1. Introduction

Confocal Scanning Laser Microscopy (CSLM) represents one of the most important advances in optical microscopy of the last decades. It is widely accepted that the confocal microscope was invented by Marvin Minsky, who filed a patent in 1957 (Minsky, 1957). However, at that time such a system was very difficult, if not impossible, to implement, due to the unavailability of required laser sources, sensitive photomultipliers or computer image storage possibilities. A laser scanning microscope using mechanical object scanning was developed in Oxford in 1975, and a review of this work was later published (Sheppard, 1990). The Oxford microscope was the first commercial confocal microscope. Other important contributors to this era of the development of confocal microscopy were Brakenhoff (Brakenhoff et al., 1979) and Cox (Cox, 1984).

The architecture of a CSLM system provides the possibility to acquire images representing optical sections on a sample’s volume. In order to achieve this, in a CSLM system an excitation source emits coherent light (laser) which is scanned across the sample surface. As it reaches the sample the light is reflected towards a detector, in reflection work mode, the same optical path being used as well in fluorescence work mode. While in conventional microscopy, the detector is subjected to light which is reflected by out of focus planes, resulting in out-of-focus blur being contained in the final image, the architecture of a CSLM system helps avoid this situation. In order to acquire images corresponding to certain optical sections, a confocal aperture (usually known as pinhole) is situated in front of the detector. More precisely, the pinhole is placed in a plane conjugate to the intermediate image plane and, thus, to the object plane of the microscope. As a result, only light reflected from the focal plane reaches the detector, out-of-focus light being blocked by the pinhole (Fig. 1). The dimension of the pinhole is variable and together with the wavelength which is being used and the numerical aperture of the objective, dictates the thickness of the volume which contributes to the collected image (Shepard et. al., 1997; Wilson, 2001).

In the case of CSLM systems, the detector is a photo multiplier tube (PMT), which presents a wide dynamic range and has high photon sensitivity suitable for detecting both strong and weak signal at a very quick refresh rates, in a time range of nano-seconds. The PMT detects light and converts photon hits into analogue electron flow as electrons leave the
photocathode of the PMT, having the energy of the incoming photon. After the electrons follow a path which amplifies their number, they reach the anode of the PMT, where the accumulation of charge results in a sharp current pulse indicating the arrival of a photon at the photocathode. The continuous analogue current signal is then sampled at separate time point, digitized into discrete digital signal by analogue to digital converter (ADC), then processed by image processor resulting in digital images of the sample area contained in the system’s field of view.

Besides CSLM specific advantages such as increased resolution and better contrast, the provided possibility of achieving images corresponding to optical sections represents as well a significant advantage to people working in fields such as biology, medicine, material science or microelectronics, as CSLM image stacks can be used for 3D reconstructions of the studied sample (Rigaut et al., 1991, Liu et al. 1997, Pironon, 1998, Rodriguez et al, 2003, Sugawara et al., 2005).

![Principle of Confocal Scanning Laser Microscopy](image_url)

**Fig. 1. Principle of Confocal Scanning Laser Microscopy**

In some circumstances, a CSLM image corresponding to an optical section may contain defocused, low contrast or over-saturated areas. This problem can be present due to various reasons, such as region non-uniformity or sample regions which contribute to the image not being in the same focal plane at the same time due to non-uniform size or sample tilt (Fig. 2). For certain types of investigations conclusions can be drawn only based on images of uniform quality or uniform focus. These types of images allow better morphological observations of the sample details. One method for obtaining this type of representation is image fusion. Image fusion will provide an artificial image, which will consist of image...
regions belonging to different images of the CSLM stack. The purpose of this operation is to achieve an image representing a better description of the imaged scene or object than any of the individual source images. Ideally, the fusion algorithm should preserve relevant information from the fused images and suppress image regions or components which are subjected to noise or which are irrelevant in respect to a defined purpose (Nikolov, 98). Applications of image fusion have been implemented with great success in different microscopy and medical applications. For example, excellent results have been achieved in the case of three-dimensional microscopy, where certain limitations imposed by the low axial resolution of the system have been overcome by fusing images acquired at different placements of the sample (Swoger et al., 2007), while in (Forster et al., 2004) wavelet based image fusion is presented as a solution which provides good results for extending the depth of field in the case of multichannel microscopy images. In (Chen, et al. 2010) an image fusion algorithm based on bidimensional empirical mode decomposition (BEMD) is applied to multi-focus color microscopic images achieving a balanced result between local feature enhancement and global tonality rendition.

