Phase contrast microscopy with full numerical aperture illumination

Christian Maurer, Alexander Jesacher, Stefan Bernet and Monika Ritsch-Marte
Division for Biomedical Physics, Innsbruck Medical University, 6020 Innsbruck, Austria
Corresponding author: Stefan.Bernet@i-med.ac.at

Abstract: A modification of the phase contrast method in microscopy is presented, which reduces inherent artifacts and improves the spatial resolution. In standard Zernike phase contrast microscopy the illumination is achieved through an annular ring aperture, and the phase filtering operation is performed by a corresponding phase ring in the back focal plane of the objective. The Zernike method increases the spatial resolution as compared to plane wave illumination, but it also produces artifacts, such as the halo- and the shade-off effect. Our modification consists in replacing the illumination ring by a set of point apertures which are randomly distributed over the whole aperture of the condenser, and in replacing the Zernike phase ring by a matched set of point-like phase shifters in the back focal plane of the objective. Experimentally this is done by illuminating the sample with light diffracted from a phase hologram displayed at a spatial light modulator (SLM). The subsequent filtering operation is then done with a second matched phase hologram displayed at another SLM in a Fourier plane of the imaging pathway. This method significantly reduces the halo- and shade-off artifacts whilst providing the full spatial resolution of the microscope.

© 2008 Optical Society of America

OCIS codes: (180.0180) Microscopy, (100.5090) Image processing - phase only filters, (090.1760) Holography - computer holography, (070.6110) Fourier optics and signal processing - spatial filtering, (170.0180) Medical optics and biotechnology - microscopy

References and links
by re-shifting the phase of the zero-order light by the amount of the diffracted light. In the central phase contrast method the image contrast is then achieved which is called the zero-order wave. If the phase shift induced by the object is small enough, fraction of the illumination wave which has transmitted the sample without interaction, and consists of components diffracted by the sample's phase structures, and a direct, undiffracted mersive phase sample is illuminated by a plane light wave. The transmitted light field then spatial resolution of the microscope, since in transmission microscopy the spatial resolution

1. Introduction

One important method to detect phase variations in an object is the phase contrast microscopy invented by Zernike in the 1930's [1, 2]. In the original central phase contrast variant, a transmissive phase sample is illuminated by a plane light wave. The transmitted light field then consists of components diffracted by the sample's phase structures, and a direct, undiffracted fraction of the illumination wave which has transmitted the sample without interaction, and which is called the zero-order wave. If the phase shift induced by the object is small enough, then the zero-order wave is advanced by approximately $\pi/2$ relative to the average phase of the diffracted light. In the central phase contrast method the image contrast is then achieved by re-shifting the phase of the zero-order light by the amount of $\pi/2$. This is done in a back focal plane of the objective, where the zero-order wave has a focus which coincides with an inserted point-like phase shifter. In the camera plane of the microscope the interference of the phase-shifted zero-order light with the undisturbed scattered light then produces an image with an intensity distribution proportional to the optical thickness of the sample.

Whereas the central phase contrast method still has some applications, as for example in a modified version for the steering of optical tweezers with SLMs [3], efficient laser projection [4], or common path interferometry [5], nowadays it is only marginally used in microscopy. The reason is that the required plane wave illumination reduces both the transverse and the axial spatial resolution of the microscope, since in transmission microscopy the spatial resolution
depends on the sum of the numerical apertures (NA) of the illumination and the imaging optics [6, 7, 8, 9, 10]. A further disadvantage of plane wave illumination is that it produces undesired sharp shadows in the image plane, which result from out-of-focus scatterers like dust particles or scratches in optical components, and which reduce the image quality.

Therefore the commonly used variant of Zernike phase contrast microscopy uses a modified illumination and filtering method: Illumination is done through a ring shaped aperture which is located in the front focal plane of the condenser lens. As a result the sample is illuminated with a uniform light field, which has a cone-shell shaped directional distribution. Behind the sample, in the back focal plane of the microscope objective, a sharp image of the illumination ring aperture is formed. This ring of light corresponds to the zero-order beam, whereas the diffracted part of the image wave is dispersed in the same plane. The Zernike phase contrast method is now realized by inserting a ring shaped phase filter into this plane (normally, this filter ring is coated directly on the surface of the rear lens of a phase contrast objective), which coincides with the imaged illumination aperture ring, and which shifts the phase of the transmitted zero-order light by $\pi/2$. A final imaging lens then generates a sharp image of the phase contrasted object in the camera plane. The advantage with respect to the central phase contrast method is that the wider directional distribution of the illumination directions increases the microscopic resolution due to the enhancement of the effective numerical aperture [5, 6, 7, 11, 12], and it reduces the disturbing sharp shadows of possibly soiled optical components by directional averaging.

