Biol. Cybern. 65, 113–119 (1991)

Biological Cybernetics

The relative importance of local phase and local amplitude in patchwise image reconstruction

M. J. Morgan [1], **J. Ross** [2], **and A. Hayes** [3]

[1] Department of Pharmacology, University of Edinburgh, Edinburgh EH9 9JZ, United Kingdom
[2] Department of Psychology, University of Western Australia, Perth, Australia
[3] Department of Psychology and Centre for Cognitive and Computational Neuroscience, University of Stirling, Stirling, United Kingdom

Received November 26, 1990/Accepted in revised form April 15, 1991

Abstract. Natural images were subjected to patchwise Fourier analysis, and the local amplitude and phase spectra were swapped between different images. When the patches were large relative to the image size, the appearance of the reconstructed image was similar to that of the image from which the phase information had been derived, in agreement with previous reports of phase-dominance in the global Fourier Transform. However, when the patch size was made sufficiently small, the appearance of reconstructed images was dominated by amplitude rather than phase. This was not simply due to the DC component of the amplitude spectrum. Prior low-pass filtering of the images enhanced the dominance of amplitude information in the patchwise transform. We conclude that patchwise-reconstructed images contain two quite distinct kinds of information for the human observer. The first is the positional information ("local sign") of the patches themselves; the second is the textural information within patches, which is dominated by amplitude rather than phase. The reason why the global Fourier Transform is dominated by phase is that in the absence of any other information about local sign, phase is necessary to reconstruct localised features such as edges.

Introduction

In a number of different contexts, ranging from X-ray crystallography to perception of human faces, it has been found that objects and people can still be recognised in a Fourier reconstruction despite defective amplitude information (Oppenheim and Lim 1981; Piotrowski and Campbell 1982; Shapley et al. 1990). Furthermore, phase information prevails over amplitude. For example, if the amplitude and phase spectra of two images are interchanged, the resulting hybrid images are recognizably those from which the phase spectrum is taken. This is true even if the two original images contain objects as different as a face and a military vehicle. Phase can even be severely quantized and still dictate the form of a reconstructed image, hybrid or otherwise (Piotrowski and Campbell 1982).

To the extent that the similarity of two images depends on the similarities of the features an observer can discern in both, the relatively greater importance of phase information is easy to understand. The amplitude spectrum reflects only the first and second order statistics of images, which fail to capture the redundant structures of images (Julesz 1980; Field 1987). Phase information is needed to fix features and to locate textures. Consider sharp edges, for example. All Fourier components are locally in the same phase (0 or 180 deg in a local sine expansion) at an edge. The amplitudes of the Fourier components in an edge can be randomly altered without having much effect upon the sharpness of edges or their locations (Morrone and Burr 1988), but if the global phase of the components is completely randomised, the local congruence is lost and so the edge is lost. Similar considerations apply to even-symmetric image features (bars) and to features of other types, such as shadow edges. Global phase information is required also to locate a patch of texture in the correct part of an image. It is thus not particularly surprising that the visual interpretation of images depends crucially upon the phase spectrum of its Fourier components.

We cannot legitimately conclude from these facts that amplitude information is unimportant. The published examples of amplitude-phase swapping all show considerable distortions in images reconstructed from phase. These distortions are perhaps best described as alterations of texture and shading. Localized features are preserved, but more subtle gradations of tone and texture tend to be distorted or lost. This suggests that both phase and amplitude information have a role to play in the perception of images, the former for reproducing features and locating textures, and the latter for reproducing shading and textures of the correct kind.

