E-BEAM GENERATED HOLOGRAPHIC MASKS FOR OPTICAL VECTOR-MATRIX MULTIPLICATION+

Steven M. Arnold
Honeywell Corporate Technology Center
Bloomington, Minnesota 55420

Steven K. Case
Department of Electrical Engineering
University of Minnesota
Minneapolis, Minnesota 55455

SUMMARY

A proposed optical vector-matrix multiplication scheme encodes the matrix elements as a holographic mask consisting of linear diffraction gratings. The binary, chrome-on-glass masks are fabricated by e-beam lithography. This approach results in a fairly simple optical system that promises both large numerical range and high accuracy. A simple holographic mask has been fabricated and tested.

INTRODUCTION

There has recently been considerable interest in optical computing since it offers very high computation throughput rates for mathematical operations amenable to parallel computation. One class of such operations, vector-matrix multiplication, can be used for performing discrete Fourier transforms, coordinate transformations, pattern classification, and many other computations. The general vector-matrix multiplication may be written as:

$$y_m = \sum_{n=1}^{N} H_{mn} x_n \quad (m = 1,2,...M)$$

One optical approach to performing this computation uses N light sources to represent the components x_n of the input vector, M detectors to represent the components y_m of the output vector, and suitable optics to assure that a fraction H_{mn} of the light from source x_n gets to each detector y_m. The problem can be suitably scaled so that all parameters fall within acceptable positive ranges. Optics to perform the function of the matrix H generally will be fixed, while the sources are modulated to represent various input vectors x.

In principle, the performance of this optical computer is dependent on a number of considerations involving the optics, detectors and sources. In practice, numerical range and accuracy will usually be limited by matrix element imperfections, while speed will be limited by the amount of light reaching the detectors. For this reason, our work has focused on efficient optics to precisely distribute light among the various detectors.

In most schemes for optical vector-matrix multiplication, the matrix is encoded as a rectangular array of apertures or gray tones in a mask. This approach, represented in Figure 1, encounters several limitations. A complex optical system is required in order to illuminate and receive light from specific columns and rows of the matrix mask. Much light is discarded in

+This work was sponsored by the Air Force Office of Scientific Research under Contract No. F49620-80-C-0029.

providing uniform illumination to the mask, with the mask passing only about half of what remains. Numerical range and accuracy are limited by the space-bandwidth product of the mask (generally less than 10^6 with conventional plotting techniques). Small matrix elements result in small apertures with low relative accuracies. If results differ from those intended, it is generally difficult to modify a mask except by starting anew.

PCGH CONFIGURATION

Our approach to optical vector-matrix multiplication is based upon an e-beam generated diffractive mask which we call a partitioned computer generated hologram (PCGH) (see Figure 2). Each of N PCGHs is illuminated by collimated light from a single element of the source array and thus represents one column of the NxM matrix mask depicted in Figure 1. Each PCGH is partitioned into M linear gratings which diffract light to the M detectors. The optical power diffracted by a particular grating is made proportional to the value of the required matrix element.

Figure 3 illustrates a PCGH intended to produce 10 equal intensity outputs when uniformly illuminated. This PCGH contains 10 equal area gratings, each with its own spatial frequency. Facets are arranged symmetrically about the center to provide immunity to beam wander.

The PCGHs are fabricated as binary chrome-on-glass holograms where the pattern is delineated via e-beam lithography. A glass plate is first coated with a layer of chrome and a layer of e-beam resist. A pattern is exposed in the resist by e-beam direct writing and the resist is developed. The developed resist then serves as a mask for etching the pattern into the chrome.