However, the Zernike phase contrast method also causes some undesired artifacts, namely the halo- and the shade-off effect [13]. Halos are bright narrow boundaries around dark image regions, and vice versa. The shade-off effect corresponds to a misleading drop-off of the image intensity in the center of extended bright sample structures, and an intensity increase in the centers of dark areas, even if the actual sample structures are completely uniform. Although some improvements of the Zernike phase contrast method were reported [11, 13], both the halo and the shade-off effects are principally not fully avoidable, since they are caused by the ring shaped aperture of the phase mask (see Fig. 1).

These artifacts arise, because not only the real zero-order part of the illumination wave (i.e. the sharply imaged ring of light behind the objective) is phase shifted by $\pi/2$, but also diffracted light which passes through the phase ring. In a thought experiment the ring of light can be considered as being composed of a ring-shaped chain of light dots, each of them corresponding to a certain incidence direction of the illumination light at the sample. If one considers only one of the plane-wave incidence directions which compose the ring of light, i.e. one of these dots (indicated as a small circle in Fig. 1), its corresponding diffracted Fourier components are distributed around it (indicated as a dimmer circle around the central zero-order spot). A part of these diffracted light components will also pass through the phase ring (indicated as "erroneously phase shifted components"). These scattered components are closely adjacent to the zero-order component and thus correspond to coarse phase structures within the sample. Since these wave components are phase shifted by the same amount as the zero-order wave, their image in the camera plane, corresponding to the central part of extended areas, will have the same intensity as the zero-order background of the image, giving rise to the shade-off effect. Furthermore, at the edges of the extended areas the interference of the zero-order wave with the diffracted light components which have acquired the erroneous phase leads to the halo-effect. Overall both the shade-off and the halo artifacts are due to scattered parts of the image wave which pass erroneously through the ring shaped phase filter [13].

Here we demonstrate an improvement of this situation by illuminating the sample with a variety of incident plane waves with randomly chosen (but known) directions of incidence. In this case the light intensity is uniform in the entire sample plane, but it focusses as a pseudo-
Fig. 1. Explanation of the artifacts in Zernike phase contrast microscopy, and by random light dot illumination: (A) sketches the phase-only Zernike filter in the back focal objective plane, whereas (B) sketches a filter used for random light dot illumination (not to scale): (A) shows the Zernike type phase ring (brighter) which coincides with the ring shaped image of the illumination aperture. In the ideal case only the zero-order wave should pass through the phase ring. In the figure only a small portion (small dot) of the illumination light ring is indicated. This dot corresponds to a part of the ring-shaped zero-order wave, with its diffracted components spread-out around it (indicated as a dimmer disk around the central dot). As shown in the figure, also a part of this diffracted light passes through the adjacent areas of the phase ring, and is thus erroneously shifted in its phase, giving rise to image artifacts. In (B) the situation is sketched for "random dot" illumination, i.e. the sample is illuminated with a variety of plane waves which are incident from randomly chosen directions. In the sketched filter plane, these illumination directions focus at randomly distributed spots, but at known positions. The corresponding phase filter is designed such that it exactly matches with the focussed points, shifting their phases by $\pi/2$ with respect to the surrounding diffracted light components. Compared to the situation (A) there is now much less intensity of the scattered light which erroneously passes through phase-shifting areas of the filter.

random dot pattern in the back focal plane of the objective. This pattern now takes the role of the zero-order component of the image amplitude. There, the role of the Zernike phase ring is now taken by a mask of small $\pi/2$—phase-shifting apertures at the known positions of the light spots (indicated in Fig. 1B), thus extending the central phase contrast method to the case of many simultaneously present plane wave illumination directions.