In all examples in the literature the global Fourier transform has been employed. It is inevitable that information about features and the position of texture in the

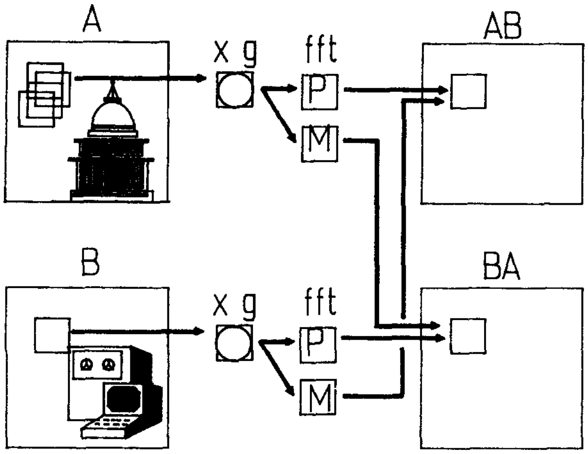

Fig. 1. The figures shows schematically the procedure used for patchwise phase-amplitude swapping between two images. A and B are the original images. These are multiplied locally by a Gaussian windowing function ($\times$ g) and the resulting patches are subjected to a Fast Fourier Transform (fft). The phase (P) and magnitude (M) components of the transform are cross-combined and subjected to an inverse FFT to yield the images AB, which has the phase of A and the magnitude of B, and BA which has the phase of B and the magnitude of A. For further explanation see the text

image providing amplitude is lost, since its Fourier components are globally repositioned, either randomly or according to a rule derived from an entirely different image. By employing a patchwise Fourier analysis of the image (Robson 1983), some positional information can be preserved, independently of phase. In a patchwise analysis the image is divided into a number of overlapping patches, each independently analyzed into its phase and amplitude components. The feasibility of image reconstruction from a local polar (amplitude and phase) representation has been demonstrated, and it has the advantage for optimal coding that low amplitude components can be given a coarse phase representation (Wegmann and Zetzsche 1990). We now consider the extent to which it is possible to reconstruct the image from phase or from amplitude information alone within patches. Figure 1 presents a schematic description of this procedure. The important point of patchwise analysis is that positional information is to a greater or lesser degree carried by the position of the patches, depending upon their size. The smaller the patch, the greater the amount of positional information carried by its location within the image. In these circumstances it is plausible that phase information within a patch will be less important for carrying positional information than is the case with the global transform. Conversely, it is plausible that amplitude information will be revealed as relatively more important.

It may be helpful to clarify this reasoning by considering some limiting cases. When the patch is equal in size to the whole image we simply have the case of the global transform, in which it is known that amplitude information by itself is virtually useless in reconstructing the image. At the other extreme, when the patch size is a single pixel, the only information in the patchwise Fourier transform is its DC component. It is a tautology that an image can be perfectly reconstructed from its single-pixel intensity values, so it follows that with a patch size of a single pixel, reconstruction by amplitude is perfect. Somewhere between these two extremes phase must give way to amplitude. Since the critical patch size for this transition cannot easily be deduced from physical principles alone, we decided to investigate the issue empirically by carrying out patchwise phase-amplitude swapping between corresponding patches in two different images. The size of the patch relative to the image was systematically altered, and we used images of two different sizes to determine whether absolute or relative size of patch is the more important variable.

Method

Image processing was carried out on a SUN 3/100G computer, with a Data Translation frame grabber, using the HIPS image processing software (Landy et al. 1984). Images were acquired either at 512×512 resolution through the frame grabber and subsequently reduced by pixel averaging to 128×128 pixels; or were obtained from digitized photographs of faces at 512×512 resolution, reduced to 128×128 or to 64×64 pixels. The nominal grey-level range was 8 bits. Before amplitude-phase swapping was carried out between two images, the mean and variance of their grey levels was equated as far as possible by DC addition and scaling.

Images of size $p \times p$ pixels were broken down into square patches of size $w \times w$ centred on a square sampling grid of periodicity $w/4$. The patches thus overlapped by $3w/4$ pixels. Patch dimension w varied from a maximum of $p/2$ down to a minimum of 4, the smallest dimension permitting an overlap of $3w/4$ pixels. Each patch was then multiplied point-by-point with a 2-D Gaussian window of standard deviation $w/4$ pixels. A new image could then be constructed by adding together all the patches, each having the same centre position in the new arrays as it had in the orignal image array. We verified that the original images were adequately reconstructed by this method, both by inspecting the reconstructed images and by calculating that the method introduces a DC ripple of no more than 4 per cent in the worst case.

