ERIM
FORMERLY WILLOW RUN LABORATORIES, THE UNIVERSITY OF MICHIGAN

3 August 1976

MEMORANDUM TO: VLA Optical Processor File (123401)
FROM: James R. Fienup J. F.
SUBJECT: Topics of Discussion with Dr. Lewis Somers on July 20, 1976

In our discussion, primarily concerned with data encoding and formatting for the VLA Optical Processor, the following topics were discussed

1. b-t to t-b sort
2. Encoding techniques
    - 2a. Major types
    - 2b. Major conclusions about encoding methods
3. Some trade-offs
    - 3a. Materials
    - 3b. Diffraction efficiency
    - 3c. Noise inherent in encoding methods
    - 3d. Computational considerations
4. Miscellaneous items

1. b-t to t-b sort. It would be highly desirable to have sorted the data into t-b order when recording on film. It may be possible to have the b-t ordering of data to be resorted to the t-b order in real time as the data is first being written on the disk, without requiring an additional digital processing step and significant additional digital

equipment, by writing successive pieces of data on different tracks. However, since the possibility of such a real-time sort is in doubt, we will continue to study the mechanics of writing data in the b-t order, but as an item of secondary priority.

2. Encoding techniques.

2a. Major types. Three radically different encoding techniques were identified as the major contenders. Let the complex visibility function be given by

$$V(u,v) = |V(u,v)|e^{j\phi(u,v)} = R(u,v) + j\ I(u,v)$$

where $|V(u,v)|$ and $\phi(u,v)$ are the modulus (or amplitude) and phase of $V(u,v)$, respectively, and $R(u,v)$ and $I(u,v)$ are the real and imaginary components of $V(u,v)$, respectively. One encoding technique is that of a <u>simple carrier</u>, using a transparency with amplitude transmittance proportional to

$$H(u,v) = B_o+2|V(u,v)|\cos\left[\omega_o u+\phi(u,v)\right]$$

$$= B_o + V(u,v)e^{j\omega_o u} + V^*(u,v)e^{-j\omega_o u}$$

where B_o is a bias term greater than the maximum of $2|V(u,v)|$ required to make the transmittance non-negative, and ω_o is a carrier frequency required to separate the desired image, $V(u,v)$, from the conjugate image, $V^*(u,v)$, and from the undiffracted term, B_o. A second encoding technique is the real-imaginary technique in which two transparencies are made, one with amplitude transmittance

$$H_R(u,v) = B_o + R(u,v)$$

and the other with

$$H_I(u,v) = B_o + I(u,v)$$

where again B_o is a bias term. A third contender is the Lohmann binary detour-phase technique which consists of a set of apertures of unit transmittance on a black background, the area of each aperture being proportional to $|V(u,v)|$ and the position of each aperture in the u-dimension relative to a specific raster of basis points being proportional to $\phi(u,v)$. When an error-compensation technique is employed, the positions of the Lohmann apertures coincide with the peaks of the function $\cos\left[\omega_o u + \phi(u,v)\right]$.

Other encoding techniques were mentioned but were rejected as incapable of providing the required encoding accuracy. For example, Fienup and Chu's ROACH technique can theoretically provide perfect accuracy, but existing photo-sensitive materials suitable for the ROACH suffer from substantial film-grain noise and inter-layer cross-talk. The kinoform, Chu's synthetic coefficient hologram, Lee's continuous-tone detour-phase hologram, and a number of types of binary detour-phase holograms suffer from various inherent sources of noise that would eliminate them from contention.

2b. Major conclusions about encoding methods. The following conclusions were reached, but should be considered to be tentative.

1. The simple carrier method is the most promising.
2. Not requiring a carrier, the real-imaginary method would be preferable if the space-bandwidth product of the recorder is the dominant constraint. This would be done at the expense of a very bright spot in the center of the image.
3. Not requiring a continuous-tone material, the Lohmann detour-phase method would be preferred if the sensitometry could not be controlled with sufficient accuracy or if film-grain noise were too severe. This would be done at the expense of a larger space-bandwidth-product requirement on the recorder and additional noise inherent in the encoding method.

3. Some tradeoffs. The following tradeoffs were discussed; however, they represent only a small number of the possible considerations.

