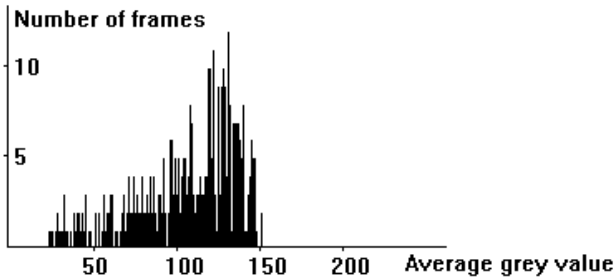


Figure 3 – Temporal variation of illumination noise.

3. TARGET SHAPE ANALYSIS

An analysis of the local characteristics of the particle detection is also required. It is important to measure the size and shape stability of this detection, allowing a better understanding of the segmentation task. This was obtained by manually annotating all occurrences of the particle in the available frames, followed by the extraction of local characteristics. An 'average shape' was obtained for a 15x15 block, centered on the manual detection pixel, for all frames available. This was compared to random shapes obtained by selecting random points on the image, as well as with each individual particle annotation, using typical block-matching cost functions (Kuhn [5]), such as mean square error (3) and mean amplitude error (4).

$$MSE = \frac{1}{15x15} \sum_{x=-7}^{7} \sum_{y=-7}^{7} [f(x,y) - \mu_{virus}(x,y)]^2 \quad (3)$$

$$MAE = \frac{1}{15x15} \sum_{x=-7}^{7} \sum_{y=-7}^{7} |f(x,y) - \mu_{virus}(x,y)| \quad (4)$$

	Original		Normalized	
	MAE	MSE	MAE	MSE
Av.err	29,9	1361,8	36,3	1437,0
Rand.	27,9	1501,7	82,85	9380,6

Table 1 – Target shape similarity for original and contrast stretched images (*Normalized*).

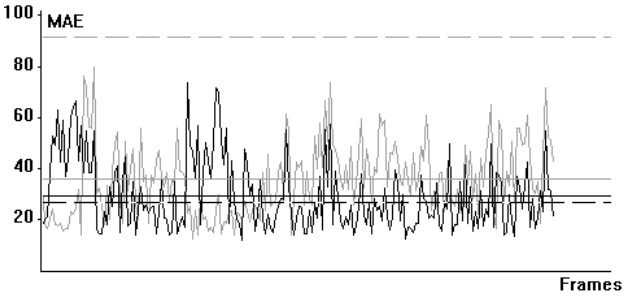


Figure 4 – MAE errors for orig. (black) and norm. blocks (light gray). Dotted lines show random block error.

Results in Table 1 show that attempting to detect particles using block matching techniques directly on the image will probably fail since we are analyzing the target shape corrupted by illumination noise. MAE errors for particle blocks (average error) and random blocks are almost the same. We are interested in the target's shape and not in its brightness (unreliable due to illumination noise), so our matching can be improved by previously applying contrast stretching (*normalized*) to the compared blocks. Figure 4 confirms this by showing that differences between particle and random blocks are now significant.

4. SEGMENTATION RESULTS

Two basic segmentation tools (brightness threshold, block matching) were selected since they are probably the simplest methods of two different approaches to the segmentation task: direct brightness (or color) and shape

analysis. Also, video samples with one or zero particles totaling 500 frames were selected, reflecting a variety of noise situations (weak, medium and strong noise). The performance of these segmentation algorithms was tested using various error measures: accuracy (5), recall (6), average error distance (7), and total error distance (8). A correct detection (*corr.det.*) means that the algorithm correctly detected that there is one particle $p_S(x_S,y_S)$ in the image while total detections is the number of manually annotated $p_{MA}(x_{MA},y_{MA})$ particles in all images. Total pixels (*tot.pixels*) are the number of pixels $p_k(x_k,y_k)$ in all frames that are above or below a threshold *k*.

$$\text{Accuracy} = \frac{\text{correct detections}}{\text{correct detections} + \text{incorrect detections}} \quad (5)$$

$$\text{Recall} = \frac{\text{correct detections}}{\text{total detections}} \quad (6)$$

$$Av.Error.Dist = \frac{1}{\text{corr. det.}} \sum_{n=1}^{\text{corr.det}} \sqrt{\left(x_{MA}(n) - x_S(n)\right)^2 + \left(y_{MA}(n) - y_S(n)\right)^2} \quad (7)$$

$$Tot.Error.Dist = \frac{1}{\text{tot. pixels}} \sum_{i=1}^{\text{tot.pixels}} \sqrt{\left(x_{MA}(i) - x_K(i)\right)^2 + \left(y_{MA}(i) - y_K(i)\right)^2} \quad (8)$$

Table 2 lists some results obtained for brightness thresholding segmentation (*Gray*), block matching using different error measures (*MAE, MSE*) and either without (*orig*) or with (*norm*) illumination normalization.