Fig. 2. Possible scenario for the acquisition of a CSLM images of non uniform focus

Image fusion can be performed in both frequency and spatial domains. Our approach, which deals with the fusion of CSLM images, was developed on a region level basis. Lately, much attention has been focused towards region-based image fusion because of its perceived advantages. With fusion rules based on combining regions instead of pixels, more useful tests for choosing the adequate regions from the source images, based on various properties of a region, can be implemented prior to fusion. Problems such as sensitivity to noise, blurring effects and misregistration in the case of pixel-fusion techniques, can be overcome by processing semantic regions rather than individual pixels (Li & Yang, 2008).

In the case of the four fusion methods that we experiment, each image in the CSLM stack is divided into the same number of square regions, same as in (Huang & Jing, 2007). Two of the proposed methods are based on a focus assessment operator, while the other two are based on a quality assessment operator. In the first two methods, which we entitled FFMAX and FFAVG, a focus assessment for the same region in all the images in the stack is calculated by Tenenbaum’s algorithm (Tenengrad). In FFMAX the region of the best focus is chosen to appear in the fused image, as in (Huang & Jing, 2007), while in FFAVG, instead of building the fused image from blocks which belong to a single image, we build it by mean averaging the blocks of all images in the stack, the contribution of each source image...
to the fused image being proportional to its response to the focus measure, as in (Stanciu et al., submitted 2010). The same approach is used in the 3rd and 4th presented methods, QFMAX and QFAVG, where instead of using a focus assessment operator as a decision criterion which dictates the inclusions of an image region to the fused image, a quality assessment operator is used. All four methods aim to obtain a fused image of better focus or quality uniformity, with morphological details of the structure being more visible than in any other image that contributed to the fusion.

2. Fusion methods

2.1 Focus assessment

The first two fusion methods we have experimented, FFMAX and FFAVG, are based on an image clarity measure, namely a focus measure. A well-focused image has the best average focus over an entire field of view, even though objects often reside at multiple focus planes in thick sample slides. In the case of the first two experimented methods the focus assessment dictates the inclusion or the contribution of an image region to the final fused image. Once an image of the stack is divided into blocks of a certain size, for all these blocks a focus measure is calculated. Focus measures have been deeply studied in the field of autofocusing. There are two kinds of focus measures, spatial domain focus measures and frequency domain focus measures. However, frequency domain focus measures will not be used in a real-time system because of their complexity. Detailed discussions on the topics of focus measures and auto-focusing can be found in the literature (Nayar & Nakagawa, 1994; Subbaro et al., 1992; Yeo et al., 1993; Geusebroek et al., 2000). In (Huang & Jing, 2007), several focus measures were compared according to the focus measure’s capability of distinguishing clear image blocks from blurred image blocks. The results of the experiments performed on natural images showed that the Sum-Modified Laplacian (SML) can provide better performance than other focus measures when the execution time is not included in the evaluation, but other measures such as Energy of Laplacian of the image, Tenenbaum’s algorithm or Energy of image gradient provided good results as well. In (Osibote et al., 2010), a comparison of automated focusing methods for brightfield microscopy was conducted. It was showed that Vollath’s F4 algorithm provided best results, but in the same time Brenner and Tenenbaum’s algorithm provided very good results as well. For estimating the focus of a certain region we use a spatial domain focus measure, Tenenbaum’s algorithm (Tenengrad) (Krotkov, 1897; Yeo et al., 1993), which is a gradient magnitude maximization method that measures the sum of the squared responses of the horizontal and vertical Sobel masks. In its original implementation, the summation is for pixels that are above a certain threshold; however, we chose to use a variation in which all pixel locations can be included in the summation (Santos et al., 1997).