Finally, as is illustrated schematically in Fig. 1, two different images having been divided into patches in parallel as just described, and Gaussian-windowed, each patch was subjected to an FFT. The amplitude and phase spectra of corresponding patches in the two images were then swapped. Two new images were reconstructed from the two sets of overlapping patches, one having amplitude from Image A and phase from Image B, the other having amplitude from B and phase from A.

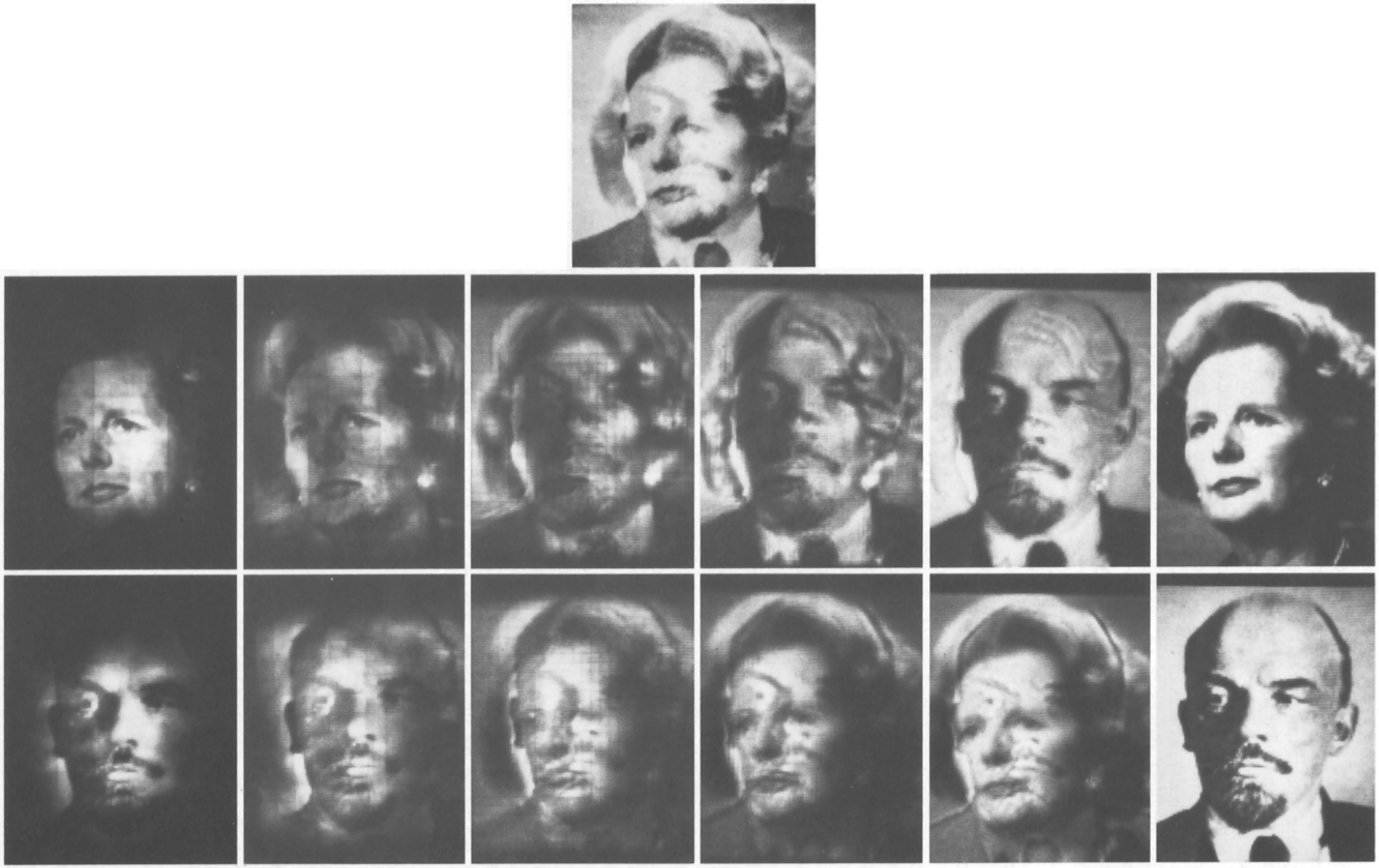

Fig. 2. The figure shows the results of patchwise amplitude-phase swapping between two images, as described in the text. The image in the top row is the sum of the two face images on the right of Rows 2 and 3. The images on the right-hand end of Rows 2 and 3 are the originals which contributed their phase information to the remaining images in the row. The amplitude information in each image was provided by the original image that did not provide the phase information. Thus, each image other than the originals obtains its phase information from one of the originals, and its amplitude from the other original. Amplitude and phase information were derived, *not from the global Fourier transform, but from within* overlapping localized image patches, as explained in the text. Along each of the two two rows, the patch size was varied. Going from left to right along the row, the patch sizes were: 64, 32, 16, 8 and 4 pixels. The original images are of size 128 × 128 pixels. It will be seen that with large patchsizes, phase information dominates the appearance of the images, while with the smallest patchsizes amplitude information becomes increasingly dominant

Results and discussion

The results for a pair of faces are shown in Fig. 2. The terminal image in each sequence is the image from which the patchwise phase spectrum was taken, the amplitude spectrum coming from the terminal image of the other sequence. Each image was 128 × 128 pixels in size; patch size varied from 64 × 64 down to 4 × 4. In each sequence reconstruction was dominated by phase at the largest patchsize (64 pixels), as it is in the global transform, but with the smallest patchsize (4 pixels) appearance was dominated by amplitude. At a patchsize of 16 pixels, phase and amplitude are equally matched in influence, with aspects of both images appearing, somewhat rivalrously. There is also ambiguity and some rivalry at a patchsize of 8 pixels, but phase tends to prevail. There is thus a crossover from phase dominance to amplitude dominance in determining the appearance of the hybrid image at a patchsize