# Confocal microscopy by aperture correlation

**T. Wilson, R. Juškaitis, M. A. A. Neil, and M. Kozubek**

*Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK*

Received August 13, 1996

Most confocal microscopes do not produce images in real time with nonlaser light sources. The tandem scanning confocal microscope does produce such images but, because the pinhole apertures of the Nipkov disk must be placed far apart to reduce cross talk between neighboring pinholes, only 1% or less of the light available for imaging is used. We show that, by using aperture correlation techniques and relaxing the requirement to obtain a pure confocal image directly, one can obtain real-time confocal images with a dramatically increased (25% or even 50%) light budget. © 1996 Optical Society of America

The unique three-dimensional imaging properties of the confocal microscope arise from the use of a point source and a point detector.[1] A laser is usually required as the light source to achieve sufficiently bright illumination. However, it is clearly attractive to remove the need for laser illumination, as this would not only increase the number of available imaging wavelengths but also reduce cost. To do so it is necessary to increase the light budget, which requires that parallel confocal systems be implemented in which the pinhole apertures are placed as close together as possible and some method of preventing cross talk between adjacent confocal systems is devised. In the following we show how this can be achieved by use of a mask in which the transmittance of the individual pinhole apertures is programmable.

The basic idea is shown in Fig. 1. The aperture mask is imaged onto the specimen, and light reflected back from the specimen passes back through the mask. The final image is captured by a CCD camera that is focused onto the mask. To determine how the apertures of the mask should be programmed to prevent cross talk between adjacent apertures in the mask, we consider, for ease of analysis, the transmission system shown in Fig. 2. The object is characterized by an amplitude transmittance function $\tau(\mathbf{x_0})$, the source mask by an intensity sensitivity $S(\mathbf{x_1})$, and the detector mask by $D(\mathbf{x_2})$. If we consider the illumination to be incoherent, the intensity just after the detector mask is

$$I(\mathbf{x}_2) = \int S(\mathbf{x_1})D(\mathbf{x_2}) \left| \int h_1\left(\mathbf{x_0} + \frac{\mathbf{x_1}}{M}\right) \times \tau(\mathbf{x_0})h_2\left(\mathbf{x_0} + \frac{\mathbf{x_2}}{M}\right)\mathrm{d}\mathbf{x_0}\right|^2 \mathrm{d}\mathbf{x_1}\,, \tag{1}$$

where $h_1$ and $h_2$ are the amplitude point-spread functions of the two imaging lenses and $M$ represents the magnification.

Let us now specialize to the reflection case and consider ideal point source and detector apertures positioned at $\mathbf{x_i}$ such that we can set $S(\mathbf{x}) = D(\mathbf{x}) = \delta(x - x_i)$, where $\delta(.)$ denotes the Dirac delta function. This gives an intensity in the detection plane at position $\mathbf{x_i}$ of

$$I(\mathbf{x_i}) = \left| \int h_1(\mathbf{x})h_2(\mathbf{x})\tau\left(\mathbf{x} - \frac{\mathbf{x_i}}{M}\right)\mathrm{d}\mathbf{x}\right|^2, \tag{2}$$

which is the well-known equation describing image formation in confocal microscopes.[1] However, Eq. (2) describes the intensity corresponding to one object point alone, so it is necessary to introduce scanning, i.e., to vary $\mathbf{x_i}$ to form an image of a finite region of the object.

Our proposed system does not require scanning, and the source and detector aperture masks consist of a fixed array of apertures with programmable pixel transmittance. Our plan is to encode each pixel transmittance with a time sequence of transmittances that is uncorrelated with that presented to every other pixel. Let us specialize to the reflection case and write

$$S(\mathbf{x}) = D(\mathbf{x}) = \sum_{i=1}^{N} b_i(t)\delta(\mathbf{x} - \mathbf{x_i})\,, \tag{3}$$

which describes an array of $N$ point apertures with transmittance $b_i(t)$ positioned at $\mathbf{x_1}, \mathbf{x_2} \ldots \mathbf{x_i} \ldots \mathbf{x_N}$. The coefficients $b_i(t)$ represent the time sequence of pixel transmittances that must be chosen to provide the desired correlation. The final recorded image is given by the time average of Eq. (1), with $S(\mathbf{x_1})$ and $D(\mathbf{x_2})$ given by Eq. (3). Because $S$ and $D$ are the only time-dependent functions, it is sufficient to consider the time average of $S(\mathbf{x_1})D(\mathbf{x_2})$. Thus

$$\langle S(\mathbf{x_1})D(\mathbf{x_2})\rangle = \sum_{i=1}^{N}\sum_{j=1}^{N}\langle b_i(t)b_j(t)\rangle\delta(\mathbf{x_1} - \mathbf{x_i})\delta(\mathbf{x_2} - \mathbf{x_j})\,, \tag{4}$$

where $\langle.\rangle$ denotes the time average. For the sequence presented to pixel $i$ to have zero cross correlation with

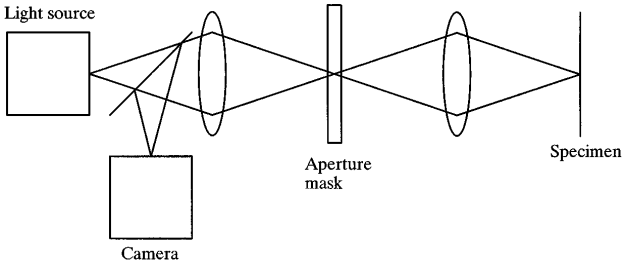


Fig. 1. Schematic diagram of the confocal microscope.