The advantage of this phase contrast variant is that now much less diffracted light passes through phase shifting areas ("the dots") of the phase filter, even if their integral area is the same as that of the Zernike ring. The reason for this is that typically the scattered light intensity around each spot strongly falls off with increasing distance from its center, such that only a small portion of the diffracted intensity will pass through other spots of the filter. This reduces both the halo and the shade-off effect. Furthermore, the average light distribution of the random spot illumination does not possess the symmetry of the Zernike ring illumination. This leads to a correct weighting of the different spatial frequency components in the image intensity (for quantitative measurements), while optimizing the image resolution by making use of the full numerical aperture range of the illumination optics.
Fig. 2. Sketch of the experimental setup: A collimated laser beam illuminates a Fourier hologram displayed at a first phase-only SLM. With a Fourier transforming lens, the holographically programmed illumination pattern is reconstructed in the plane of a rotating diffuser, acting as the effective incoherent illumination source. A further Fourier transforming (condenser) lens leads to a uniformly illuminated sample, which is then imaged with a microscope objective. In its back focal plane the programmed illumination pattern (which was displayed at the rotating diffuser screen) is sharply imaged in the plane of a second SLM which acts as a programmable phase-only Fourier filter, displaying for example the phase masks of Fig. 1. A final Fourier transforming imaging lens then produces a sharp, processed image of the sample at a camera.

2. Experimental setup

In order to test different variants of phase filtering methods, we use a high resolution phase-only SLM as a filter in the back focal plane of the microscope objective. In earlier experiments Ng et al. [5] already used a SLM as a central phase contrast filter with a variable phase-shift for common path interferometry, and they also demonstrated the resolution enhancing effect of averaging over multiple illumination directions in a sequential way. In [14] it has been demonstrated that an SLM displaying off-axis phase holograms can switch a microscope between phase contrast and dark-field imaging modes, and also enables new phase filtering methods such as spiral phase contrast [15], or spiral phase interferometry [16].

However, all of these methods were performed with plane wave illumination, and thus have the above mentioned disadvantages of reduced resolution and the appearance of disturbing shadows from out-of-focus scatterers. In the present paper we introduce a second SLM in the condenser beam path for shaping also the illumination light. The use of SLMs as holographic projectors has been primarily developed for the steering of so-called holographic optical tweezers [17, 18, 19, 20], since a hologram projection makes a more efficient use of the available light intensity than a direct projection of intensity modulated patterns.

The principle setup is sketched in Fig. 2. An expanded laser beam illuminates a phase-only, off-axis Fourier hologram which is displayed at the first high resolution SLM. It projects a tailored intensity distribution through a Fourier transforming lens onto a rotating ground glass.
diffuser screen. For example, if the Zernike phase contrast method is simulated, then the programmed intensity pattern just corresponds to a ring of light, as indicated in the sketch. The rotating diffuser screen is employed (similar to the experiment in [5]) in order to avoid disturbing static speckle patterns by averaging over time-varying speckle fields if the image integration time is chosen long enough (which is in our case > 1 ms). Behind a condenser lens the random spatial and temporal phase produced by the rotating diffuser generates a homogeneous sample illumination for all kinds of illumination patterns. The sample is then imaged with a microscope objective. In its rear focal plane, where a second high resolution phase-only SLM2 is located, a sharp image of the illumination pattern is reconstructed, which corresponds to the zero-order image wave, whereas the scattered light components are diffusely distributed. The SLM2 acts as a phase-only spatial Fourier filter, which is used as a phase shifter for the zero-order part of the image wave. A filtered sharp image of the sample is then produced by a final Fourier transforming lens and recorded by a CCD camera.

The actual experimental setup is slightly more complicated, since the two employed SLMs are reflective devices (not transmissive as indicated in the sketch in Fig. 2 in order to reduce its complexity), which act as off-axis reflection holograms, and thus the beam path is actually folded. In the experiment two microscope objectives (Zeiss 40x EC Plan Neofluar NA = 0.75, and Zeiss 40x NA = 1.3) were used as the condenser and the imaging objectives, respectively. Since the pupil plane of these objectives are not accessible, these were imaged with two telescope systems onto the rotating diffuser and at SLM2, respectively. The first telescope consists of two f=100 mm lenses, the second one is built with the tubus lens of the microscope (f = 160 mm) and a second achromatic lens (f = 150 mm). For the illumination a frequency doubled Nd:Yag laser with a maximal output power of 200 mW is used. The beam is expanded by a factor of 30 and illuminates the first SLM1. The two SLMs are high resolution, phase only light modulators. SLM1 (Holoeve HEO 1080 P) has a resolution of 1920x1080 pixels with a pixel size of 8 µm, whereas SLM2 (Holoeve LC-R 3000) has a resolution of 1920x1200 pixels with a size of 9.5 µm. Both of the SLMs are used to display off-axis holograms, which means that the desired phase-front modulations are generated in their first diffraction orders, whereas the residual diffracted orders are blocked. The holograms displayed at SLM1, which produce the illumination patterns, are calculated using the so-called Gerchberg-Sexton algorithm [21, 22].

