MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY
and
CENTER FOR BIOLOGICAL INFORMATION PROCESSING
WHITAKER COLLEGE

A.I. Memo No. 1271
C.B.I.P. Memo No. 62

January 1991

Synthesis of visual modules from examples: learning hyperacuity

Tomaso Poggio Manfred Fahle Shimon Edelman

Abstract

Networks that solve specific visual tasks, such as the evaluation of spatial relations with hyperacuity precision, can be easily synthesized from a small set of examples. This may have significant implications for the interpretation of many psychophysical results in terms of neuronal models.

This report describes research done at the Massachusetts Institute of Technology within the Artificial Intelligence Laboratory and the Center for Biological Information Processing in the Department of Brain and Cognitive Sciences and Whitaker College. The Center's research is sponsored by grant N00014-88-K-0164 from the Office of Naval Research (ONR), Cognitive and Neural Sciences Division; by the Alfred P. Sloan Foundation; and by National Science Foundation grant IRI-8719392. The Artificial Intelligence Laboratory's research is sponsored by the Advanced Research Projects Agency of the Department of Defense under Army contract DACA76-85-C-0010 and in part by ONR contract N00014-85-K-0124. MF is at Universität Augenklinik, Schleichstr. 12, D7400, Tübingen, Germany. SE is at the Department of Applied Mathematics and Computer Science, The Weizmann Institute of Science, Rehovot 76100, Israel.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704 0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	January 1991	memorandum

4. TITLE AND SUBTITLE

Synthesis of Visual Modules from Examples: Learning Hyperacuity

5. FUNDING NUMBERS

N00014-88-K-0164
IRI-8719392
DACA76-85-C-0010
N00014-85-K-0124

6. AUTHOR(S)

Manfred Fahle

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Artificial Intelligence Laboratory
545 Technology Square
Cambridge, Massachusetts 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

AIM 1271
(C.B.I.P. 62)

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
Information Systems
Arlington, Virginia 22217

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

None

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution of this document is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Networks that solve specific visual tasks, such as the evaluation of spatial relations with hyperacuity precision, can be easily synthesized from a small set of examples. This may have significant implications for the interpretation of many psychophysical results in terms of neuronal models.

14. SUBJECT TERMS (key words)

learning
hyperacuity

15. NUMBER OF PAGES
18

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED

1 A general framework for psychophysical modeling

We wish to propose a single, new hypothesis instead of the many specific models that are invoked to explain a broad range of visual abilities as measured in psychophysical tests. We will consider, in particular, hyperacuity tasks as an example for our claims.

For any given visual competence, it is tempting to conjecture a specific algorithm and a corresponding neural circuitry. It has been often implicitly assumed that this machinery may be hardwired in the brain. This extreme point of view, if taken seriously, may quickly lead to absurd consequences. Consider for instance the many different hyperacuity tasks, some of which are outlined in Figure 1. The underlying reason for the spectacular performance of human subjects in these tasks is that the information sampled by the photoreceptors and relayed to the brain does contain the information necessary for precise localization of image features, since the spacing between photoreceptors and the eye's optics satisfy (in the fovea) the constraints of the sampling theorem [5]. More specifically, it has been shown that, in principle, spatial mechanisms that account for grating resolution are sensitive enough to support hyperacuity-level performance [13,4,26]. Furthermore, some of the hyperacuity tasks can be solved by detecting "secondary" cues such as luminance difference (as in the bisection task) or orientation (as in the detection of vertical vernier stimuli). The detailed structure of the neural circuitry that subserves the detection of these cues, or hyperacuity performance in other tasks is, however, unknown.