0146-9592/96/231879-03$10.00/0 

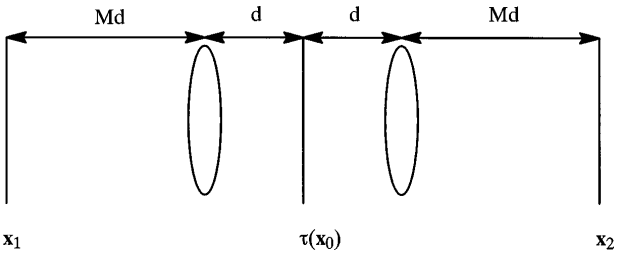


Fig. 2. Schematic diagram of the optical system.

that presented to any other pixel $i \neq j$, the sequence $b_i(t)$ must be chosen such that

$$\langle b_i(t) b_j(t) \rangle = \delta_{ij}\,, \tag{5}$$

where $\delta_{ij}$ represents the Kronecker delta. Any orthonormal sequence can be used for $b_i(t)$. One example would be an infinite random sequence of +1 and −1 or, more practically, finite-length complementary Golay sequences.[2] Unfortunately these sequences require negative values for $b_i(t)$, which is impossible to achieve optically because the pixel intensity transmittance cannot be negative. To overcome this problem we introduce a dc shift and encode the pixel transmittances according to

$$S(\mathbf{x}) = D(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^{N} [b_i(t) + 1] \delta(\mathbf{x} - \mathbf{x_i})\,, \tag{6}$$

where the pixel transmissivities are now given by $[b_i(t) + 1]/2$, which switch from 0 to 1 as $b_i(t)$ switches from −1 to +1. The final image is given by the time average of Eq. (1) together with $S$ and $D$, which are given by Eq. (6). The only time-dependent part of this is given by

$$\langle S(\mathbf{x_1}) D(\mathbf{x_2}) \rangle = \frac{1}{4} \sum_{i=1}^{N} \sum_{j=1}^{N} \langle [b_i(t) + 1][b_j(t) + 1] \rangle \times \delta(\mathbf{x_1} - \mathbf{x_i}) \delta(\mathbf{x_2} - \mathbf{x_j})\,. \tag{7}$$

If the sequence $b_i(t)$ is chosen according to Eq. (5) together with

$$\langle b_i(t) \rangle = 0\,, \tag{8}$$

then we obtain

$$\langle S(\mathbf{x_1}) D(\mathbf{x_2}) \rangle = \frac{1}{4} \left[ \sum_{i=1}^{N} \delta(\mathbf{x_1} - \mathbf{x_i}) \delta(\mathbf{x_2} - \mathbf{x_i}) + \sum_{i=1}^{N} \sum_{j=1}^{N} \delta(\mathbf{x_1} - \mathbf{x_i}) \delta(\mathbf{x_2} - \mathbf{x_j}) \right]. \tag{9}$$

The first term in Eq. (9) can be written in exactly the same form as in Eq. (2) and represents the pure confocal image.

The second term can be thought of as that obtained when the pixels of the aperture mask are all switched to the same state, which will be a nonconfocal image. In the limit that the pixel apertures are adjacent to each other with no dead space between them, $S(\mathbf{x}) = D(\mathbf{x}) = \sum_{i=1}^{N} \delta(\mathbf{x} - \mathbf{x_i}) = 1$, and hence the second term represents a pure conventional image.

Equations (6) and (9) suggest the basis of our method. We first obtain a composite image when the aperture pixels are encoded according to Eq. (6). This image will consist of a confocal image superimposed upon a nonconfocal image. We then obtain the nonconfocal image by switching all the pixel transmittances to the same value. We finally obtain the confocal image by subtracting the nonconfocal contribution from the composite image.