Our e-beam PCGH optical vector-matrix multiplication scheme has several advantages over the scheme in Figure 1. E-beam lithography offers a higher space-bandwidth product, which can translate into greater numerical accuracy. Also, the PCGHs are in the Fourier plane of the transform lens with respect to the detectors. This means that the only requirement for light to reach a particular detector is that it be traveling in the right direction upon leaving the PCGH. Therefore, the input modules, consisting of source, collimating lens and PCGH, may be located anywhere within the aperture of the transform lens. This same immunity to shifts allows a PCGH to be partitioned into facets in any manner consistent with dividing up the available light amongst the various detectors (providing, of course, that the facets do not become too small). Other advantages relate to optical efficiency. All light striking the PCGH can be used. Light need not be wasted in achieving uniform illumination; non-uniform illumination is acceptable so long as its effects are accounted for in the partitioning. Small facets associated with lesser outputs can be made physically larger by placing them where PCGH illumination is lowest. The various considerations which go into the design of a PCGH are discussed in the following sections.

PCGH DIFFRACTION ANALYSIS

Each facet of the PCGH contains a linear grating to diffract incident light to the appropriate detector. The spatial frequencies of these linear gratings are determined by the system geometry. First-order diffracted light

from a facet of spatial frequency ν will be focused in the detector plane a distance $\nu\lambda F$ from the transform lens axis (Figure 1), where λ is the wavelength and F the transform lens focal length. Our design is for a 10-element linear detector array. This requires 10 equally spaced grating frequencies. The widest possible detector separation is achieved for grating frequencies $n\Delta\nu$, where $\Delta\nu$ is the frequency separation and n = 10, 11, 12, 13, 14, 15, 16, 17, 18, and 19. Then the unwanted harmonic frequencies from the square wave gratings begin at $20\Delta\nu$ and will not coincide with the desired outputs.

MODULAR FACETS

The matrix values are encoded into the PCGH via grating area modulation. The hologram must therefore be divided into facets such that the amount of light diffracted by a facet to its detector is proportional to the required matrix element. Various considerations lead us to partition the PCGH into facets along a square grid.

Mathematically, the transmittance of a facet can be regarded as the product of its aperture and an infinite linear grating. By the Fourier convolution theory, the diffraction pattern of this facet is the diffraction pattern of its aperture convolved with the delta function from the infinite linear grating. In other words, the effect of the linear grating is to shift the location of the diffraction pattern of the facet aperture.

Figure 4 indicates the diffraction pattern due to a square aperture of dimension D. The main lobe has a width $2\lambda F/D$ and contains 81.5% of the energy passing through the aperture, which we label as 0 dB. The sidelobes form a rectangular array with sidelobe energy diminishing inversely as the square of the distance from either axis. The figure indicates the energies (in dB) of the various sidelobes relative to the main lobe. It is seen that their energies diminish most rapidly along a diagonal.

Crosstalk between channels depends on how these diffraction patterns overlap adjacent detectors. As indicated earlier, the separation of detectors is $\Delta\nu\lambda F$. As facets are made smaller, their diffraction patterns become larger, requiring higher grating frequencies to separate them. For this reason, we impose a minimum square facet size, based on detector separation and an acceptable level of crosstalk. We size our detector aperture to capture only the main lobe of this minimum facet diffraction pattern. A larger aperture would capture lesser lobes of the channel of interest, but also some greater lobes of adjacent channels, thereby degrading the signal-to-noise ratio. Also, zero intensity at the edges of the detector aperture eases mechanical tolerances.

In the diffraction pattern of Figure 4 the square box in the center represents the detector aperture. Similar boxes are used to indicate possible locations of adjacent detectors. These have been separated diagonally to use the more rapid sidelobe decay (implying grating fringes which run diagonally within the facets). A separation of $2\lambda F/D$ in each dimension results in 29 dB of crosstalk with some very bright axial sidelobes just outside the detector aperture. A separation of $3\lambda F/D$ yields a much more comfortable 37 dB crosstalk. With this as our choice, the minimum facet size will be $D = 3\sqrt{2}/\Delta\nu$.

The above discussion of crosstalk implies that each facet is a minimum square facet. This would be a severe constraint on system numerical range. In practice, we form the facet for each channel from many minimum sized modular subfacets. Therefore, the diffraction pattern for a single channel is not that of a single subfacet, but rather the pattern is due to the aperture consisting of all subfacets for that channel. Crosstalk can be minimized by clustering all required subfacets of a given spatial frequency into one or two large facets with minimum perimeter, such as was done in the PCGH of Figure 3.

