PERGAMON

Vision Research 38 (1998) 3025–3029

Vision Research

Rapid Communication

How much of the visual object is used in estimating its position?

Leonid L. Kontsevich *, Christopher W. Tyler

Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, CA 94115, USA

Received 12 November 1997; received in revised form 4 May 1998

Abstract

Localization accuracy for a wide Gaussian bar (sigma 1°) was essentially invariant with sampling density over a wide range. Luminance contrast, on the contrary, had a profound effect on localization accuracy. This difference suggests that the position is estimated from a few (3–4) samples from all those available in the image. A subsequent experiment confirmed that limiting the display to four samples did not impair localization accuracy. Computer simulations show that the result cannot be explained by the peak or centroid rule for position. The results imply that the visual system uses only a few samples to interpolate the luminance profiles, regardless of how many samples are available in the image.

Keywords: Position coding; Localization accuracy; Visual object

1. Introduction

Visual perception allows the localization of visual stimuli with remarkably high precision even when they are large entities with poorly defined features and edges [1]. A major problem in understanding the neural mechanisms of visual localization is its accuracy in localizing the peaks of stimuli that are represented by isolated elements (samples), so that the peak position is typically not present in the stimulus (Fig. 1b). A widely accepted theory for this performance says that the visual system integrates the samples into a fuzzy blob by means of spatial filtering [2,3]. The blob position is then estimated as the centroid [4] or peak [5] of the blob profile. Morgan and Watt [4] demonstrated that the accuracy of position estimation drops precipitously beyond about 3 arc min separation between the samples representing the luminance profile, implying that this distance is the extent of the spatial filter. Such an interpretation, therefore, attributes localization to low-level visual processing. To test this proposition, we measured the effect of sparse sampling over a wide range, well beyond that of local optical blurring in the retina. We also used a simple Gaussian target to minimize any aliasing effects of the sampling. The results invalidate low-level filtering theory, suggesting that localization of both continuous and sampled objects is based on interpolation of the samples selected by a high-level process.

* Corresponding author. Tel.: +1 415 5611793; fax: +1 415 5611610; e-mail: lenny@skivs.ski.org.

2. Methods

The stimuli were presented in a dark room at 1.13 m distance from observer on the screen of a calibrated monochrome monitor. The observers were instructed to judge the peak position of the test Gaussian profile relative the reference (leftward or rightward shift) while fixating on a cross in the middle.

The stimuli consisted of two ribbons of light (24 arc min × 8°) separated vertically by 6 arc min. The top ribbon contained a position reference, which was a dark 1 arc min line in the center. The bottom ribbon had a one-dimensional Gaussian modulation of luminance, which could be continuous (Fig. 1a) or sampled (Fig. 1b). For sampled stimulus the spatial phase of the samples was randomized across trials. The experimental variable was displacement of the peak of the modulation relative to the center.

0042-6989/98/$19.00
PII: S0042-6989(98)00148-5

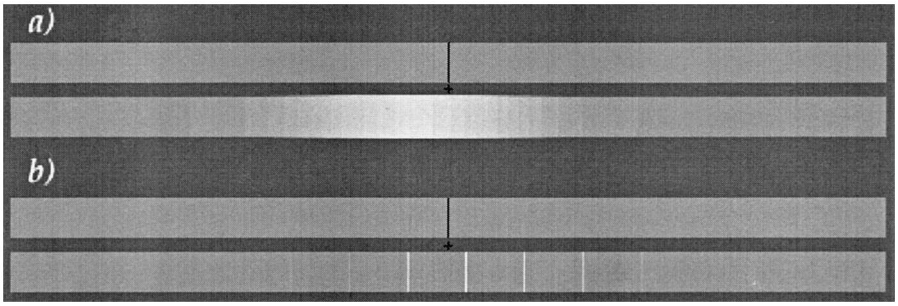

Fig. 1. Examples of (a) continuous and (b) sampled stimuli, which were Gaussian profiles with 1° sigma. The contrast in both cases was set at the level of 0.5. The sampling steps in the continuous and sampled stimuli are 1 and 32 arc min, respectively. The task for the observer was to tell in which direction the peak of the profile displaced relative to the dark reference bar.

The background luminance was 17 cd/m^2; the base luminance of the ribbons L_0 was 34 cd/m^2. The Gaussian profile $G = L_0(1 + Ce^{(x-\Delta x)^2/\sigma^2})$ always had $\sigma = 60$ arc min and was always brighter than the base luminance. The parameter C is the Weber contrast for continuous profiles; for sampled profiles it may deviate from actual Weber contrast of the stimulus because the peak may be not sampled. The luminance of the reference line was set at 0.2 cd/m^2 to diminish involvement of luminance cues in determining the position.