	Acc.	Rec.	Av.Err	Tot.Err
Gray	59%	100%	19,3	
Orig. MAE	59%	100%	47,1	
Norm. MAE	59%	100%	37,7	
Norm. MSE	59%	100%	39,5	
Gray (k = 200)	59%	100%	19,3	46,3
Gray (k = 250)	59%	100%	19,3	18,6
N.MAE (k=30)	94%	52%	29,4	39,7
N.MAE (k=40)	72%	83%	33,0	45,8
N.MSE (k=2000)	86%	66%	27,9	40,9
N.MSE (k=2500)	78%	80%	31,5	42,3

Table 2 – Segmentation errors (*Acc.* – Accuracy; *Rec.* – Recall; *Av.Err* – Av.Error.Dist; *Tot.Err* – Tot.Error.Dist)

Lines 1-4 measure the unlikely case where we know that there is one target in the image and we just want to find it. *Gray* clearly obtains smaller average distance errors and while block-matching results are worse, they are improved significantly by the normalization step, which is not surprising given Table 1 results. However, when we attempt to improve our accuracy by using thresholds thus classifying images as having either one or zero targets, only block-matching algorithms improve their performance (higher accuracy, lower average distance errors) although they don't have perfect recall (some targets are not detected). No significant difference is noticeable between using MAE and MSE error measures.

A typical limitation of this type of methods is that they rely heavily on what are usually called "magic numbers". The choice of threshold value is usually empirical and may lead to instable behaviors when test conditions change. It is therefore important to measure the sensitivity of the system to variations in threshold values. One way to do this is to plot our chosen error measures against various threshold values (Figure 5).

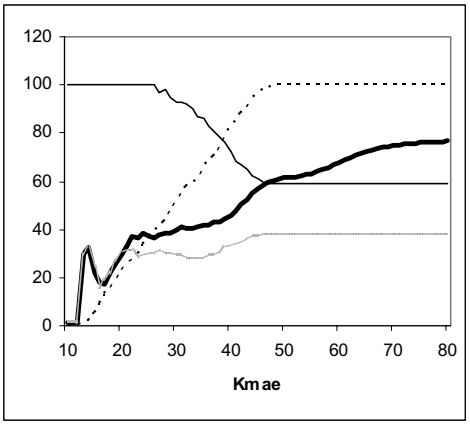


Figure 5.a) – Variation of errors using *MAE*

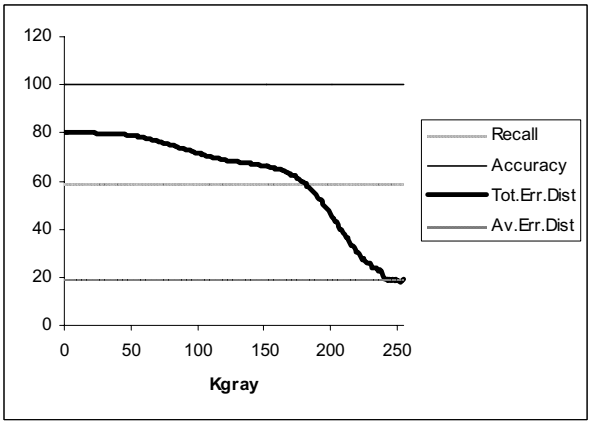


Figure 5.b) – Variation of errors using *Gray*

Figure 5 allows some interesting observations. As expected, we have an accuracy/recall tradeoff in MAE, obtaining a common 75% at around k_{MAE} = 35. 100% recall is only obtained when all images (and thus 59% accuracy) are selected. The average distance error is relatively stable and independent of k_{MAE} for recall values above zero. The most striking results is the apparent independence of *Gray* results from k_{gray}. This is not true since the total error distance slowly decreases with *k* until it converges on the average error distance line. The imaging system uses some form of auto-gain mechanism that makes simple thresholding unable to distinguish between a strong signal or amplified illumination noise.

It is common signal processing practice, especially when the signal is an image, to apply low-pass filtering to remove spot noise. This type of filtering might not be

adequate for certain situations since the signal might contain significant high-frequency information that we want to extract (i.e. sharp edges). The same segmentation procedure that was used to obtain the results from Table 2 was used, but now the images were previously filtered using various low-pass filters.