somewhere between 8 and 16 pixels. None of the hybrid images resembles the composite image formed by simple addition of the two originals. In the composite image both originals appear, each diluted in contrast by one half, and rivalling one another. Attention can be concentrated on one face or the other. In the hybrid image, the face from which amplitude was taken is not visible at larger patch sizes. As patch size reduces it becomes visible, but by then only sketch-like traces remain of the originally dominant face (that providing phase information).

To quantify the similarities between the mixed images and the originals, we correlated each of the two originals with each of the mixed images over a range of patch sizes. The results are shown in Fig. 3. Filled symbols represent correlations in the case where the mixed image has the same amplitude as the original with which it was being correlated; open symbols refer to the same-phase case. It will seen from this figure that the similarity of an image to one containing only its amplitude information decreases as the patch size is made larger. Conversely, the similarity of an image to one containing only its phase increases as the patch size is made larger. The correlation between the two original images was small (0.02). These findings confirm the informal impressions gained from viewing the figures. Figure 3 also shows the results of correlation from a different set of images (cars and trees) illustrated in Fig. 4. In this case also there was a high correlation between images and mixtures containing their amplitude information, and this correlation declined as patch size was increased. However, in this case there was little systematic effect of patchsize upon the similarity between like-phase images, despite visual appearances to the contrary (Fig. 4).

There is evidence that face-recognition ability of human observers is tied to the spatial frequency content of the image of the face viewed (Hayes et al. 1986). To see whether the limiting patchsize was proportional to the size of the face, rather than dependent on absolute size in pixels, we repeated the experiment with faces of dimensions 64 × 64 pixels. The results were, broadly speaking, the same as for the larger faces, except for a scaling factor depending on original image size. The crossover patchsize is closer to 8 pixels than to 16 in the smaller image series. Thus the limit at which magnitude information produces the same degree of reconstruction as phase seems to correspond to a fixed proportion of original size, rather than to an absolute patchsize.