Placement of the Gaussian profile relative to the center (parameter Δx) was controlled by a new adaptive PSI procedure (Kontsevich and Tyler, submitted), which maximizes information gain on each trial based on the Bayesian rule. The procedure estimated both threshold, which reflects width of the transitional range of the psychometric function, and point of subjective equality, which determines the position of the transitional range. The threshold was defined as one σ of the cumulative Gaussian parameterization of the psychometric function. This threshold corresponds to the distance between the point of subjective equality and the point where the shift in rightward direction was reported in 81% of trials. The miss rate was set at 5%. The bias and threshold were estimated after completion of 70 trials. Each point shown is a result of averaging of at least four threshold estimates. In each trial both profile and reference were presented simultaneously for 0.5 s; between the trials both ribbons were blank. Auditory feedback was provided only during short training sessions prior to the experiment.

One of the authors (LK) and a naive observer (NF) participated in the experiments.

3. Experiment 1

In this experiment the effect of contrast on position error was studied for continuous (one-pixel sampling step) and sampled (32-pixel step) stimuli. For observer NF (Fig. 2a), contrast increase reduced localization error across the studied range of contrasts (from 0.0625 to 1.0). Observer LK (Fig. 2b) shows similar results except at the unity contrast point for the continuous profile. The slopes for continuous profiles in double-logarithmic coordinates were slightly shallower than -0.5 for both observers. This result is consistent with numerous studies showing that localization error decreases reciprocally with the square root of contrast at low contrasts and becomes contrast-independent at higher contrasts [6,7]. This transition contrast between the contrast-dependent and contrast-independent regimes is between 0.5 and 1.0 for LK and somewhere above 1.0 for NF.

The localization errors for phase-randomized sampled profiles show stronger contrast dependence than for continuous profiles: the slope for both observers was -0.76. This relationship was evident through the contrast range.

4. Experiment 2

This experiment evaluated the effect of phase-randomized sampling density on localization error. The contrast of the profile was set to the value of 0.5, which falls in the range where localization accuracy depends on contrast for both continuous and sampled profiles. For each observer there are conditions where the localization threshold is lower than the value obtained at that contrast (see Fig. 2). The localization thresholds to be studied, therefore, are not a result of some intrinsic factor limiting the localization accuracy.

The results shown in Fig. 3 indicate little effect of sampling density on localization accuracy up to extremely high separations. The accuracy dropped abruptly (threshold rose) at sampling step of 64 for the observer NF and at 96 for LK. At these separations only 1–4 samples were visible. For sampling steps starting from 1 arc min (where the profile was represented by every pixel) to 32 arc min (where only one of every 32 pixels carried the luminance information), sampling density had little effect on localization accuracy.

These results are incompatible with an ideal observer model of information integration. If all information available from the stimulus were utilized, then according to signal detection theory the localization error would be reciprocal to the square root of sampling density [8] regardless of the origin of noise in the samples. This predicted rate of increase is shown in Fig. 3 by a dashed line. The scale of the assumed noise was adjusted to match the datum point at the smallest sampling step. The data fall close to this line for the sampling steps from 1 to 2 arc min but then show no further increase until very wide sampling steps.

The result is also incompatible with explanations attributing localization to the signals from spatial filters