3a. Materials. Pure-phase materials can be used for the encoding methods discussed and have high diffraction efficiency and low film grain noise, but are less desirable than amplitude (absorbing) materials because they introduce higher-order spurious terms that overlap the desired image. The use of a positive-working amplitude material results in opaque areas where there are gaps between tracks, whereas negative materials result in clear areas. Consequently, for negative materials, an additional term $1 - A(u,v)$ is present, where $A(u,v) = \{1$, for areas covered by a track; 0, for areas between tracks$\}$. This results in additional terms $\delta(x,y)-a(x,y)$ in the image, where $a(x,y) = \mathcal{F}\{A(u,v)\}$ is the impulse response due to the aperture of ellipses. There already exists a term $B_o\, a(x,y) = \mathcal{F}\{B_o A(u,v)\}$, where B_o is the bias term required to make the transmittance non-negative. Therefore, for negative materials the sum of on-axis terms is $\delta(x,y) - (1-B_o)a(x,y)$. Thus, the intensity $(1-B_o)^2|a(x,y)|^2$ is minimum for a bias amplitude transmittance $B_o = 0.5$ and increases as B_o departs from 0.5. For the simple carrier method of encoding, a somewhat greater carrier frequency would be required with a negative material than with a positive material in order to keep the image a safe distance from the on-axis terms. For the real-imaginary method, a positive material is even more necessary since the on-axis terms are in the center of the image. On the other hand, the processing of negative materials is more direct than that of positive materials, possibly resulting in more accurate sensitometry.

Binary detour-phase encoding requires only clear and opaque areas, allowing for high-contrast materials and greatly simplified sensitometry. Since film grain noise is at a minimum both for very high and for very low transmittances, it is not a problem for a binary encoding.

3b. Diffraction efficiency. A higher diffraction efficiency is desirable because it results in a higher signal-to-noise ratio (since more signal is present) and reduces the integration time required of the output detector. The diffraction efficiency for a given image is given by $\eta_m\overline{|V|^2}/|V|^2_{max}$ where $\overline{|V|^2}$ is the average of the modulus-squared of the visibility function, $|V|^2_{max}$ is the maximum value of $|V|^2$, and η_m is the theoretical maximum diffraction efficiency for the encoding method (η_m is arrived at by assuming that the image consists of a single point source). While η_m gives the percentage of light intensity going into the desired image, we may also be interested in $\sqrt{\eta_m}$ which is porportional to the amplitude of the image. For the simple carrier method, $\eta_m = 6.25\%$ ($\sqrt{\eta_m} = 25\%$); for the real-imaginary method, $\eta_m = 25\%$ ($\sqrt{\eta_m} = 50\%$), whether the real and imaginary components are read out singly or combined in an interferometer (but not counting a probably 50% loss due to reflection or diffraction during the interferometric combination of the two); and for the Lohmann method, $\eta_m = 1/\pi^2 \simeq 10.13\%$ ($\sqrt{\eta_m} = 31.8\%$).

3c. Noise inherent in encoding methods. Assuming ideal conditions (perfect materials, sensitometry, recording devices, etc.), the simple carrier will produce a perfect image plus an undiffracted beam, so the only inherent noise is that due to the sidelobes of the undiffracted beam (i.e., the impulse response of the system centered at the optical axis). By using a carrier frequency somewhat greater than the minimum required to separate the desired image from its conjugate, the desired image can be moved sufficiently far from the undiffracted beam so that those sidelobes are of no consequence. This is done, however, at the expense of increasing the space-bandwidth product and accuracy requirements of the recorder. Under ideal conditions the real-imaginary method also produces a perfect image plus an undiffracted beam, but then the undiffracted beam is in the center of the image and wipes out a substantial number of picture elements in the image. If a carrier it to be used to move the desired image off-axis, then there is no advantage to the real-imaginary concept and it is better to use the simple carrier method. The Lohmann method produces the ideal image plus an undiffracted beam plus a number of other spurious terms, some of which do overlap the desired image. The errors inherent in the Lohmann method, which are substantial under some conditions, have been discussed in the literature.

The effect of the sidelobes of the undiffracted beam on the desired image depends both on the sidelobe levels and on the maximum amplitude in the image. The form of the aperture function determines the sidelobe structure of the impulse response of the system. Sidelobe levels can be decreased by an appropriate weighting of the aperture. The maximum value of the image amplitude depends on the diffraction efficiency and on the nature of the image. An image of a few point-like stars will have much higher maximum values than will an image of a large extended object, and so will be more resistant to sidelobes from the undiffracted beam.

3d. Computational considerations. The real-imaginary technique requires a search for the maximum value of V(x,y) to determine the proper scaling of the data, a scaling of the data according to the maximum value, and the addition of a bias. The simple carrier method requires the above computations plus a conversion of real-imaginary to amplitude-phase with linear phase offset. The Lohmann method requires all the above plus an interpolation step that may be iterative.

4. Miscellaneous items. A few other topics were briefly discussed, including sources of errors, effect of track overlap for different encoding methods, methods of measuring phase errors, the relationship between recorder positional accuracy and phase errors, and time-reference and space-reference recorder calibration methods.

JRF:sd

cc: C. Aleksoff
M. Carter
I. Cindrich
M. Hidayet
A. Klooster