As a check on the patchsize at which the crossover from phase to amplitude dominance occurred we free-fused original and reconstructed images. Binocular rivalry was observed for patchsizes below 16, but not above, for 128 × 128 images, and for patchsize below 8 for the 64 × 64 images.

Some features of the image that contributes phase survive even at the smallest patch size, 4 pixels. The outline of Thatcher's hair is visible in the penultimate image in the top sequence; so too is the outline of Lenin's moustache in the bottom sequence. This survival of features due only to phase, and independently

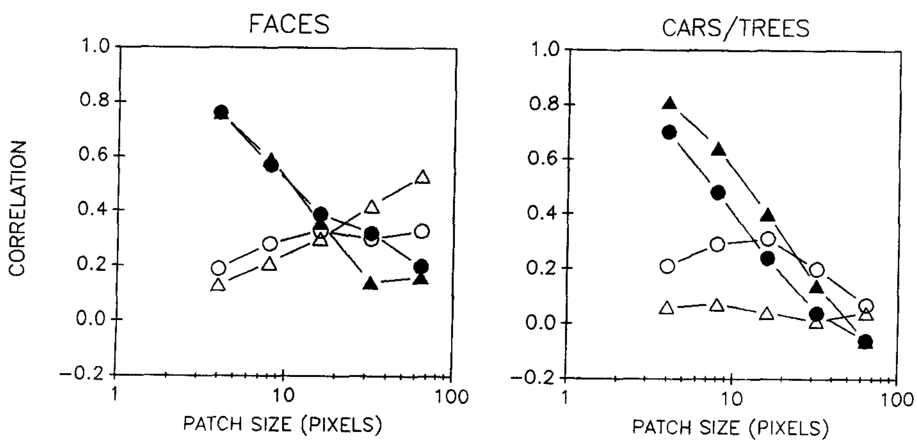

Fig. 3. The figures show the results of correlating the original image with hybrid phase-amplitude images at a range of patch sizes. Solid symbols refer to correlations between an image and a hybrid containing its amplitude information; open symbols refer to correlations between an image and a hybrid containing its phase information. Left hand figure: Faces. (*Circles*: Lenin's phase and Thatcher's amplitude; *Triangles*, the reverse). Right-hand figure: Trees and cars (see Fig. 4)

Fig. 4. The figure shows the results of patchwise amplitude-phase swapping between two images, as described in the text. The two original images are on the left. In the middle are the results with a patchsize of 64, on the right results with a patchsize of 4. Images in the same row have the same local phase information. The original images are of dimensions 128 × 128

of amplitude, was confirmed by a local-energy analysis, locating points of high phase congruence. The outlines of hair, eyebrows and other facial features from the phase-contributing image were evident in the local-energy map, even at the smallest patch size.

Within a patch of size 4 × 4 pixels, only two Fourier components of the original image can be sampled, and within a patch of size 8 × 8, only four. Such limited sampling can preserve only the two or four components of highest frequency from the original image, and those few only at a few orientations. There will nevertheless be peaks in the local-energy map of even the smallest patches, where components are congruent in phase. These peaks will be preserved when the amplitude spectrum of a corresponding patch from another image is substituted for the original amplitude spectrum of the patch. Thus some micro-features within a patch will survive a change in the amplitude spectrum, and alignments of micro-features from within different patches will outline macro-features from the face from which the phase spectra of patches were taken, such as Lenin's moustache. The smaller the patch size, the sparser the surviving features dictated by phase congruence should become, and they should appear more high-pass filtered as fewer components from the lower end of the frequency spectrum are available to describe them.

