Optical Coherence: Recreation of the Experiment of Thompson and Wolf

David Collins

Senior project

Department of Physics,

California Polytechnic State University

San Luis Obispo

June 2010
Abstract

The purpose of our experiment is to recreate the experiment done by Thompson and Wolf in 1957 on the measurement of optical coherence. What is novel about our approach is that it allows one to view the effects of source size on optical spatial coherence in real time.

Background and Introduction

The main idea in optical coherence is the degree of correlation between different points on an optical wave field. In order to measure or quantify the degree of correlation one can use the apparatus shown in figure 1, which is a modification of Young’s double slit interference experiment.

First let us consider an extended source illuminating two pinholes (P₁ and P₂) in an opaque screen. We are interested in the field at some point P behind the screen, and in the plane going through that point and parallel to the screen. Figure 1 illustrates the setup for a source S of quasi-monochromatic light. The source S can be thought of as a primary source, while the points P₁ and P₂ can be thought of as secondary sources.

If there is perfect correlation of phase between P₁ and P₂, then the interference pattern produced will have fringes which are perfectly visible, meaning that in some locations in the plane parallel to the screen and at the distance P, the points will completely destructively interfere, and in other locations in this same plane they will be perfectly in phase and completely constructively interfere. If there is no correlation of
phase between \( P_1 \) and \( P_2 \), then there will be no visible interference. We define the visibility with the following equation:

\[
V = \frac{I_{\text{max}} - I_{\text{min}}}{I_{\text{max}} + I_{\text{min}}}
\]  

(1)

where \( I_{\text{max}} \) and \( I_{\text{min}} \) are the maximum and minimum intensities of the interference fringes. The visibility of the fringes increases with coherence, and is always between 0 and 1; 0 indicates no visible fringes while a 1 indicates maximum visibility, with \( I_{\text{min}} = 0 \).

![Diagram](image)

**Figure 1:** Method for determining the degree of spatial coherence. An extended source (S) illuminates the screen. Light passing through pinholes at positions \( P_1 \) and \( P_2 \) on the screen aperture (A) form interference patterns for a coherent source.

Another way of thinking about the formation and visibility of the fringes is that with an extended source, it will lead to overlapping fringe patterns which cancel each other out when the peaks and troughs of each pattern do not coincide, and results in low visibility fringes. However, there will still be a small degree of coherence between
points $P_1$ and $P_2$ because although light coming from different parts of the source is incoherent, radiation coming from the same part of the source will be coherent. In general, the degree of coherence is a measure of the ability to make interference fringes. Figure 2 demonstrates how an extended source leads to lower visibility.

![Diagram](image)

Figure 2: An extended source can be thought of as a composition of many individual sources. Because they are oriented differently with respect to the double slit, the wave fronts will have different angles as they pass through the slits which results in the waves having different phases in each slit than if they were parallel to the slits. This shifts the fringe pattern slightly with respect to the others, blurring the pattern and reducing visibility.

The experiment of Thompson and Wolf [1] involved studying the general interference law for partially coherent fields. The arrangement is shown in figure 3. They used a mercury-vapor compact lamp for their primary source ($S_0$), and used a filter with a pass band of about 40 nm centered on the mercury yellow doublet, which has an average wavelength of 579 nm. The filter blocks out all other spectral lines. The source is then imaged by a lens onto a pinhole (of diameter $9 \times 10^{-2}$ mm) located at $S_1$, after
which the light is rendered parallel by lens \( L_1 \). This arrangement was used to reduce aberrations. A second lens \( L_2 \), which is strictly similar to \( L_1 \), is placed in front of the two pinholes at \( A \). The light is brought to a focus at \( F \), and the diffraction pattern is viewed using a microscope. A mirror (M) is used to reduce the overall length of the equipment. The two pinholes were separated by distances varying between 0.2 cm and 6 cm, which were placed symmetrically about the axis at \( P_1 \) and \( P_2 \). Because they had a low level of light in the final diffraction pattern, special measures were taken – such as using special film and a long exposure time – to produce visible images.

![Figure 3: Apparatus for viewing diffraction patterns. [1]](image)

The way the visibility varies with the separation of the pinholes can be calculated from the van Cittert-Zernike theorem [1]. In the case of a circular aperture and varying the pinhole separation the visibility is of the form \( |J_1(x)/x| \) (where \( J_1 \) is a Bessel function of the first kind and \( x \) is proportional to the pinhole separation) and is shown in figure 4.
Figure 4: The degree of coherence as a function of pinhole separation. [1] Note that the points are not actual measurements of visibility – they correspond to the expected visibility for images shown in the paper.

