Masking in Visual Recognition: Effects of Two-Dimensional Filtered Noise

Abstract. *It is difficult to recognize portraits that have been coarsely sampled and quantized. Blurring such images improves recognition. A simple, straightforward explanation is that high-frequency noise introduced by the sampling and quantizing must be removed by low-pass filtering to improve the signal-to-noise ratio and hence signal detectability or recognition. Experiments reported here, suggested on the basis of a different model, show instead that noise bands that are spectrally adjacent to the picture's spectrum are considerably more effective in suppressing recognition.*

Recent investigation of the recognizability of coarsely sampled and quantized portraits showed that faces can be recognized more easily after blurring the sampled and quantized picture (*1*). An example of such a picture, obtained by computer processing of a high-quality photograph, is shown in Fig. 1a. Each square block of uniform density represents the average density of that region in the original photograph. This averaging within a block removes the high-frequency spectral components of the image and thus is a low-pass filtering operation. We shall refer to such processed pictures as block pictures.

In asking why recognition is improved by blurring a block portrait (for example, by squinting, defocusing, or viewing at a distance), one is first led to the following explanation. Whenever a signal, limited to a spectral band of 0 to w, is discretely quantized, "noise" artifacts whose spectrum extends above w are introduced. This quantization noise is conventionally stripped off by low-pass filtering the signal plus noise; when this is done so as to remove all spectral components above w, most of the quantization noise is removed while the desired signal information is retained.

In block-portrait pictures like that of Fig. 1a, the most obvious sampling-noise components seem to be those introduced by the sharp edges of the squares. Although Fourier analysis shows the energy content of those high frequencies to be relatively small, one might speculate that since the eye is particularly sensitive to straight lines and regular geometrical shapes, such square-patterned noise masks especially well (*2*). That is, this image-correlated quantization noise may mask more effectively than randomly distributed noise of equal energy. Hence, moderate low-pass filtering of such presentations should enhance perception. And indeed, as one progressively defocuses the projected images of presentations like that of Fig. 1a, or as the viewing distance from the sharp image is increased, recognition improves rapidly.

This rather obvious explanation, however, is not the only candidate. Another consideration, quite different in kind, comes from what is called critical-band masking. This phenomenon, which is now known to exist for both audition (*3*) and vision (*4, 5*), demonstrates that spectral proximity of signal and noise components drastically influences detection thresholds. In particular, the detectability of a signal of a certain frequency is impaired by the presence of noise in its spectral vicinity.

In audition, the threshold for detecting a single sinusoid anywhere in the spectrum is elevated when a second (noise) sinusoid is introduced, if the noise lies well within an octave of the signal (*3*). But if the noise lies outside this band, masking does not occur.

Correspondingly, in vision there is recent perceptual evidence for spectral decomposition based on spatial-frequency-tuned analyzers in the nervous system. This was first found for one-dimensional single-frequency presentations. Subjects were shown a high-contrast single-frequency sinusoidal grating, and then a determination was made of the raised visibility threshold for a similar but low-contrast "test" grating (*4*). The visibility threshold for this test grating was raised if its spatial frequency fell within ±1.5 octaves of that of the adaptation grating. A subsequent study of the effect, in which signals well above threshold were masked by simultaneously displayed high-contrast one-dimensional