It will also be noticed that to some extent the appearance of the reconstructed image depends upon viewing distance. As viewing distance is increased the influence of amplitude becomes more pronounced, particularly with the smaller patch sizes. Considerations of local energy would suggest that features would survive in small patches only to the extent that components of high spatial frequency remain visible. This being so, we would expect amplitude to be made more dominant by low-pass filtering the images before patchwise analysis, because such images will lack sufficient high frequency energy for phase to create edges and other features. An example of this procedure is illustrated by Fig. 5, which reveals that the influence of amplitude is indeed enhanced by prior low-pass filtering. The high-pass filtered edges found in the broad-band images at small patch sizes (Fig. 2) are absent after low-pass filtering of the original image.

It might be argued, from the dominance of phase in global reproduction, that the amplitude spectrum contributes no more than the DC value of a patch. At small patchsizes DC sampling alone may be sufficient to

Fig. 5. The figure illustrates the results of low-pass filtering before carrying out the patchwise phase-amplitude swapping procedure described in Figs. 1 and 2. The two images on the left are low-pass filtered versions of the originals in Fig. 2. Those on the right are obtained by phase-amplitude swapping with a patchsize of 4 pixels. The originals on the left contribute phase information to the image to their right on the same row, and their amplitude to the image in the other row. Note the absence of intrusion by edges from the image contributing phase, which is seen in the full-bandwidth images of Fig. 2

Fig. 6. The figure demonstrates that the dominance of amplitude information with small patch sizes is not solely due to the DC level in the patch. The original image contributing phase information is shown in the middle. Phase information was provided by the other face in Fig. 2. The top image was obtained by phase-amplitude swapping with another face, while the bottom was obtained by averaging within each patch. The patch size was 4×4 pixels

ensure that amplitude dominates phase. As we pointed out earlier this is true, though trivial, for a single pixel patchsize. However, it does not follow from this that phase makes no contribution, even when the patch size is as small as 4×4. To check on this point we analyzed a single image without phase-amplitude swapping. All of the pixels within a patch were replaced by the mean intensity value of that patch, and a new image was constructed from densely overlapping patches, as in the original procedure. The results for various patch sizes are illustrated in Fig. 6. They show quite clearly that images reconstructed from local DC and the right phase are substantially inferior in reproducing detail to those made by reconstruction from the right local amplitude spectrum and the wrong phase. It follows that the local amplitude spectrum contributes more than the local DC level. We assume that this result is due to the amplitude spectrum providing first and second order statistical information.

The concept of a patchwise analysis of the visual image arose in the first instance from psychophysics and cortical physiology. Psychophysics clearly points to the importance of retinally-localised phenomena, such as local adaptation. At a given cortical location cells have receptive fields centred on approximately the same point of the retina, but with some degree of apparently random positional scatter. Single-unit recording in the cortex shows that at approximately the same location, cells have receptive fields covering a full range of preferred orientations and a wide range of spatial frequencies (Hubel and Wiesel 1974; Hubel and Wiesel 1977). It is, of course, possible that the exact position of each cell is known to a higher level, but there is no compelling reason to suppose so. A plausible model of cortical processing is that cells are organised into a number of spatially overlapping patches, covering the whole of the visual field. Within each patch a number of cells cover the complete range of orientations, and a wide range of spatial frequencies, the centre of the latter range probably depending upon visual eccentricity (that is, upon distance from the centre of gaze). However, there may be no positional label ("local sign") attached to cells within a patch, the only source of positional information coming from the position of the patch itself, and from interpolation across patches where appropriate. On such a model cells would be insensitive to global phase, but not necessarily to local phase. There is evidence for the latter in the grouping of some cells into spatially-coincident odd-symmetric and even-symmetric pairs; an arrangement first described in the cat by Pollen and Ronner (1981). Such a grouping may have the function of detecting peaks in local energy, and thus the presence of phase congruence at features.

Translating this model crudely into Fourier terminology, we can say that the activity of cells within each patch carries the amplitude spectrum of the Fourier transform, but that phase information within a patch is not preserved. Global phase information is carried implicitly only, in the distribution of activity across patches. Our results suggest that this is an efficient method of image encoding, but they also set lower limits to the sizes of images that can effectively be encoded in this way. Fine detail, marked by peaks in local energy, may need to be encoded by more precise local signs.