		Gray (k = 200)		MAE (k = 40)	
Type	Size	A/R (%)	Av.Er	A/R (%)	Av. Er
No filter		59/100	19,3	72/83	33,0
Mean	3	64/99	14,8	67/90	40,0
Mean	5	60/63	10,1	65/87	40,8
Mean	7	55/31	13,8	74/71	36,0
Med.	3	64/99	14,7	69/88	37,5
Med.	5	60/64	9,2	66/92	37,3
Gaus.	3	62/99	15,1	66/90	39,9
Gaus.	5	65/97	14,4	65/92	41,0

Table 3 – Low-pass filtering results (A – Accuracy; R – Recall; Av.Er. – Av.Error.Dist)

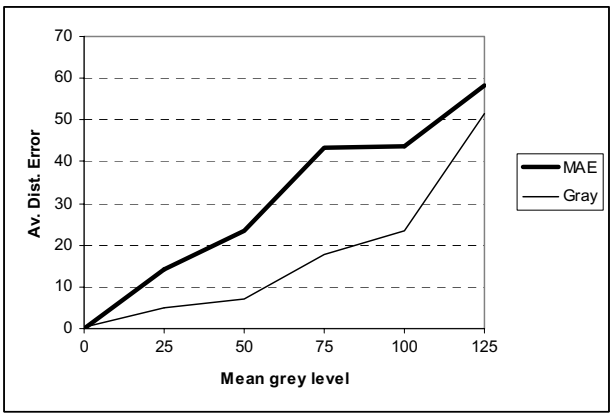


Figure 6 – Influence of the temp. heterogeneity of noise

Results presented in Table 3 show that low-pass filtering is not necessarily adequate for all segmentation methods. The performance of the *Gray* algorithm seems to improve: higher accuracy and lower average distance errors are generally obtained. Mean or median filters of size 5 obtain the lowest *Av.Err.* but at the expense of a 40% drop in recall. There is no apparent difference between filters of size 3. On the other hand, the performance of block-matching algorithms seems to degrade (lower accuracy, higher *Av.Err.*), hinting that high-frequency components are quite important for such shape matching methods. Not surprisingly, best results are obtained by the median filter, due to its edge preserving characteristics.

Finally, we've tested how the temporal heterogeneity of illumination noise affects segmentation results. The behavior of the average distance error with the average gray level of each frame was measured for two situations: Gray and norm. MAE results (Figure 6) show a clear increase of segmentation errors for high mean gray levels, confirming the suspicions of the noise studies of Section 2. The results measuring the relevance of the distance to the image center were inconclusive.

5. DISCUSSION

This paper has shown that, even using a relatively small data set, some important characteristics of the imaging system used for single particle tracking have been highlighted. Tests hint at the spatial heterogeneity of noise (Figures 1 and 2), stronger in the center of the image, as well as temporal heterogeneity (Figure 3) when illumination noise is suddenly amplified, probably by some form of auto-gain mechanism of the imaging system. Target shape is affected by illumination noise but using a normalization step improves results (Table 1, Figure 4).

Simple segmentation methods have been tested and their performance measured using various error measures. Results confirm the negative influence of illumination noise on low-level image processing techniques, especially the effect of its temporal heterogeneity (Figure 6). Other observations also show the danger of relying on fixed thresholds (by plotting error variation with threshold value - Figure 5) as well as the positive and negative effect of low-pass filtering (Table 3).

In the future we intend to tighten the cooperation between computer vision and biomedical science, by improving single particle imaging systems based on noise and low-level segmentation analysis. A much larger data set is needed where such data should ideally be obtained by different imaging systems. A parallel task is to develop robust automatic segmentation and tracking tools for such systems, based on the knowledge gained by the experiments here presented.

REFERENCES

[1] Seisenberger G., Ried M.U., Endress T., Büning H., Hallek M., Bräuchle C., "Real-time single molecule imaging of the infection pathway of an adeno-associated virus", in Science, 294 (2001) pp. 1929-1932.

[2] H.P. Babcock, C. Chen, and X. Zhuang, "Using single-particle tracking to study nuclear trafficking of viral genes", in Biophysical Journal, vol.87, 2004, pp. 2749-2758.

[3] M. Lakadamyali, M.J. Rust, H.P. Babcock, and X. Zhuang, "Visualizing infection of individual influenza viruses", in PNAS, vol.100, no.16, 2003, pp. 9280-9285.

[4] P. Gravel, G. Beaudoin, and J.A. Guise, "A method for modeling noise in medical images", in IEEE Trans. On Medical Imaging, vol.23, no.10, 2004, pp.1221-1232.

[5] P. Kuhn et al., "Complexity and PSNR-comparison of several fast motion estimation algorithms for MPEG-4", in Proc. of SPIE, Application of Digital Image Processing XXI, vol. 3460, 1998, pp. 486-499.