To have the greatest flexibility in partitioning the PCGH, we would like the modular subfacets to be as small and numerous as possible. Since they can be no smaller than $D = 3\sqrt{2}/\Delta\nu$ (with our 37 dB crosstalk limit), we want a large spatial frequency separation $\Delta\nu$. However, a large spatial frequency separation implies large spatial frequencies and hence small grating periods. If the grating period becomes comparable in size to the e-beam spot size, considerable grating duty cycle errors with corresponding diffraction efficiency errors will result. For these reasons we have elected to start with grating frequencies of 60 to 114 lp/mm (measured along either axis) and 0.5 mm subfacets. This gives us 400 subfacets in a 1 cm x 1 cm PCGH and a maximum numerical accuracy of about 20 dB. With Gaussian illumination, the effective number of subfacets is extended to more than 5000 since the corner facets have about 6% the intensity of illumination of the central facets.

TRIMMING METHODS

To extend our numerical accuracy beyond about 20 dB, it is necessary to employ a separate lithography step to adjust the relative amplitudes of the outputs. The ability to do this trimming is one of the chief advantages of e-beam lithography and the PCGH. There are several possible methods of trimming a PCGH once it has been made and tested. The best trimming method is to add a negative facet, that is to say, a facet exactly out of phase with the existing facet. This is much easier than it might sound: the negative facet can be written in space already occupied by the existing positive facet, leaving a completely open area. Linewidth and phase problems are overcome since, if we write over the whole area, the phase and duty cycle will be exactly the complement of what was already there. The easiest and probably best way to trim, because it minimizes registration requirements, is to draw a square box of the appropriate size somewhere near the middle of the existing facet.

PARTITIONING

The task of partitioning the PCGH to achieve the correct relative outputs is greatly simplified by the decision to adopt module facets. The intensities of rectangular facets are easy to compute for either uniform or Gaussian illumination. We need only determine which subfacets are to be assigned to which channels. We do this by means of an algorithm which maximizes accuracy and minimizes crosstalk. Each subfacet is assigned entirely to one channel. Trimming is then used to achieve the desired numerical accuracy. With a 1 cm x 1 cm PCGH, 0.5 mm subfacets and a Gaussian illumination diminishing to e^{-2} at the edges, a subfacet in the corner will receive only 0.00019 of the total energy. Therefore, it is possible to obtain as much as 37 dB of numerical range, with no channel receiving less than a full subfacet.

EXPERIMENTAL RESULTS AND DISCUSSION

The experimental setup for demonstation and evaluation of PCGHs is indicated in Figure 5. Light from a He-Ne laser is passed through a pinhole spatial filter to create a point source which is then imaged in the detector plane by a lens. The PCGH is kinematically mounted in a micropositioner placed immediately after this lens. This mount is designed to allow removal for trimming and subsequent replacement of the PCGH without disturbing the alignment. The distance from the PCGH to the detector plane is 50 cm. In the detector plane, we have a single UDT-455 photodetector mounted on a motorized translation stage. One of several square apertures, selectable in size from 1 to 2 mm, is placed in front of this detector to define and limit its effective area. A second identical photodiode directly monitors the laser output. By ratioing the signals of these two detectors, we eliminate a number of potential noise sources, including laser power fluctuations. The two detector signals can be fed into a logarithmic amplifier for strip chart traces and rough measurements, or else the signals can be measured directly when the greatest accuracy is required.

Our experimental work thus far has been aimed chiefly at constructing experimental apparatus capable of yielding the precise measurements needed for verifying the 40 dB numerical accuracy and/or 40 dB numerical range we seek. In this regard, we consider the setup as described to be adequate with only minor refinements.

We have fabricated the PCGH of Figure 3. This hologram utilizes diagonally separated outputs, compact facets and symmetry about the axis. Figure 6 shows the resultant diffraction pattern at the output plane. Since the grating fringes are written at 45° relative to the facet boundaries, the many on-axis sidelobes from each output are seen to be diagonally separated from the adjacent output signals. Figure 7 is an intensity trace through the centers of the outputs. The spacing of the outputs was selected to give 37 dB worst case crosstalk between adjacent channels.