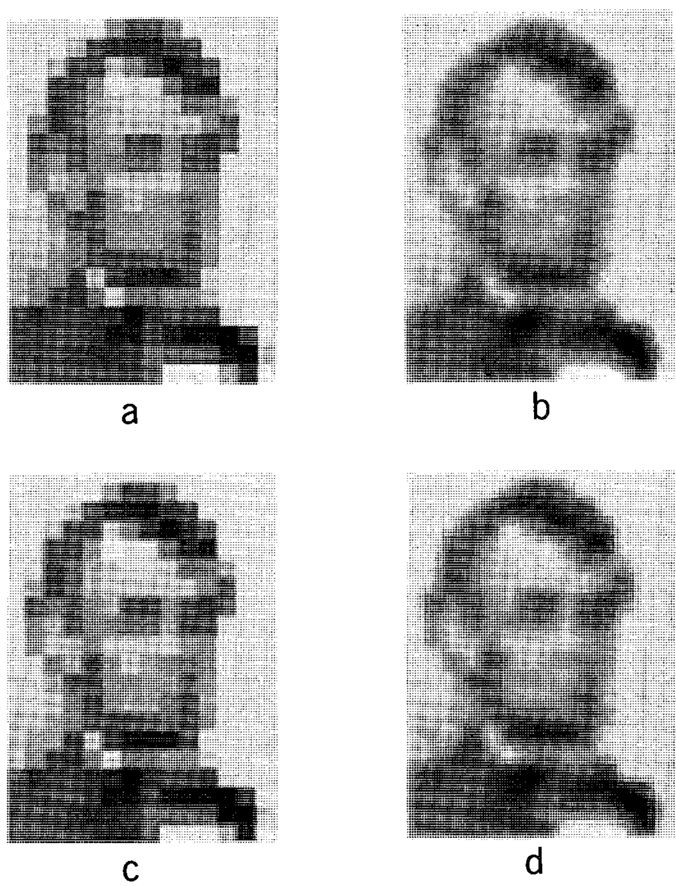


Fig. 1. (a) Computer-processed block portrait. A high-resolution photograph was scanned with a flying spot, digitized, and stored on magnetic tape. A digital computer averaged the brightness values in each of 15 by 20 small square domains. Each of these squares consists of 25 by 25 samples. New, average brightness values were quantized linearly into 16 levels (4 bits), written on digital tape, and converted to analog signals, which were printed on a Facsimile receiver. The print was cropped to 14 by 19 blocks; this portrait is thus represented by 14 × 19 × 4 = 1064 bits. (b) The block portrait of (a) was low-pass filtered just above the highest spatial frequency, w, of the picture information present. Since the intact block portrait was 20 blocks high, $w = 10$ cycles. The picture in (b) has been low-pass filtered to $1.26w = 12.6$ cycles. (c) Block portrait of (a) low-pass filtered at $4.06w = 40.6$ cycles. (d) Block portrait of (a) filtered to eliminate the frequency band from $1.22w$ to $3.94w$; that is, all spatial frequencies from 12.2 to 39.4 cycles, roughly the two octaves just above the picture information, have been removed.

noise, yielded similar results (5). This masking noise consisted of a sum of vertical sinusoidal gratings whose luminance values formed a Gaussian distribution in the horizontal direction. After spectral shaping, this noise, uncorrelated frame by frame at a 60-hertz rate, was superimposed on a single-frequency, vertical sinusoidal grating. This study showed that a sinusoidal grating is masked effectively only if the noise falls within approximately ± 1.5 octaves of the frequency of the grating. If no spectral component of the masking noise comes nearer than about two octaves to the frequency of a grating, the visibility of the grating is not affected by the noise (6).

Suppose we assume that a similar sort of critical-band masking effect may hold for pictures like Fig. 1a. We then would expect that the components of the quantization noise that fall within about two octaves above w would be primarily responsible for perceptual masking. All the rest of the spectral components of the noise, including the high-frequency portions contributing to the sharp edges of the blocks, should have little or no masking influence. Consequently, the explanation for recognition effects in these pictures may be far different from the simple notion about low-pass filtering first espoused (7).

To resolve this question we prepared a series of pictures which were spectrally manipulated in the computer. An input picture was transformed to obtain its Fourier spectrum, filtered to specification, then transformed back and printed out.

The system utilized a Facsimile transmitter/receiver in conjunction with a Honeywell DDP-224 computer to convert photographic images to digital representation, and back again. The input image was sampled into 500 by 500 points, and a black margin 125 points wide was added in both dimensions (to avoid spurious edge effects due to filter ringing). Thus, the subsequent Fourier transforms and the filters worked with a 625 by 625 array.

The digitized information was processed by a Honeywell 6078 computer with a conventional two-dimensional fast-Fourier-transform program. The transformed image was processed by a two-dimensional filter; the result, when transformed back, provided the digitized output. After conversion from digital to analog representation, the data were converted to a final picture by using the Facsimile writer. In a first