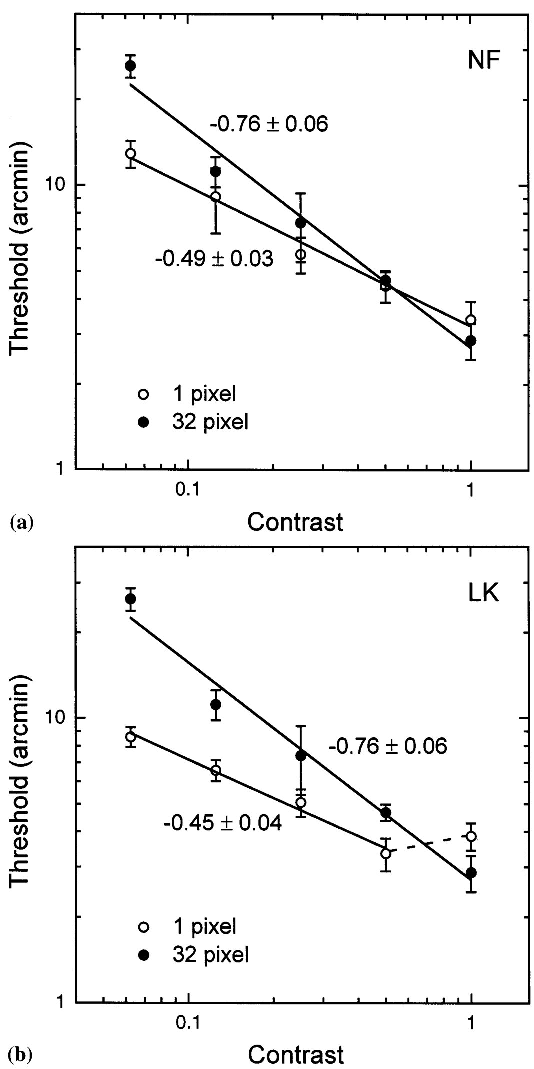

Fig. 2. Variation of localization threshold for two observers as a function of contrast for continuous (1 arc min step) and sampled (32 arc min step) stimuli. Linear regression lines and their slopes are shown for each data set. Error bars depict ± one standard error of the mean (S.E.M.).

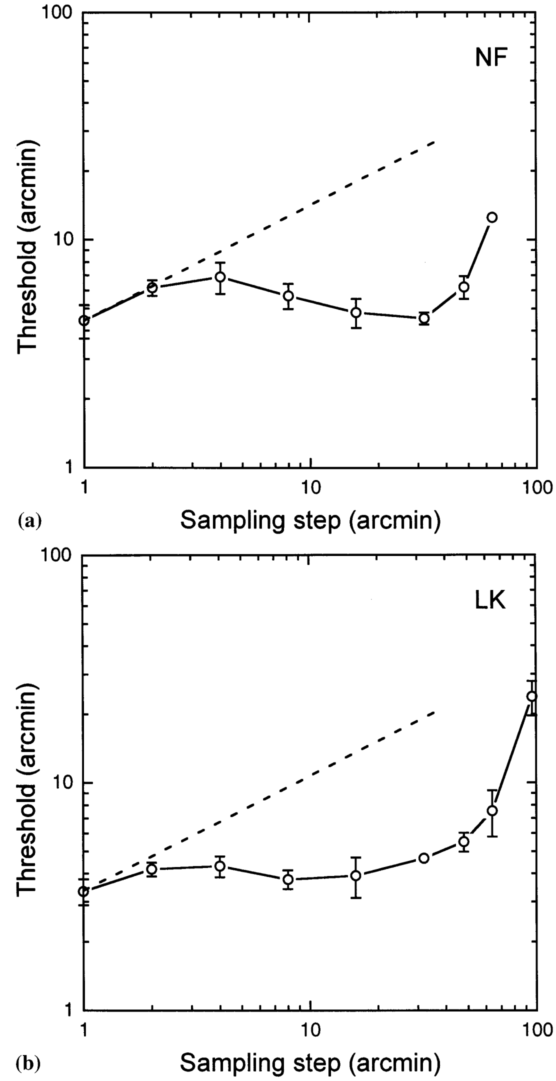

Fig. 3. Localization threshold as a function of sampling step with a profile contrast of 0.5. The dashed lines show the prediction of the ideal observer model with the sampling noise parameter adjusted to fit the data for the datum point at 1 arc min sampling step. Error bars depict ±1 S.E.M.

integrating information from the samples with their receptive fields. The receptive field size of such filters would need to be at least one degree in diameter to accommodate three samples separated by 32 arc min. For smaller sampling steps effective contrast in the linear filters is proportional to the number of samples present in the receptive field; i.e. sampling density should affect the localization accuracy exactly as contrast does. Localization threshold as a function of the sampling step, therefore, should have a slope about 0.5–0.75 in double-logarithmic coordinates to match the data from the previous experiment. This is certainly not the case in Fig. 3.

The small rise of thresholds at the smallest sampling steps in Fig. 3 can be attributed to the Morgan/Watt reduction of the effective contrast in the front-end filters of the localization mechanism. The detrimental effect of the sampling step on localization accuracy disappears between 2 and 4 arc min steps where our observers start to resolve the samples, which confirms the finding of Morgan and Watt [4] that sample integration disappears at a 150–200 arc sec sampling step for DoG stimuli. Levi et al. [9] reported a somewhat smaller critical sampling step of 90 arc sec for sinusoidal stimuli. Thus, if there is some integration of the samples by spatial filters, it is limited to the smallest sampling steps.