In light of our preliminary measurements, the PCGH method of optical computing appears capable of high numerical accuracy over a wide numerical range.

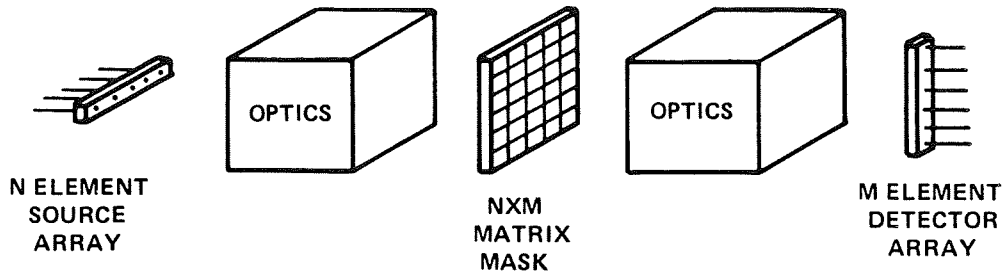

Figure 1.- A general scheme for optical vector-matrix multipliation.

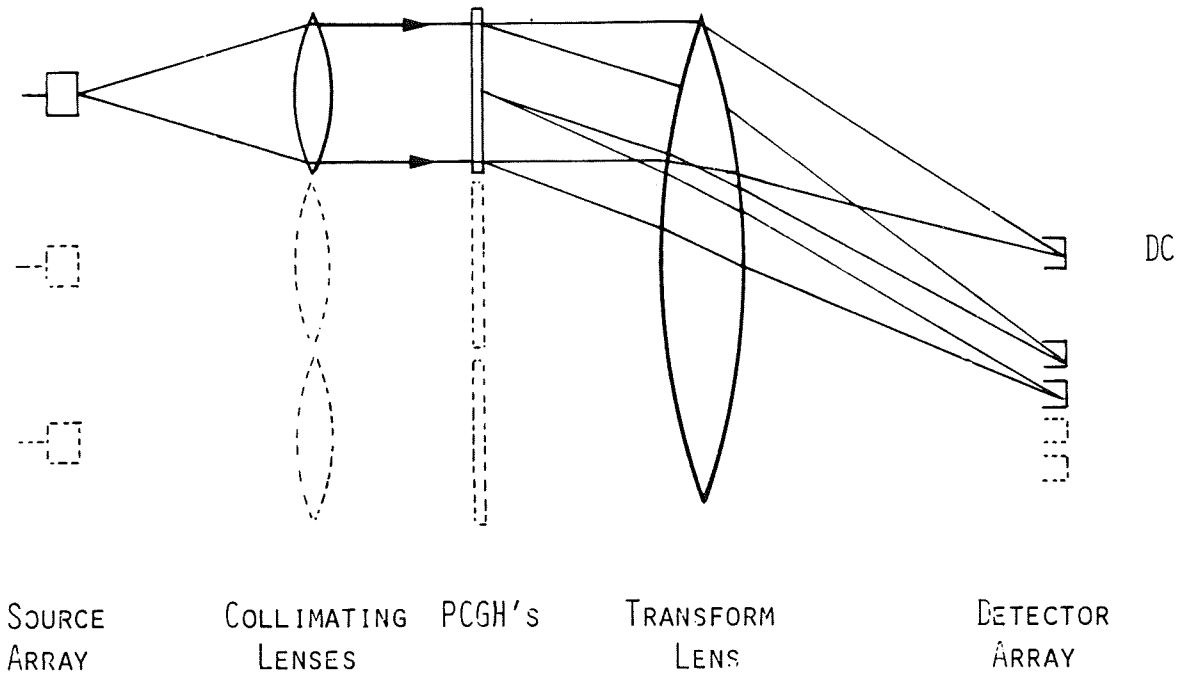

Figure 2.- PCGH vector-matrix multiplier.