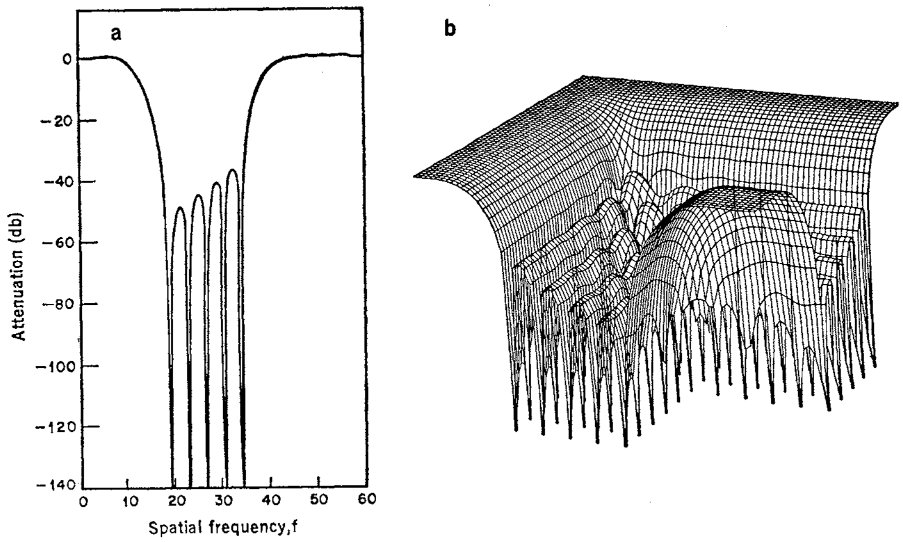


Fig. 2. (a) Attenuation characteristic of the one-dimensional band-rejection filter used to process Fig. 1d. The stop-band extends from 12.2 to 39.4 cycles. The minimum attenuation is 37 db. (b) Plot of the two-dimensional attenuation characteristic for the cascaded band-rejection filter of (a). (A floor of -100 db was inserted in the plot for format convenience.)

test of the system, Fig. 1a was made in this way. The filter was all-pass (that is, no spectral shaping occurred), and the output picture is indistinguishable from the original block-picture input.

In order to experimentally test critical-band masking in images which are more complicated than simple gratings, we proceeded as follows. First, we filtered the picture of Fig. 1a by a low-pass filter designed to attenuate all frequencies just above the upper limit of the picture's spectrum. Since the original picture height was 20 blocks, at most 10 cycles per vertical dimension could be represented. Thus, the upper limit of the picture spectrum, w, was 10 cycles per vertical dimension. The cutoff frequency of the filter was set to $1.26w$, that is, 12.6 cycles. The minimum attenuation in the stop-band of the filter was 68 db. The resultant picture output is shown in Fig. 1b.

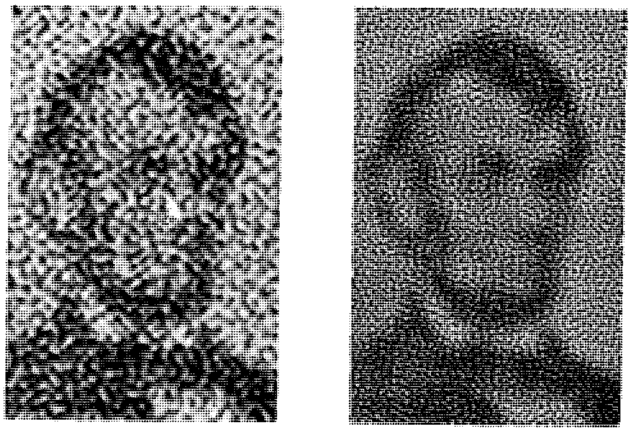
As one would expect, the result looks very much like a blurred version of the original high-resolution photograph. Except for this degradation, recognition of a particular face is apparently unimpaired—there is no evident masking.

In the second experiment, the image of Fig. 1a was again low-pass filtered. This time the cutoff frequency was $4.06w$, or 40.6 cycles per picture height, the attentuation being 67 db. Thus, no quantization-noise components appear beyond two octaves above the picture's spectrum. The result appears in Fig. 1c. Despite the softening of the block edges due to reduced high-frequency content, the portrait's aspect is still very much like that of Fig. 1a, and recognition remains difficult. Considerable visual integration (blurring) is still required for satisfactory perception.