SLM2 just displays the selected phase filters, like a Zernike ring or an array of dots, but the corresponding phase structures are superposed by a blazed grating which diffracts the filtered image wave to the camera. The lenses and objectives of the setup are chosen such that the maximal resolution of the system can be detected with a CCD camera (DVC 1412, DVC Co.) with a pixel size of 6.5µm.

An initial calibration routine was necessary in order to map the filter structure displayed at SLM2 to the size and position of the holographically projected pattern by SLM1. This was done by displaying test patterns at SLM1, consisting of blazed gratings with different periods and orientations. Each of them produces a single spot at a certain position of the rotating diffuser screen, which is then sharply imaged at the surface of SLM2. There a corresponding test hologram is displayed consisting of an off-axis spiral phase hologram [23] which has a phase discontinuity in its center. As soon as the projected light spot hits the center of the spiral phase hologram, a strong intensity decrease is observed at the camera, since the light is scattered out of the imaging pathway by the phase discontinuity [24]. Thus, by shifting the center of the spiral phase hologram under computer control across the SLM display, it is possible to map the grating vectors in the SLM1 plane to the corresponding focal spot positions in the SLM2 plane. With this mapping stored in the computer, matched source and pupil masks in the condenser and filtering plane can be generated in a straightforward way.
3. Results

For comparing the performance of the Zernike phase contrast and our suggested random dot illumination phase contrast, we imaged two test samples with both of the methods (Fig. 3). The first test sample consisted of polystyrene beads (refractive index 1.59) with a diameter of 10 µm, which were suspended in immersion oil (refractive index 1.52). The left image was recorded with the Zernike phase contrast method by holographically projecting a ring of light as an illumination source onto the rotating diffuser (indicated in the inset of the image), and by displaying a corresponding ring-shaped π/2 phase shifter at the SLM2 in the back focal plane of the microscope. In the presented experiment the Zernike ring produced a numerical aperture of the illumination light of NA = 0.5. Below the image, a plot of the intensity along an exemplary horizontal line (indicated in the image) is shown. As expected, the halo artifact manifests itself as a dark boundary of the bead (arrows pointing downwards), and an increased intensity further away.

The next image of the same sample was recorded under the same conditions by the random dot phase contrast method (indicated in the inset). There, 50 randomly chosen dots with a uniform distribution over the numerical aperture of the illumination condenser (NA=0.75) were projected onto the rotating diffuser to act as the effective illumination source, and subsequently filtered by 50 corresponding π/2-phase shifting disks in the back focal plane of the microscope objective. For a “fair comparison”, the integrated phase shifting area of the 50 dots was chosen equal to the phase-shifting area of the Zernike ring used before. The intensity profile along the same horizontal line as in the previous plot shows that the halo artifact is significantly reduced with respect to the Zernike phase contrast method.

The next two images show the same type of comparison using another phase sample consisting of an oil smear surrounded by water, sandwiched between two glass cover slips. There one can expect that the optical thicknesses of the oil smear and the surrounding water are constant, such that a quantitative phase contrast method should generate a constant image intensity within the oil and the water-filled areas. The images where recorded under the same conditions as before. In the Zernike image (left) the halo artifact appears again as a dark boundary of the oil smear (arrows pointing downwards) followed by an intensity increase. Since in this case the optical thickness of the sample is uniform, it is now also possible to quantitatively observe the shade-off artifact as an intensity decrease of the central oil smear area with respect to its surrounding (arrow pointing upwards).
The last image of the same sample was again recorded by the random dot phase contrast method. Similar to the last example a reduction of the halo-artifact is obtained, but additionally it is now demonstrated that also the shade-off artifact is suppressed, leading to an almost constant image brightness along the cross section of the oil smear, as indicated in the intensity profile plot sketched below.

In order to further characterize the method, we now give a short summary of the results of additional measurements. First of all an optimal number of randomly distributed dots should exist between the two limiting cases, namely illumination with just one single dot, and with an infinite number of dots. The first case corresponds to the central phase contrast method, which provides on the one hand the best suppression of artifacts, but on the other hand it has the disadvantages of a reduced resolution and of noise from out-of-focus scatterers. The second case corresponds to a uniform bright-field illumination, which generates no phase contrast. Thus a compromise between high image resolution, low noise from out-of-focus scatterers, and a high contrast enhancement has to be found, which depends on the number of illumination dots.