5. Experiment 3

According to the analysis presented for the previous experiment, increase of the sampling step should increase the localization threshold. Conversely, we argue that lack of threshold variation across a wide range of sampling densities implies that the visual system takes into account only a limited number of samples while estimating position. This hypothesis was evaluated in the following experiment.

The localization error was measured for the stimuli represented by the few most central samples with a 32 arc min step. The contrast of the test profile was set at 0.5. The results shown in Fig. 4 demonstrate that the number of samples had little effect on localization accuracy unless the number was reduced to three for observer NF and to two for LK where it impairs their performance. These data suggest that observer NF made position judgments for the profile tested on the basis of exactly four samples regardless of how many samples were present. Observer LK needed only three samples.

This result is consistent with the small number of samples estimated from the previous experiment. The sampling step at which the localization accuracy starts deteriorating is smaller for NF and longer for LK. Given the fixed size of the test stimulus, this difference means that LK needed fewer samples than NF.

6. Computer simulations

How does the visual system estimate the position of the peak from such a few samples? To answer this question we ran simulations of the previous experiment for observers with ideal centroid- and peak-based localization rules.

In the simulations we neglected noise in the sample readings; the only source of localization error was the random phase of the samples. Since adding sample noise would raise the threshold, the simulations provide a lower limit for the threshold estimates.

The thresholds were evaluated for number of samples ranging from 2 to 9. Percent correct for a given profile displacement was computed as a proportion of correct responses across all possible (32) sample phases. The displacement increased until the percent correct exceeded the value of 81%; this displacement was taken as a threshold estimate.

The simulation results are shown in Fig. 4 by the dashed lines. Threshold estimates for the peak rule remains at 10.1. Insensitivity of the threshold to the number of samples has a simple explanation: since the threshold displacement was smaller than the sampling step, the peak was always present in one of the two samples nearest to the center and the other samples were irrelevant to the task.

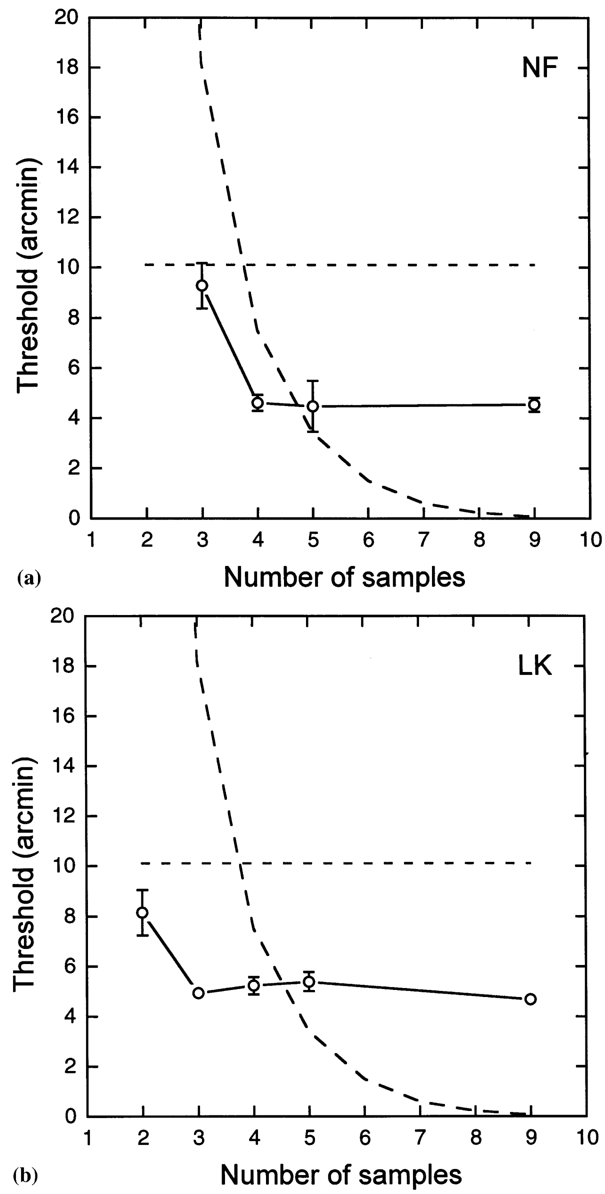

Fig. 4. Localization threshold as a function of the number of samples measured for the 32-pixel sampling step and a contrast of 0.5. The horizontal and hyperbolic dashed lines are the predictions of ideal peak-based and centroid-based localization rules. Error bars depict ±1 S.E.M.