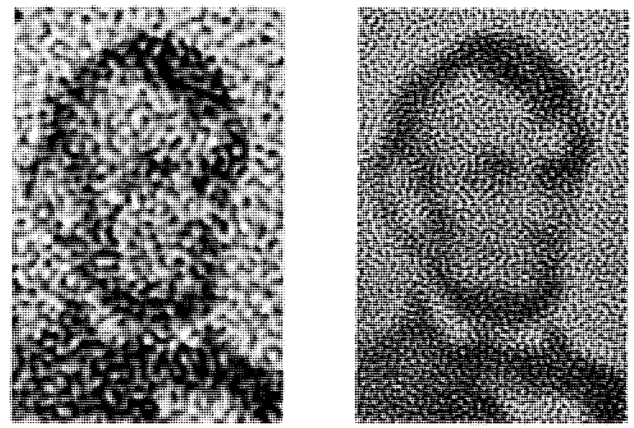
In the third experiment, the image of Fig. 1a was filtered by a band-rejection filter whose stop-band extended from $1.22w$ to $3.94w$, with a 37-db minimum attenuation. The result is shown in Fig. 1d. Although the quantization noise (especially at the block boundaries) is apparent, it does not seem to affect recognition of the portrait very much; it is almost as though the sharp lines were superimposed on the picture rather than being part of it (unlike the aspect presented by Fig. 1c). This rendition is much closer to Fig. 1b than is Fig. 1c.

One- and two-dimensional plots of the filter characteristic used to process Fig. 1d are shown in Fig. 2, a and b. This band-rejection filter suppresses all frequencies from 12.2 to 39.4 cycles, with 37-db minimum attenuation.

In this pilot exploration, cursory inspection reveals a great similarity between Fig. 1a and Fig. 1c, and between Fig. 1b and Fig. 1d. This suggests that the quantization noise spectrally adjacent to the picture is most effective in suppressing recognition. At least, under the conditions employed, the very high frequencies do not have the strongest influence, contrary to what would be predicted by the progressive defocusing effects. Thus, the hypothesis for critical-band masking in two-dimensional scenes, which initiated



a b



a b

Fig. 3 (left). (a) Low-pass filtered image with band-limited noise added. The spatial frequencies of the portrait were suppressed above $w = 10$ cycles while noise of uniform amplitude was added between w and $4w$, that is, just adjacent to the portrait's spectrum. (b) Same as (a) except that the added noise was in the spectral range $4w$ to $7w$; that is, the noise was of the same amplitude and spectral extent as in (a), but this time no closer than two octaves to the portrait's spectrum. The filters in both (a) and (b) were sequentially applied Cartesian filters; the filtering was done first in the x direction, then in the y direction. Fig. 4 (right). Same as Fig. 3, a and b, except that the filters were circularly symmetric rather than rectilinear.

these experiments, might seem to be supported.

It is impossible, however, to establish a firm case for critical-band masking from this result, for three reasons: (i) the block sampling introduces picture-correlated spectral components that confound any results (noise amplitudes depend on picture densities); (ii) these components are spatially periodic as well, introducing a peculiar quantization noise (the block process introduces spectral components at the block frequency and its harmonics); and (iii) the noise energy introduced by the block-sampling process decreases with increasing frequency; hence the adjacent-band noise should be more effective than the remote-band noise in masking.

The way around this difficulty is simple. We use portraits that are not quantized but simply low-pass filtered. And instead of having noise that is periodic, picture correlated, and nonuniform in energy, we introduce uncorrelated, random noise that has constant average energy over the spectrum. This was done for our final experiment, as follows.

The original high-resolution portrait was first low-pass filtered to 10 cycles per picture height, yielding a smooth, blurred image which is virtually identical to that of Fig. 1b. Noise was obtained from a random-number generator which supplied one value of a uniformly distributed random variable for each picture point (500 by 500 points, as before). The sampling rate of the statistically independent noise was high compared to the block frequency; hence, the spectral shaping that was done represents the central portion of the noise distribution and is therefore essentially flat. After the noise was filtered to occupy a specified region of the spatial spectrum, it was added to the picture. The average noise amplitude was empirically chosen to provide considerable masking of the signal.

Figure 3 shows typical resultant pictures; the noise spectrum in Fig. 3a covered the range w to $4w$ (that is, 10 to 40 cycles) and thus was immediately adjacent to the portrait's signal band (0 to 10 cycles). Figure 3b also contains 30 cycles of noise, but the spectral position of that noise, $4w$ to $7w$, was no closer than two octaves to the portrait's spectrum. For both conditions the masking-noise band was $3w$ wide; thus, noise of the same bandwidth was provided in each case.

The result is unequivocal. The most stringent test is obtained by viewing from a distance of no more than 18 times the picture height (8). But even casual inspection shows the portrait of Fig. 3b to be considerably easier to recognize than that of Fig. 3a. Critical-band masking thus appears to hold for complex pictures, confirming our initial conjecture.