For an experimental determination we have measured the noise from out-of-focus scatterers as a function of the number of illumination points. For this kind of measurement we removed the sample from the microscope, such that the system imaged only scattered light from unavoidable contaminations in the optical path, like dust particles or scratches at lenses. Then the standard deviation of the spatial image intensity distribution was computed as a probe for the out-of-focus noise. As expected, a single dot illumination produces sharp shadows of the out-of-focus scatterers in the camera plane, corresponding to a high standard deviation. Increasing the number of illumination dots reduces the standard deviation by averaging over multiple equal, but spatially displaced out-of-focus images, which overlap quasi-incoherently due to the temporal phase variations introduced by the rotating diffuser plate, and the time-averaging of the camera. It turned out that for few illumination dots the image noise decreases strongly with an increasing number of dots, but it soon reached a constant level for more than 20 illumination points.

In the next experiment test phase samples like those displayed in Fig. 3 were imaged with an increasing number of illumination dots. There it turned out that their image contrast seems to strongly fluctuate for a low number (10 or less) of illumination directions, which is actually due to the significant noise contribution from out-of-focus scatterers, as explained above. For more than 10 and up to 100 source points the contrast becomes constant, and from then on it slowly reduces until it vanishes almost completely for 2000 or more points. Practically it turned out that a range between 20 to 50 illumination dots seems to be ideal. However, the optimal number may vary for different setups, since it depends on their explicit optical parameters, particularly on the diameter of the sharply imaged dots in the plane of the second SLM with respect to the total imaged area in this plane. This depends on the optical resolution of the illumination pattern projected at the rotating diffuser plane, and on the further resolution when imaging this pattern into the plane of the second SLM, which is determined by the numerical apertures of the condenser and the objective lenses. In our case the size of the focused light dots in the plane of SLM_2 is on the order of 40 \( \mu \text{m} \), such that the diameter of each phase shifting disk is chosen to be 50 \( \mu \text{m} \), whereas the total aperture of SLM_2 has a diameter of 10 mm.

A further question may arise, namely whether there is a more preferable illumination pattern than a uniform random dot distribution. Even in the case of such randomly distributed spots in the Fourier filter phase mask, a small fraction of scattered light from one illumination source will pass through adjacent phase-shifting dots, which produces artifacts. However, the random distribution of the neighboring dots assures that the erroneous phase shift appears for different scattered Fourier components of each source point, such that the artifacts do not "accumulate"
for certain Fourier components, like in the case of the Zernike method. Such accumulated artifacts would also appear for a regularly spaced array of illumination sources, since then corresponding higher Fourier components of each source point would be erroneously phase shifted, resulting in a distortion of all image structures with a certain size and orientation.

However, a random but non-uniform distribution may be advantageous in some situations. For example, a higher concentration of dots at the outer part of the aperture results in an apparent increase of the image resolution, since now more light rays are incident under oblique directions, and thus have to be scattered by a larger angle in order to be collected by the objective. This over-weights the intensity distribution of fine sample details in the image, since these provide the largest scattering angles.

4. Conclusion and outlook

We have demonstrated a method to reduce artifacts in phase contrast microscopy, and to increase its resolution by using the full numerical apertures of the condenser and of the objective for imaging. This is achieved by extending the original central phase contrast method to a situation where the sample illumination is performed simultaneously with various temporally incoherent plane waves coming from different directions, and by using a matched phase filter in the back focal plane of the microscope. For demonstration of the method we used two phase-only SLMs, which holographically generated the illumination distribution and the Fourier filter masks, respectively. However, in principle the method can also be employed in a standard Zernike phase contrast microscope by replacing the condenser annulus by a random dot illumination mask, and the Zernike phase filter by a matched random dot phase shifter. The required random dot phase mask can be produced for example by photolithography.

Although the setup using two matched SLMs is complex, it also offers new possibilities. In the present experiment we "destroyed" the temporal coherence of the imaging light by projecting it onto a rotating diffuser screen and by time-averaging in order to avoid speckled images. However, if this diffuser is omitted, the light keeps its full spatial and temporal coherence, allowing new possibilities to perform quantitative interferometric measurements. With the electronically steered SLMs several kinds of illumination and matched filter functions can be displayed, as for example spiral phase filtered images from several oblique imaging directions. By averaging over a number of interferograms that are taken from different oblique incidence directions, one can obtain the phase information of a narrow sheet within a bulk sample, with a thickness comparable to the depth of focus of an image. This might allow a kind of interferometric tomography by recording a stack of "multiple-direction" interferograms at different sample depths [5].

Acknowledgment

This work was supported by the Austrian Science Fund (FWF) Project No. P19582-N20.