One way of interpreting our results is to consider the amplitude spectrum as providing crucial information about local texture. Julesz (1962, 1975) proposed that texture is carried mainly by first and second-order statistical information. Although psychophysical investigations have complicated the picture somewhat (Julesz et al. 1973), the overwhelming weight of evidence is in favour of Julesz's original conjecture. If, for the moment, we neglect localised features, an image can be considered as a set of spatially-localised texture patches, each defined by its statistics up to and including the second order. The principal implication of our experiments is that images reconstructed in this way do indeed provide adequate information for the human observer for recognition of objects such as faces, provided there is a ratio in the region of 8 patches/object. It is known (Kolers et al. 1985) that face recognition falls off sharply below a width of 1.2 deg. This would imply a smallest patch size of about 10 min, which is broadly consistent with estimates of the smallest aggregate receptive field (Hubel and Wiesel 1974).

Acknowledgements. This work was supported by a grant from the Image Interpretation Initiative of the Science and Engineering Research Council.

References

Field D (1987) Relations between the statistics of natural images and the response properties of cortical cells. J Opt Soc Am A 4:2379–2394

Hayes A, Morrone MC, Burr DC (1986) Recognition of positive and negative bandpass-filtered images. Perception 15:595–602

Hubel DH, Wiesel TN (1974) Uniformity of monkey striate cortex: a parallel relationship between field size, scatter, and magnification factor. J Comp Neurol 158:295–305

Hubel DH, Wiesel TN (1977) Ferrier Lecture: functional architecture of macaque monkey visual cortex. Proc R Soc London Ser B 198:1–59

Julesz B (1962) Visual pattern discrimination. IRE Trans Inf Theory IT-8:84–92

Julesz B (1975) Experiments on the visual perception of texture: Sci Am 232:34–43

Julesz B (1980) Spatial-frequency channels in one-, two-, and three-dimensional vision: variations on a theme by Bekesy In: Harris CS (ed) Visual coding and adaptability. Lawrence Erlbaum, New Jersey, pp 263–316

Julesz B, Gilbert EN, Shepp LA, Frisch HA (1973) Inability of humans to discriminate between visual textures that agree in second-order statistics—revisited. Perception 2:391–405

Kolers PA, Duchnicky RL, Sundstroem G (1985) Size in the visual processing of faces and words. J Exp Psychol 11:726–751

Landy MS, Cohen Y, Sperling G (1984) HIPS: A Unix-based image processing system. Comput Vision Graph Image Process 25:331–347

Morrone MC, Burr DC (1988) Feature detection in human vision: a phase-dependent energy model. Proc R Soc London Ser B 235:221–245

Oppenheim AV, Lim JS (1981) The importance of phase in signals Proc IEEE 69:529–541

Piotrowski LN, Campbell FW (1982) A demonstration of the visual importance and flexibility of spatial-frequency amplitude and phase. Perception 11:337–346

Pollen DA, Ronner SF (1981) Phase relationships between adjacent simple cells in the visual cortex. Science 212:1409–1411

Robson JG (1983) Frequency domain visual processing. In: Braddick OJ, Sleigh AC (eds) Physical and biological processing of images. Springer, Berlin Heidelberg New York, pp 73–87

Shapley R, Caelli T, Grossberg S, Morgan MJ, Rentschler I (1990) Computational theories of visual perception In: Spillman L, Werner JB (eds) Visual perception: the neurophysiological foundations. Academic Press, New York, pp 417–447

Wegmann B, Zetzsche C (1990) Visual system based polar quantization of local amplitude and local phase of orientation filter outputs. In: Rogowitz BE, Allebach JP (eds) Human vision and electronic imaging: models, methods and applications. Proc SPIE 1249:306–317

Dr. Michael J. Morgan
Department of Pharmacology
University of Edinburgh
Medical School
1 George Square
Edinburgh EH8 9J2
United Kingdom