The use of rectangular filters introduces some seemingly regular granularity in the pictures of Fig. 3. These filters have square-shaped passbands in the two-dimensional Fourier domain, as can be seen in Fig. 2b. Consequently, the diagonal frequency components of the noise are 1.4 times as long as the horizontal and vertical components. The resultant granularity is avoided in the pictures of Fig. 4, where the spectral manipulation was done with a filter having a circular-shaped passband; the fundamental result, greater masking by spectrally adjacent noise, is not noticeably influenced.

It remains to be seen how various noises of different spectral shapes but equal spectral energy and bandwidth may mask. However, with block noise (see Fig. 1, c and d) and with uniform noise (see Fig. 3, a and b) spectral proximity masks more effectively. This suggests that the location of the noise band is more important than its shape.

LEON D. HARMON*
BELA JULESZ

Bell Laboratories,
Murray Hill, New Jersey 07974

References and Notes

1. L. D. Harmon, in *Pattern Recognition in Biological and Technical Systems*, O.-J. Grusser and R. Klinke, Eds. (Springer-Verlag, New York, 1971), pp. 196–219.
2. L. D. Harmon, A. B. Lesk, J. Z. Levinson, in preparation.
3. E. Zwicker, G. Flottorp, S. S. Stevens, *J. Acoust. Soc. Amer.* **29**, 548 (1957).
4. These adaptation experiments were reported by A. Pantle and R. Sekuler [*Science* **162**, **1146** (1968)] and C. Blakemore and F. W. Campbell [*J. Physiol. London* **203**, 237 (1969)]. The existence of spatial-frequency-tuned channels was first proposed by F. W. Campbell and J. G. Robson [*ibid.* **197**, 551 (1968)] in an investigation of the visibility thresholds of sinusoidal or square-wave gratings. Adaptation to high-contrast gratings also produces an apparent shift of neighboring frequencies of test gratings, as shown by C. Blakemore and P. Sutton [*Science* **166**, 245 (1969)]. For discussions of physiological evidence for frequency-tuned channels, see: C. Enroth-Cugell and J. G. Robson, *J. Physiol. London* **187**, 517 (1966); F. W. Campbell, G. F. Cooper,

C. Enroth-Cugell, *ibid.* **203**, 223 (1969); F. W. Campbell, G. F. Cooper, J. G. Robson, M. B. Sachs, *ibid.* **204**, 130 (1969).
5. B. Julesz, in *Proceedings of the Seventh International Congress on Acoustics, Budapest, 1971* (Akadémiai Kiadó, Budapest, 1971), pp. 446–447; C. F. Stromeyer and B. Julesz, *J. Opt. Soc. Amer.* **62**, 1221 (1972).
6. Besides single-sinusoid gratings, square waves (bar gratings) were also used in these experiments. The influence of spectrally proximate noise was similar.
7. It must be noted that this expectation will be tempered by two provisions. First, previous results on visual critical-band masking are based on one-dimensional presentations; the experiments reported here deal with two-dimensional displays. Second, in the prior work only single sinusoids or square waves were used; the present material is pictorial. So, provided these differences are unimportant, we may be able to establish a case for critical-band masking in the present experiments.
8. Although the noise amplitudes and bandwidths were the same for the two cases, the noise energies were not thereby made equal. This can be seen by considering that in the two-dimensional Fourier space, the ratio of the areas of the annuli described by the noise bands 4*w* to 7*w* and *w* to 4*w* are unequal in the ratio 2.2 : 1. Furthermore, the human contrast-sensitivity curve (modulation transfer function) peaks around 7 cycles per degree, and thus the eye's noise-sensitivity is different for the two cases. To obtain equal masking energies in Fig. 3a and Fig. 3b, it is necessary for the average observer to view Fig. 3 from a distance of about 18 times the picture height. For closer viewing distances, the perceived noise in Fig. 3b becomes relatively greater, and the test is even more stringent.
9. We acknowledge the assistance of A. B. Lesk for the computer programming and picture processing. We thank L. R. Rabiner and R. W. Schafer for their consultation on digital filtering techniques, and we are grateful to them and to other colleagues for critical discussions and manuscript review.
* Present address: Department of Biomedical Engineering, Case Western Reserve University, Cleveland, Ohio 44106.
6 November 1972; revised 30 January 1973

Temperature-Sensitive Pawns: Conditional Behavioral Mutants of Paramecium aurelia