\[
\text{Tenengrad} = \sum_{x=2}^{M-1} \sum_{y=2}^{N-1} |\nabla S(x,y)|^2
\]  

(1)

Where is the Sobel gradient magnitude given by:

\[
\nabla S(x,y) = \sqrt{\nabla S_x(x,y)^2 + \nabla S_y(x,y)^2}
\]  

(2)

where \(\nabla S(x,y)\) is
\[ \nabla S_x(x,y) = \{-[f(x-1,y-1) + 2f(x-1,y) + f(x-1,y+1)] + [f(x+1,y-1) + 2f(x+1,y) + f(x+1,y+1)] \} \]

\[ \nabla S_y(x,y) = \{+[f(x-1,y-1) + 2f(x,y-1) + f(x+1,y-1)] + [f(x-1,y+1) + 2f(x,y+1) + f(x+1,y+1)] \} \]

2.2 Quality assessment

For estimating the quality of image regions we have chosen to use the same quality metric as defined in (Stanciu et al, 2010): \[ q_f = \mu_f \sigma_f \mu_g, \]
where \( \mu_f \) is the average grey level of the image, \( \sigma_f \) is the standard deviation of the image pixels, and \( \mu_g \) is the mean intensity of the gradient image. We have chosen this quality factor, as it takes into consideration three important aspects which define quality when referring to CSLM images: image brightness, image contrast and presence of edges and boundaries.

A good measure of image brightness is the average gray level of the image. Let us consider the analyzed square region as a discrete image \( f:[0,M-1] \times [0,N-1] \rightarrow [0,L-1] \) and let \( H = \{h(0), h(1), \ldots, h(L-1)\} \) be its histogram. The average gray level, \( \mu_f \), immediately follows:

\[ \mu_f = \frac{1}{MN} \sum_{i=0}^{L-1} h(i) \]  

(4)

The standard deviation can be regarded as a measure of image contrast. This can be easily understood since \( \sigma \) is a measure of how widely spread the values in a data set are. An unbiased estimate of the standard deviation is:

\[ \sigma_f = \sqrt{\frac{1}{MN-1} \sum_{i=0}^{L-1} h(i)(i - \mu_f)^2} \]  

(5)

Another important factor that we take into consideration when choosing the reference image is related to the edges contained in the image. If an image is of good quality we can discern very clearly the objects contained in it. Edges characterize boundaries and are therefore a problem of fundamental importance in image processing. Edges represent discontinuities between image regions of rather uniform graylevel or color. In a fashion similar to the focus assessment method described in 2.1, we considered the Sobel edge detector (Eq. 2), where \( S_x \) estimates the gradient in the x-direction (columns), while \( S_y \) estimates the gradient in the y-direction (rows). We consider the measure of the edges contained in image \( f \) as the mean intensity of its gradient magnitude image, namely \( \mu_g \).

2.3 Fusion of square regions

In all four methods which we have experimented each of the images in the stack is divided into a set of square regions. The dimension of the square regions can be chosen according to the content of the images that are to be fused. Higher region size is equivalent to less discriminative power between image areas, while a lower region size will bring a larger number of disturbing artifacts at the boundaries of the fused regions, also known as seams.
The computational time is also directly linked to the size of the square region, larger regions being equivalent to faster processing time. Because of these aspects, a compromise should be made when choosing the size of the square region. Usually for an image of 1024 x 1024 pixels we have obtained best results for square regions of 32 and 64 pixels.

FFMAX, the first method we experimented, is similar to the methods described in (Huang & Jing, 2007), for each square region, its inclusion in the fused image is decided by calculating its response to the Tenengrad operator in all the images in the stack (source images). The block with the maximum response to the Tenengrad operator will be included in the fused image (Fig. 3), while others will be discarded. A similar approach is used for the QFMAX method (Stanciu et al, 2009), where instead of deciding a region’s inclusion into the fused image based on a focus assessment operator, we use the quality estimate defined in section 2.2. For both FFMAX and QFMAX methods, a decision operator (F) is calculated for each square region in all images. The number of the image in the stack which contains the region of maximum response to the decision operator is introduced into the correspondence matrix. Once the correspondence matrix is completed, a fused image is constructed as presented in Fig. 3. In the case of FFMAX the decision operator is represented by the Tenengrad focus assessment operator, while in QFMAX it is represented by the quality estimate operator described in 2.2.