Notice that the idea of a fine-grid reconstruction of the image in some layer of the cortex [1,5] is unsatisfactory, because it still requires a homunculus looking at the reconstructed image and applying *a different routine for each specific hyperacuity task*. We propose instead [16] that the brain may be able to synthesize - possibly in the cortex - appropriate modules for specific tasks after a quick training phase in which it is exposed to examples of the task. In most psychophysical experiments, subjects are actually shown several examples of the task before testing takes place. Hyperacuity tests, in particular, require a significant training period in order to achieve good performance (thresholds typically decrease by a factor of two to four during the first several hundreds of stimulus presentations [24]; on the other hand, some subjects have thresholds of 10″ or less upon the first testing). A broad prediction of our conjecture is that almost any psychophysical task could be performed after suitable training, provided the necessary information is available in the stimulus.

Synthesizing a module from examples for a specific task may be often regarded as approximating a multivariate function from sparse data. An efficient scheme for the approximation of smooth functions was proposed recently under the name of HyperBF networks [19]. Detailed descriptions of it, its theoretical underpinnings and its performance can be found in [19], [16],

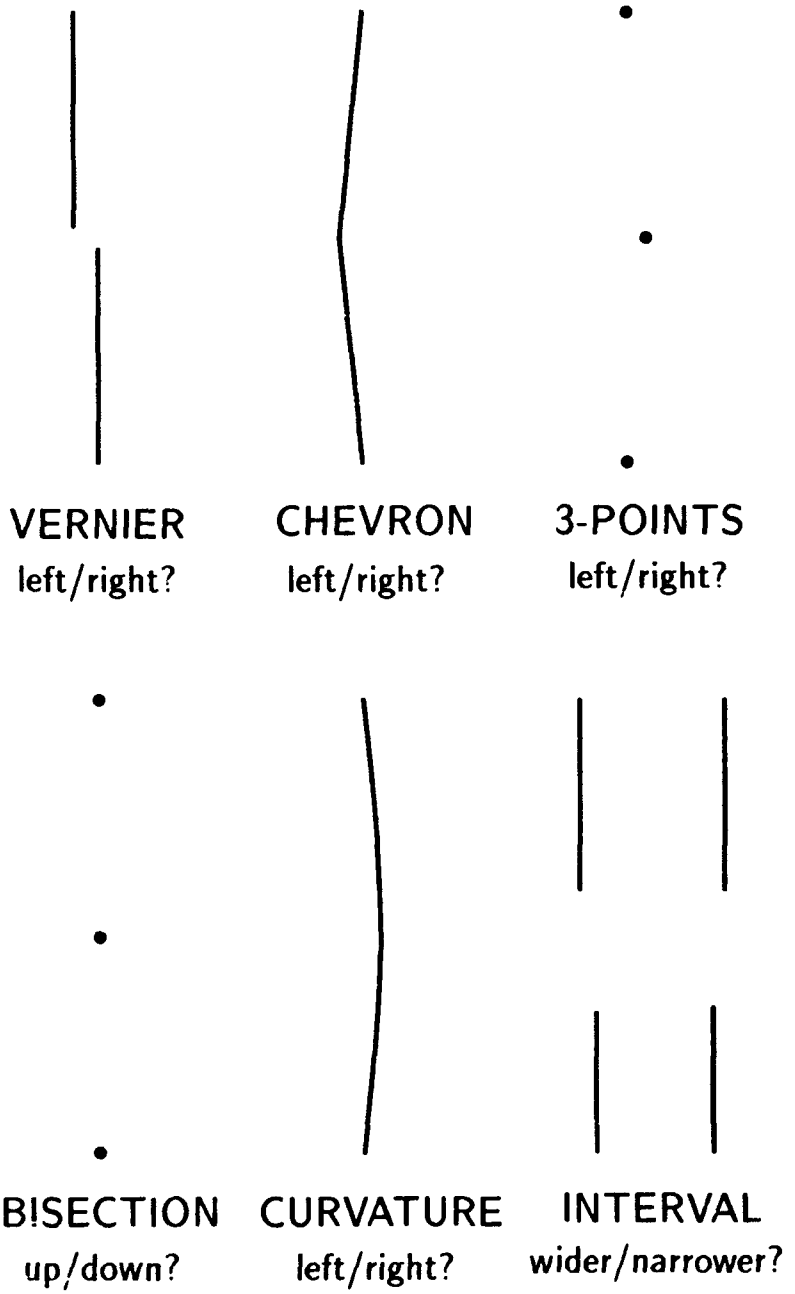

Figure 1: Examples of six tasks in which human subjects perform at hyperacuity levels (that is, exhibit resolution finer than the spacing between individual photoreceptors). Many other variations are possible, such as, for instance a horizontal vernier.