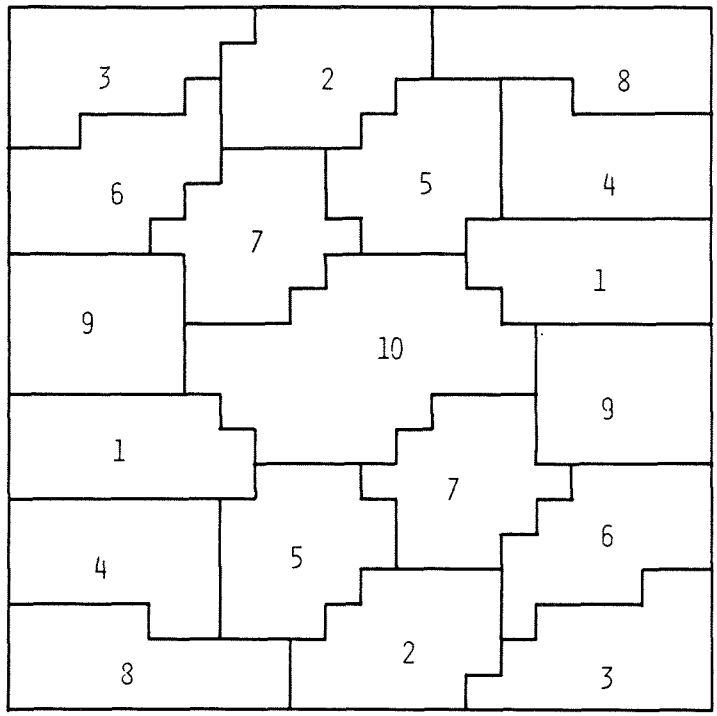

Figure 3.- PCGH design showing facet boundaries. (Fringes within facets not shown.)

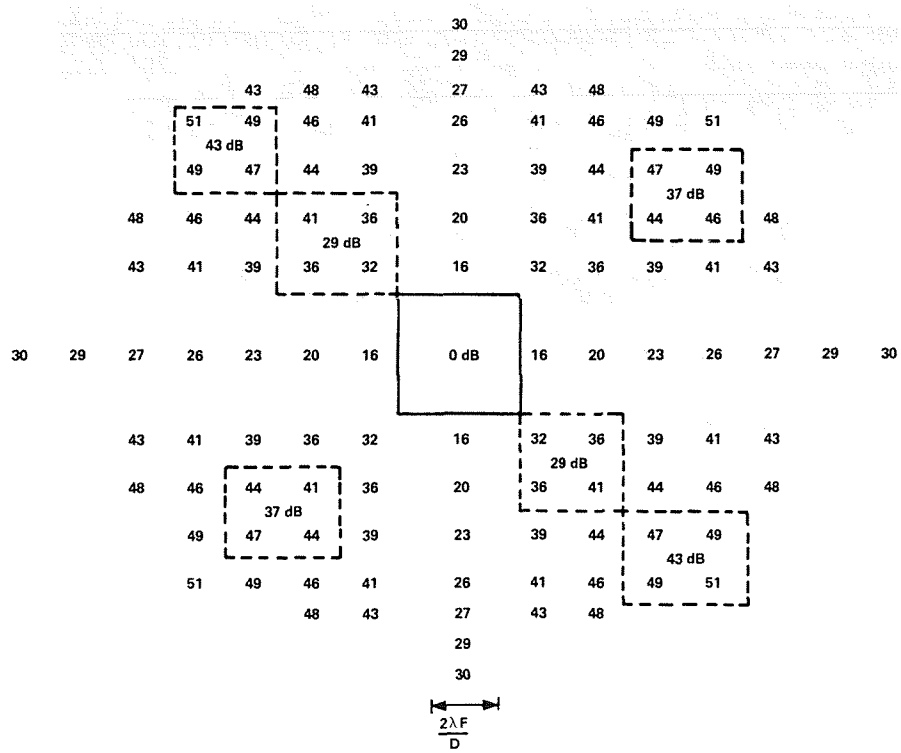

Figure 4.- Crosstalk from square subfacets.