The centroid-based rule, unlike the peak rule, integrates over samples and therefore benefits from the increase of the number of samples. Its threshold estimates produce a hyperbolic curve whose horizontal asymptote is a small positive value.

Since the simulations provide the lower bound on threshold estimates, these estimates ought to be smaller than the experimental ones to be consistent with the data. None of the simulated rules passes this test. Consider the experimental data and model predictions for four samples. The threshold was 4.61 ± 0.32 S.E.M. for observer NF and 5.22 ± 0.35 S.E.M. for LK. These values are significantly ($P < 0.01$) smaller than the lowest estimates provided by the centroid-based (7.5) and peak (9.1) localization rules. The difference is even more dramatic for observer LK when the number of samples is equal to three: 4.92 ± 0.24 S.E.M. for the measured threshold versus 10.1 for the peak-based model and 18.2 for the centroid-based one. We conclude, therefore, that when the number of samples is small none of the tested localization models can match the accuracy of our observers.

Failure of the standard localization models to explain the thresholds for the profile represented by a few samples requires a new model. Our data show that localization process selects a few (3–4 in our case) samples to localize the stimulus. This selectivity suggests that localization is a feature of the late stages of visual processing. After the samples are chosen, they need to be processed into a position code. Our simulations suggest that simple rules for position estimation are too imprecise. It seems that the only way to achieve accuracy comparable with that demonstrated in the experiment would be to interpolate the profile between the samples. We see two ways in which this interpolation can be implemented in the visual system: (1) with a specialized interpolation mechanism or (2) with a match of the sampled profiles at the input with complete representations of these profiles stored in visual memory and projecting the peak position from memory back into the visual field. The data available are insufficient to make a choice between these two options.

7. Related work

The paradigm employed in our study is similar to that used by Morgan and Watt [4]. Despite this similarity, Morgan and Watt's results are quite disparate from ours. They show that for DoG profiles of a range of sizes thresholds have approximately constant localization threshold up 2–3 arc min, while at longer sampling steps the thresholds rapidly deteriorate. Our data, to the contrary, show that no deterioration occurs up to 48–64 arc min. Why the discrepancy?

This discrepancy may be a consequence of slight difference in the stimulus arrangement. The threshold levels imply that the two studies were operating in different ranges. The thresholds in the Morgan and Watt study were about 0.5 arc min for the largest stimulus, whose size is comparable with ours. This value is an order of magnitude lower than the 5 arc min thresholds typical in our study. The test and reference stimuli in Morgan and Watt [4] both had identical luminance profiles and displacement could be determined based on low-level comparison of the luminances. In our experiments the reference stimulus had a different profile and luminance in comparison with the test, which precluded the involvement of low-level luminance-based mechanisms and forced the visual system to rely on some high-level encoding mechanism to make the comparison.

This low-level explanation of Morgan and Watt's results gets additional support from their auxiliary experiment where they measured stereoscopic acuity for the sampled stimuli and found a similar limit of 2–3 arc min. This limit, as mentioned in the discussion of the Section 4, corresponds to the point where the effective contrast in local filters stops affecting the localization accuracy. Thus, it seems likely that a dual-process theory is going to be needed for positional localization of sampled gratings for local tasks, one reflecting a filter-like integration, and the other a more cognitive interpolation process for more abstract tasks.

Acknowledgements

This work was supported by NIH Grant 7890.

References

[1] Toet A, Koenderink JJ. Differential spatial displacement discrimination thresholds for Gabor patches. Vis Res 1988;28:133–43.
[2] Barlow HB. Reconstructing the visual image in time and space. Nature 1979;279:189–90.
[3] Watt RJ, Morgan MJ. A theory of the primitive spatial code in human vision. Vis Res 1985;25:1661–74.
[4] Morgan MJ, Watt RJ. Mechanisms of interpolation in human spatial vision. Nature 1982;299:553–5.
[5] Hess RF, Holliday IE. The coding of spatial position by the human visual system: effects of spatial scale and contrast. Vis Res 1992;32:1085–92.
[6] Levi DM, Klein SA. 'Weber's Law' for position: the role of spatial frequency and contrast. Vis Res 1992;32:2235–50.
[7] Whitaker D, Latham K. Disentangling the role of spatial scale, separation and eccentricity in Weber's law for position. Vis Res 1997;37:515–24.
[8] Watt RJ, Morgan MJ. Spatial filters and the localization of luminance changes in human vision. Vis Res 1984;24:1387–97.
[9] Levi DM, Klein SA. Vernier acuity with simusoidal gratings: what information is needed for precise localization? Invest Ophthalmol Vis Sci 1991;32:1269.