To implement our method it is necessary to decide how to realize the aperture mask. In essence we need to be able to vary the transmittance of each pixel in the detector plane according to the appropriate correlation sequences. We consider two equivalent approaches. One is to use a programmable spatial light modulator that might, for example, be based on ferroelectric liquid crystals; these devices have fast enough switching times that real-time imaging is possible. However, these devices are currently expensive; a much simpler and cheaper alternative is to impress the desired transmittance sequences onto a disk photolithographically. If the disk is rotated, the pattern presented in time to any fixed pixel in the detector plane is equivalent to that impressed in space along the corresponding arc of the disk. Mathematically, this means that Eq. (6) must be modified to

$$S(\mathbf{x}) = D(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^{N} [b_i(\mathbf{x_r}) + 1] \delta(\mathbf{x} - \mathbf{x_i})\,, \tag{10}$$

where $\mathbf{x_r}$ denotes the rotation coordinate. We now obtain the final image by substituting Eq. (10) into Eq. (1) and averaging over $\mathbf{x_r}$. Because this is formally identical to averaging Eqs. (1) and (6) over time, we see that this technique will also produce a composite confocal and conventional image. It is then necessary to obtain a conventional image to extract the required confocal image.

To demonstrate the imaging capabilities of our new confocal system, we produced the aperture mask by photolithographically defining 80-$\mu$m-square pixels on an 80-$\mu$m pitch in aluminum on a Perspex disk. The pixel transmittance was either zero or unity, and the pattern of transmissive pixels was randomly distributed over a sector of the disk such that the number of pixels along an arc within the sector was typically 400. A complementary pattern was impressed upon the adjacent sector such that the track along any given arc sees an average transmittance of exactly 0.5. This ensures that $\langle b_i(t) \rangle = 0$, as in Eq. (8). Because it is necessary to subtract a conventional image from the composite image to obtain the pure confocal image, we also provided a blank sector on the disk to produce the conventional image. The system consisted of a 250-W tungsten halogen incandescent projector lamp together with a narrow-band green filter ($\lambda = 546$ nm, $\Delta\lambda = 10$ nm). A simple CCD camera was used to capture both the composite and the conventional images. Our image-capture system consisted of a PC together with an imaging board that allowed us to perform near-real-time subtraction of the conventional image from the composite image and to display the confocal image at the rate of 17 frames/s. This rate was limited by the equipment available to us, and TV rate imaging is available by upgrading to

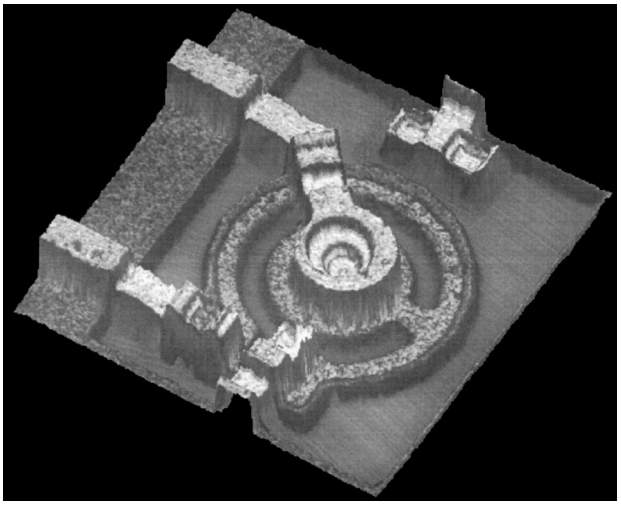

Fig. 3. Three-dimensional image of a transistor. The height projection image is combined with an autofocus image to provide surface rendering. A $50\times/0.8$ numerical aperture objective was used. The image field is 120 $\mu$m $\times$ 120 $\mu$m.

existing PCI bus systems. Figure 3 shows the confocal capabilities of the system. A $50\times/0.8$ numerical aperture objective was used, and 34 through-focus image slices were recorded over a 6.2-$\mu$m focal depth. The image shows a three-dimensional height representation of the transistor onto which an autofocus image has been superimposed to render the surface texture. This image is indistinguishable from those obtained with our other traditional confocal systems.

In our system the encoded sectors contained an equal number of transmissive and nontransmissive pixels, and hence the confocal image is formed by use of 25% of the available light. Tandem scanning instruments,[3] by contrast, use only 0.5%–1% of the available light. Indeed, in an implementation using ferroelectric liquid-crystal spatial light modulators and polarization effects to separate the confocal from the combined image a 50% light budget can be obtained.

By relaxing the requirement to obtain a pure confocal image directly, we have built a very simple, light-efficient microscope. The key to our approach has been to use aperture encoding to ensure zero cross correlation between different pixels in an aperture mask. This removes the need to space the apertures far apart and hence provides for a highly light-efficient confocal microscope.

M. Kozubek acknowledges the support of the Soros Foundation. This study is supported by the Paul Instrument Fund of the Royal Society.

## References

1. T. Wilson, ed., *Confocal Microscopy* (Academic, London, 1990).
2. M. J. E. Golay, J. Opt. Soc. Am. **39,** 437 (1949).
3. M. Petran, M. Hadravsky, M. D. Egger, and R. J. Galambos, J. Opt. Soc. Am. **58,** 661 (1968).