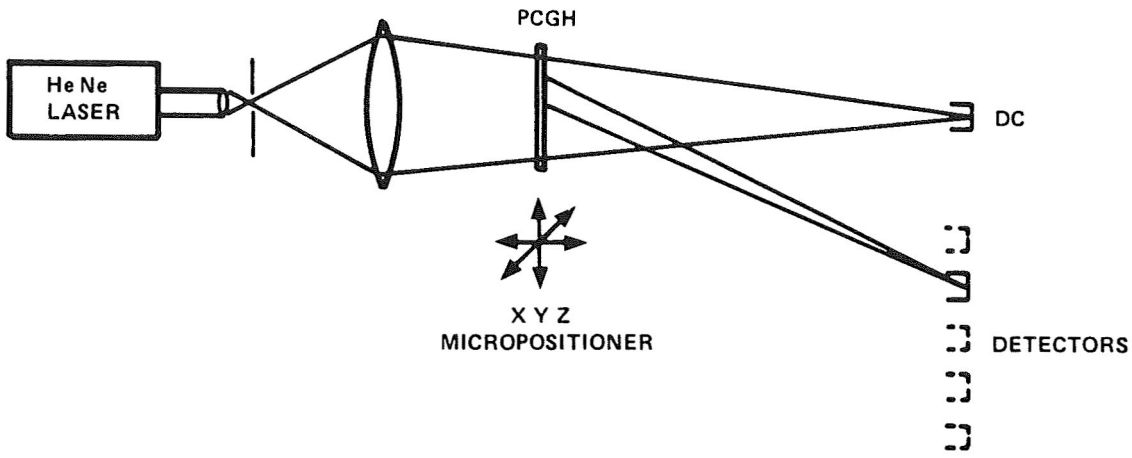

Figure 5.- Experimental setup for demonstration and evaluation of PCGHs.

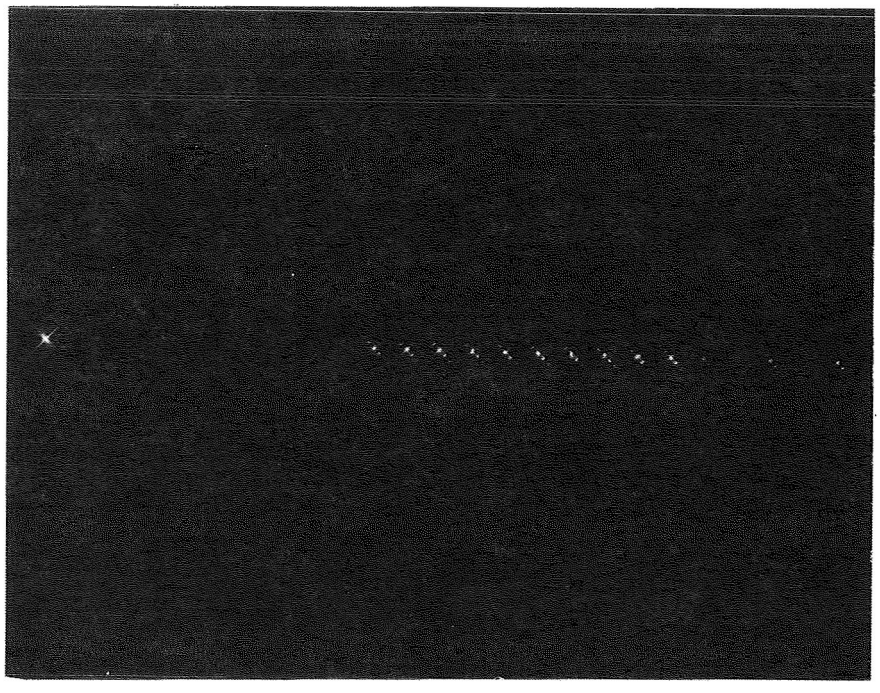

Figure 6.- Diffraction pattern from the PCGH. The undiffracted beam, all 10 first-order diffracted beams, and the first few second-order beams are visible.