One can see that the visibility falls to zero and then gets larger again. After it goes to zero, the fringe patterns exhibit a phase reversal where the bright fringes become dark and vice versa (one can see this later in figure 9).
Since the experiment of Thompson and Wolf, some people have recreated the experiment with slight alterations, or with better technology. One such example is the experiment of Ambrosini, et al. [2] They repeated the Thompson and Wolf experiment using a CCD camera and a measurement method which utilizes the fast Fourier transform.

Their experimental setup was slightly different than that of Thompson and Wolf. Instead of using an incoherent primary source, they used an argon laser (λ = 0.5145 µm) and illuminated a continuously rotating ground glass plate. Because the laser spot size is large compared to the ground glass inhomogeneities, the cross section of the beam that emerges from the ground glass is essentially spatially incoherent. This beam is then focused by a lens L₁ (f = 50 mm) onto a pinhole of diameter 100 ± 1 µm, which acts as an incoherent source. A second lens L₂ (f = 500mm) is used to collimate the light as it passes through the two identical pinholes P₁ and P₂ (each of diameter 150 µm) which are separated by a distance 2d (ranging from 0.98 and 4.10 mm) that can be adjusted by a micrometer screw to a precision of ± 0.01 mm. A third lens L₃, which is identical to L₂, focuses the collimated light onto the CCD camera, which is cooled to -30° C to minimize dark currents due to thermally generated electrons. Figure 5 illustrates their experimental setup.
They estimated their error in measurements from two sources which they summed to get their total error. The first method involved the accuracy of the measurement algorithm and was determined by numerical simulations. Computer generated fringe patterns with known visibilities were analyzed using the fast Fourier transform. The error in visibility for the numerical method was $\pm 3 \times 10^{-4}$.

They also accounted for the error resulting from low light levels used. The noise of the system was about 80 electrons per pixel, evaluated under dark conditions with a 10 second exposure time.

They summed the total error from both contributions. The overall error was low, ranging from about .1% to 3% error for different pinhole separations.

**Experiment and Results**

The objective of our experiment is a recreation of the Thompson and Wolf experiment, but in a new manner which demonstrates the phase reversal of the fringes in real time by using fixed pinholes and a variable slit aperture. We used a sodium lamp
which emits principally at 589.0 and 589.6 nanometers, making it quasi-monochromatic. We also placed a piece of ground glass in front of the source to reduce effects resulting from the structure of the lamp. However, the lamp and ground glass by themselves are insufficient; for interference to occur, the light entering the slits must be partially coherent. We achieve this by placing a variable slit taken from a monochromator in front of the lamp, the width of which can be manually adjusted to be anywhere between 100 micrometers, and 3 millimeters to an accuracy of at least ± 10 µm. The slit acts as our source, and horizontally it is most similar to a point source (which emits coherent light) at the minimum width of 100 micrometers, and increasingly less like a point source as you widen the source slit. About 28.5 centimeters in front of the adjustable slit, we placed a double slit which acts as our two pinholes in figure 1. The slits were separated by 250 µm and had a width of 40 µm. After the double slit, we placed an achromatic lens of focal length 15 cm which focuses the light. At the focus point, we secured the CCD camera (Pulnix TM250). In order to reduce background noise, a cardboard box was placed around the apparatus with a hole where the source aperture is to block out stray light that did not pass through the double slits, and other unwanted radiation. In addition, a black cylindrical tube about 10 cm long was screwed onto the front of the camera, further reducing background noise. Figure 6 shows the basic setup, while figure 7 is a photograph of the laboratory set-up.
Figure 6: The experimental setup.

Figure 7: Picture of the experimental setup.

A great advantage of this setup is it provides the ability to observe the interference pattern change in real time as the source width is increased. This has the
same effect as using double slits of varying separations, but rather than being limited to a small set of data points, a large number are attainable.

If the source width is set to the minimum of 100 micrometers, then the visibility is very close to 1, as it would be with a coherent source such as a laser. As the width is widened, the visibility changes and can be shown to follow the function:

\[
V = \left| \text{sinc}\left(\frac{du}{\lambda L}\right) \right| \tag{2}
\]

where \(\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}\), \(d\) is the width of the source aperture, \(u\) is the separation between slits, \(\lambda\) is the wavelength of the light, and \(L\) is the distance between the source and the slits.