![Image 1](image1.png)

Fig. 3. Image fusion process for FFMAX and QFMAX methods

Further on, we propose two other methods FFAVG and QFAVG. In these methods each of the source images will contribute to the fused image in a certain proportion. The contribution of a square region belonging to a source image to the value of the correspondent square region in the fused image is proportional to its responses to the decision operator. Hence, the responses to the decision operator (F) represent weights in a weighted mean based image fusion process (Fig 4). In FFAVG the decision operator is represented by Tenengrad focus assessment operator, while in QFAVG it is represented by the quality estimate operator described in 2.2.

While in FFMAX and FFAVG only information regarding gradient magnitude is used in the decision regarding a region’s inclusion or contribution to the fused image, in QFMAX and QFAVG along with information regarding gradient magnitude, estimates on image brightness and image contrast contribute as well to the decision criterion, which results in brighter resulted images.
3. Objectives and results

3.1 Objective of the technical work carried out

The PQR ‘mesa’ lasers are three-dimensional (3D) whispering gallery (WG) mode lasers with doughnut type Laguerre-Gaussian (LG) beam patterns (Ahn et al., 1999). During our investigations (Stanciu et al., 2008) on this type of devices several aspects related to their geometry could not be resolved from the original images obtained by CSLM as it was not possible to have all the regions of the device structure in focus at the same time. In Figure 5 we present a stack of images obtained on PQR devices by CSLM. The number in the top left corner depicts the numerical order of optical sections in the full series. The stack consisted of 50 CSLM images acquired at different levels along the Z axis. In order to enhance the results of our investigations on PQR devices, it was needed to construct an artificial image, constructed based on the CSLM set, that would contain information from different focal planes (thus from different images corresponding to different optical sections). The four image fusion algorithms presented in 2.3 have been experimented as a solution for this problem.

The CSLM system that was used is a Leica TCS SP. The images of the PQR structures were obtained by scanning a HeNe laser beam (633nm). The power of the laser beam on the sample surface was kept at 10μW. The objective that was used was HC PL FLUOTAR 20.0 X, with a numerical aperture of 0.5.

By looking at the images in the stack (Fig. 5), we can observe that each one contains different details with their origin in different optical sections; to be more precise there are images with more details from the top of the structure (see images 25,31), and others with more details from the background (see images 37, 43).

For our investigations on the PQR devices Laser Beam Induced Current (LBIC) investigations were conducted in order to study the distribution of photocurrents density when illuminating the PQR structure with a laser beam, Fig. 6. In order to establish the relationships which occur between the sample’s geometry and the photocurrent distribution comparisons over CSLM and LBIC representations of the devices were conducted.
Our investigations on PQR devices relied on obtaining images at the same Z levels, for both induced photocurrent map obtained in Laser Beam Induced Current (LBIC) mode, and for the reflection signal of the PQR structure collected by CSLM, in order to determine the region where the current was generated. Due to the device geometry and the limitations of the investigation technique (limited depth of field and limited axial resolution), we had obtained images of generated current, by LBIC, even for Z levels for which there was no image in the reflection workmode of the CSLM, due to the structure slope. In this situations we were not able to link the physical regions of the device to the regions in which photocurrent was present. A solution to this problem was represented by the fusion of images of the CSLM stack acquired on a PQR device as described in 2.3. Comparing the photocurrent image acquired in LBIC to an artificial image containing details from different optical sections (the image resulted after image fusion) had enhanced our understanding of the phenomena which takes place in studied devices.
3.2 Results

In Fig 7 we illustrate the fused images constructed based on the CSLM stack presented in Fig. 5, by using the FFMAX method.

![Example images showing fused images](image_url)

Fig. 7. Resulted images achieved by FFMAX method. The dimension of the square fused regions: a) 16 pixels, b) 32 pixels, c) 64 pixels, d) 128 pixels

The fused images consist of blocks belonging to different images in the stack according to the calculated correspondence matrix (Table 1). Each number in the correspondence matrix represents the number of the image in the stack which contributes with the respective region to the fused image.