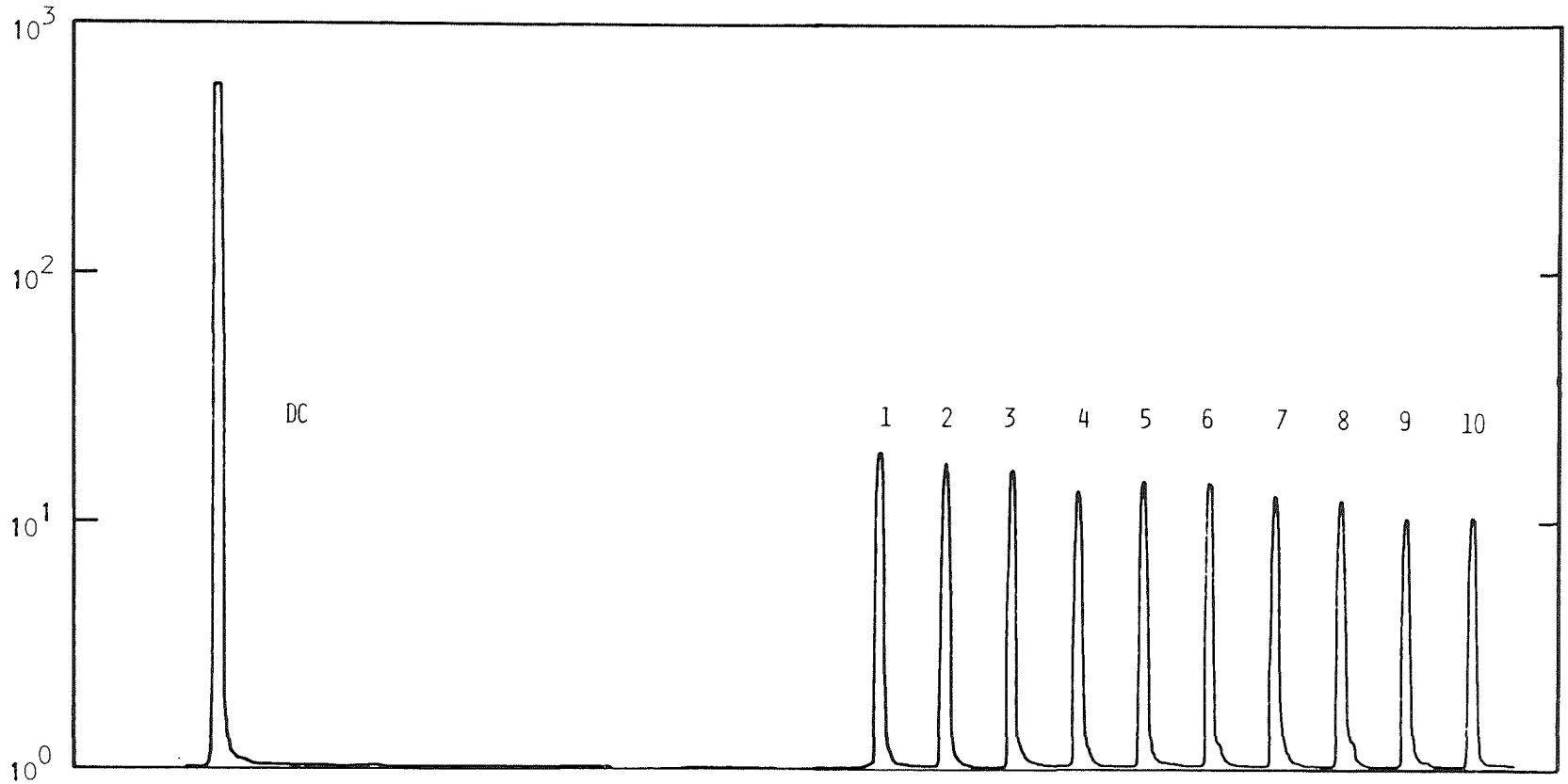

Figure 7.- Scan through diffraction pattern from the PCGH.