Data was taken for slit widths of 100 µm to 1500 µm, and the fringe patterns were saved as .bmp files using a National Instruments PCI 1407 frame-grabber. Four examples of such patterns are shown in figure 9. The gain on the camera can be turned up at 100 µm so that an interference pattern is still visible, and turned down at wider aperture sizes to avoid saturation. Because the fringe patterns are produced by interference between two slits of finite width, it can be shown that the intensity profile would follow the following equation:

\[
I(x) = A + Bsinc^2\left(\alpha(x - x_0)\right)[1 + V \cos(\omega(x - x_0) + \phi)] \tag{3}
\]

\(A\) is a uniform background resulting from any stray radiation and dark currents in the CCD camera, \(B\) is proportional to the source brightness and aperture width, \(x_0\) defines the center of the fringe pattern, and \(\phi\) is the phase of the fringe pattern. The variables \(\alpha\) and \(\omega\) are dependent on the width of the slits \((w)\), the slit separation \((u)\), the
wavelength of the light ($\lambda$), and the focal length of the final lens ($f$) which focuses the image on the camera. They are defined as:

$$\alpha = \frac{w}{\lambda f}, \quad \omega = \frac{2\pi\mu}{\lambda f}$$

In order to measure the visibility, we created a one-dimensional intensity profile by averaging the central 100 columns of the fringe patterns, and then fitting them to equation 3 using a MATLAB program (see appendix). Sources of error included aperture width, which was accurate to ± 10 µm, and error in the fitting program. We estimated about ± 10% error from the program because when multiple images were taken without changing any settings, there was about ± 10% variation in the visibility. Figure 8 compares our experimental results with theory, where the solid line is attained from equation 2 using a distance $L = 28.5$ cm.
Figure 8: Visibility as a function of source width for theoretical and experimental results.
Figure 9: The upper images are the interference pattern corresponding to its graph below. You can see the visibility start high, then lower as the aperture size is increased, and then rise again as it is further increased. The points on the graphs are data points taken from the fringe patterns by averaging the central columns in the images, and the solid line is the representative graphs correspond to source aperture sizes of: a) 400 \( \mu \)m, b) 600 \( \mu \)m, c) 700 \( \mu \)m, and d) 900 \( \mu \)m

Conclusion

In this experiment, we have recreated the work of Thompson and Wolf in a new manner which utilizes more current technology and allows a large number of data points to be taken with ease. It also allows one to view the phase reversal of the fringes in real time as the aperture width is slowly changed. The data follows the expected visibility closely. It should be noted that since the source is very large compared to other experiments, it is a challenge to get uniform illumination of the slit aperture. Also,
great care must be taken to ensure that the all equipment is aligned properly, or the results will not be accurate.
References


Appendix

This MATLAB program was used to read in image files, average the central columns, and fit the intensity profile to equation 3. The fitting program uses the fminsearch routine that is included with MATLAB.

```matlab
% plot the profile of fringes
a = imread('..\..\mar_05_10\700mic.bmp');

x = 1:480;
x = x + 0.000000001;
xd = x;
% xdata = x(50:410);

ydata = sum(a(:, 250:400)) / 151;
ydata(1:10) = ones(1, 10) * ydata(20);
% ydata = ydata(50:410);

[estimates, model] = fit_fringe_profile(xdata, ydata)

plot(xdata, ydata, 'x', 'Linewidth', 2)
hold on
[sse, FittedCurve] = model(estimates);
plot(xdata, FittedCurve, 'r', 'Linewidth', 2)

xlabel('position (pixels)', 'FontSize', 16)
ylabel('Intensity (arb.)', 'FontSize', 16)
set(gca, 'FontSize', 13)
% title(['Fitting to function ', func2str(model)]);
% legend('data', ['fit using ', func2str(model)])
hold off

function [estimates, model] = fit_fringe_profile(xdata, ydata)
% Call fminsearch with a starting points close to parameters of the profile.
% A=20; the offset of the pattern
% B=100; the max amp of the fringes
% alph=0.01; the multiplier in the sinc2 function
% x0=250; the center of the fringes
% V=1; the visibility
% omeg=0.2; the frequency of the fringes
% phi=0; the phase

% start point is A, max amp, alph, center, vis, fringe freq, phase
start_point = [3 30 0.0041 230 0.09 0.16 pi];
model = @fringefun;
estimates = fminsearch(model, start_point);
    function [sse, FittedCurve] = fringefun(params)
```
A = params(1);
B = params(2);
alph = params(3);
x0 = params(4);
V = params(5);
omeg = params(6);
phi = params(7);
sinc2 = (sin(pi*alph*(xdata-x0))./(pi*alph*(xdata-x0))).^2;
FittedCurve = A + B*sinc2.*(1 + V*cos(omeg*(xdata-x0)+phi));
ErrorVector = FittedCurve - ydata;
sse = sum(ErrorVector .^ 2);