In Fig. 8 we illustrate fused images obtained by using the QFMAX method, while in Table 2 one of the correspondences matrix resulted after a quality estimation of square regions of the initial images is presented. Both FFMAX and QFMAX methods provide fused images with better focus or quality uniformity than the initial images in the stack. However, the images provided by both methods contain a large number of artefacts around the borders of the fused square regions. This problems are attenuated in the second group of methods, FFAVG and QFAVG, which are based on an averaging of the square regions having the response to a decision operator as weight.
Table 1. Correspondence matrix for FFMAX method when division into 64 pixel square regions was considered.

Fig. 8. Resulted images achieved by QFMAX method. The dimension of the square fused regions: a) 16 pixels, b) 32 pixels, c) 64 pixels, d) 128 pixels.
Table 2. Correspondence matrix for QFMAX method when division into 64 pixel square regions was considered

In Fig. 9 and Fig 10 we illustrate the results obtained when fusing images by the FFAVG and QFAVG methods. Artefacts around square fused regions are still visible but their intensity is diminished. However this advantage offered by averaging comes at the expense of losing image sharpness, thus a compromise between image sharpness and intensity of border artefacts must be considered when choosing to use one of the experimented methods.

Fig. 9. Resulted images achieved by FFAVG method. The dimension of the square fused regions: a) 16 pixels, b) 32 pixels, c) 64 pixels, d) 128 pixels
4. Conclusion

In this chapter we describe the results obtained by experimenting four region based fusion methods on CSLM image stacks. The four methods have been experimented on image stacks collected on PQR devices. The presented results highlight some of the advantages and limitations of the experimented image fusion methods in connection to CSLM imaging. In the case of CSLM when different regions of the investigated area are not in focus at the same time, the collected image will contain both sharp and bright areas corresponding to the regions in focus, but also blurred areas of low contrast or over saturated areas corresponding to the sample regions which contribute to the image but are not in focus. By the experimented image fusion methods we have obtained representations of the investigated sample constituted from image regions belonging to different images in the CSLM stack, corresponding to different optical sections, thus containing details from various focal planes. In two of the experimented methods, FFMAX and FFAVG, the fused image consists of image blocks of a fixed size which have been extracted or calculated from various images in the stack based on the response to the Tenengrad focus assessment operator, while in the remaining methods, QFMAX and QFAVG the contribution of an image in the stack to the fused image is decided based on a quality estimate of the square regions. In two of the methods, FFMAX and QFMAX the fused image consists of blocks which provide a maximum...
response to a decision operator, while in the other two, QFAVG and FFAVG all square regions from the images in the stack contribute to the fused image proportional to their response to the decision operator. In the case of our experiments both types of methods provided artificial images which had enabled us to have a better estimate on the morphology of the studied sample in the purpose of correlating the photocurrent distribution to the device geometry than any of the source images. In the methods where the maximum response to a decision operator is considered the resulted images preserve the original sharpness but contain a large number of high intensity artefacts around the borders of the fused regions, while in the methods where averaging is performed using as weights the response to the decision operator, the intensity of the border artefacts is reduced at the cost of image sharpness.

5. Acknowledgement

The research presented in this publication has been supported by the Romanian National Program of Research, Development and Innovation PN-II-IDEI-PCE, Grant 1566/2009, UEFISCU CNCSIS.

6. References


Chen, Y; Wang, L; Sun, Z; Jiang, Y and Zhai, G. Fusion of color microscopic images based on bidimensional empirical mode decomposition, *Optics Express* 18, 21757-21769 (2010), ISSN 1094-4087.


Image fusion technology has successfully contributed to various fields such as medical diagnosis and navigation, surveillance systems, remote sensing, digital cameras, military applications, computer vision, etc. Image fusion aims to generate a fused single image which contains more precise reliable visualization of the objects than any source image of them. This book presents various recent advances in research and development in the field of image fusion. It has been created through the diligence and creativity of some of the most accomplished experts in various fields.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