Abstract. *"Pawns" are mutants of* Paramecium aurelia *in which the process of calcium activation during membrane excitation is genetically impaired, with a corresponding loss of avoiding reactions. Mutants are selected that behave normally when grown at 23°C but as pawns at 35°C. The normal excitation can now be disrupted and restored in the same strain at will.*

We have succeeded in selecting temperature-dependent mutants by modifying a previous method based on the principle of chemotactic interference with geotaxis (*1*). These mutants behave normally at room temperature and are virtually indistinguishable from the wild type. When grown at restrictive temperatures, they exhibit behavioral aberrations which can be grouped with the phenotypes of the previously analyzed, temperature-independent behavioral mutants. Of major interest are the temperature-sensitive "pawns" that lose their avoiding reactions completely when cultured at 35°C.

Ciliated protozoa perform the avoiding reaction in response to various stimuli (*2*). This reaction involves a period of backward swimming which results from reversing the beating direction of cilia. Such reversal is correlated with membrane depolarization (*3*). Eckert proposed that the influx of Ca^{2+}, as the result of such depolarization in normal membranes capable of Ca activation, causes the ciliary reversal (*4*). Kung reported the discovery of the behavioral mutant pawn, which was completely lacking the avoiding reaction (*1*). The genetic defect was traced to the specific loss of proper Ca activation during excitation of the membrane (*5*). We then induced and searched for heat-sensitive mutants in order to understand the nature of various mutational effects which may lead to the pawn phenotype and as an expansion of our program to dissect genetically the excitable membrane of *Paramecium*.

We used the standard culture technique (*6*) and mutagen treatment (*1*). The mutagenized exautogamous cells were cultured for a phenomic lag period of four to eight fissions, and the descendant populations were used to select for behavioral mutants. In previous studies, pawns had been selected by first injecting a mutagenized population into the bottom of a screening column filled with solution of high Na^+ concentration and later taking a selected fraction from the top of this column. In this solution, the normal animals in the population exhibited repeated avoiding reactions, moving randomly about the position in which the population was placed. Mutants unable to avoid Na^+ retained the natural tendency to swim upward, completing the negative geotactic migration in a few minutes. Thus the selected fraction collected from the top of the column after migration time was greatly enriched with mutants unable to avoid Na^+, such as the pawns (*1*). Two modifications were made in the present study.

1) Mutagenized exautogamous populations of about 10^4 cells were condensed by centrifugation (250*g*) into 10 ml of culture medium. Each population was gradually adapted into a sucrose solution of a final concentration of 65.2 m*M* through three steps spaced 20 minutes apart. Each step involved slowly dripping 0.5 ml of 500 m*M* sucrose into the culture medium. The populations that had been adapted to sucrose were then gently injected through a lower spout into a column filled with a solution rich in Na (*7*). After the migration time, a top fraction of 5 ml was removed through an upper spout of the column. All cells in this fraction were subsequently cloned separately. Sucrose was used to ensure that the injected fraction was evenly layered at the bottom of the column. This method successfully prevented an undesirable flaring, often encountered in previous studies in which the injected material and the liquid in the column had a similar density (*1*, *8*). In our method, the injected population formed a dense and even layer just below the boundary of the two liquids with the individuals in the population performing repeated avoiding reactions.

2) To screen for heat-sensitive mutants, the above procedure was followed at a restrictive temperature. From the point of exautogamous expansion during the phenomic lag until cloning of the cells from the selected fractions, all steps were performed in a 37°C walk-in incubator. Some clones obtained from such screening were found to be pawns at all temperatures and were phenotypically identical to the previ-

Table 1. Reactions of various strains of *P. aurelia* to different cationic stimuli at two temperatures; Na^+ means that the test medium contains 20 m*M* NaCl and 0.3 m*M* $CaCl_2$; K^+ means 8 m*M* KCl and 0.1 m*M* $CaCl_2$; and Ba^{2+} means 8 m*M* $BaCl_2$ and 1 m*M* $CaCl_2$. All solutions contain 1 m*M* tris buffered at *p*H 7.2. Figure 1 records the reaction of the Ba^{2+} solution. +, indicates the presence of obvious avoiding reactions; −, indicates the complete lack of avoiding reaction.

Strains	Reactions of cells grown at: 23°C			35°C		
	Na^+	K^+	Ba^{2+}	Na^+	K^+	Ba^{2+}
51s (wild type)	+	+	+	+	+	+
d4-133 (*ts* pawn)	+	+	+	−	−	−
d4-95 (pawn)	−	−	−	−	−	−