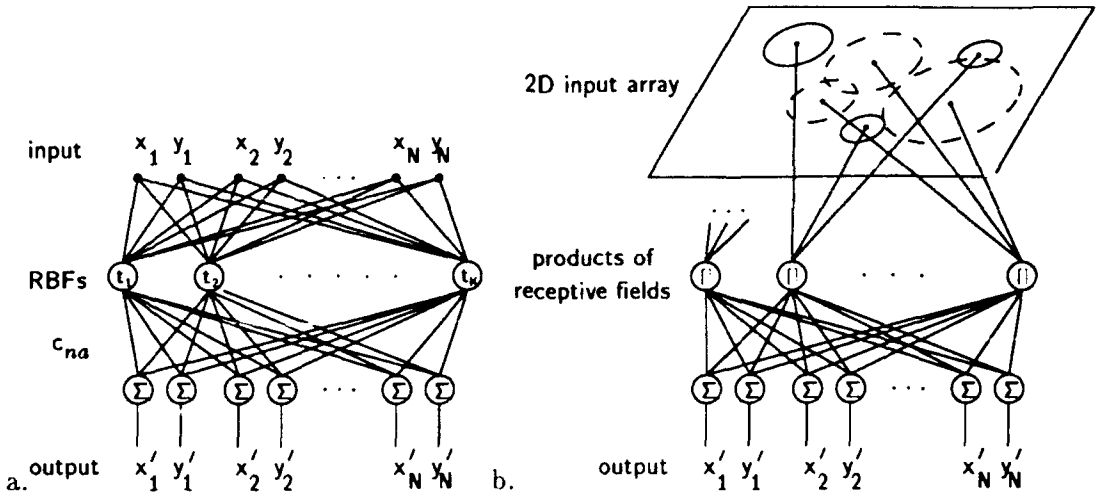

Figure 2: (a) A network representation of approximation by Hyper Basis Functions. (b) shows an equivalent interpretation of (a) for the case of Gaussian radial basis functions. Gaussian functions can be synthesized as the product of two-dimensional Gaussian receptive fields operating on retinotopic maps of features. The solid circles in the image plane represent the 2D Gaussians associated with the first radial basis function, which represents the first view of the object. The dashed circles represent the 2D receptive fields that synthesize the Gaussian radial function associated with another view. The Gaussian receptive fields transduce positions of features, represented implicitly as activity in a retinotopic array, and their product "computes" the radial function without the need of calculating norms and exponentials explicitly.

[18], . The module is an approximation of a multivariate function in terms of basis functions with parameter values that have to be found –i.e. "learned" – from the data – i.e. the examples. The expansion has the form

$$f^*(\mathbf{x}) = \sum_{\alpha=1}^{n} c_\alpha G(\|\mathbf{x} - \mathbf{t}_\alpha\|_W^2) + p(\mathbf{x}) \tag{1}$$

where the parameters $\mathbf{t}_\alpha$ that correspond to the centers of basis functions, and the coefficients c_α are unknown, and are in general much fewer than the data points ($n \leq N$). The norm is a *weighted norm*

$$\|\mathbf{x} - \mathbf{t}_\alpha\|_W^2 = (\mathbf{x} - \mathbf{t}_\alpha)^T W^T W (\mathbf{x} - \mathbf{t}_\alpha) \tag{2}$$

where W is an unknown square matrix and the superscript T indicates the transpose. In the simple case of diagonal W the diagonal elements w_i assign a specific weight to each input

coordinate, determining in fact the units of measure and the importance of each feature [19]. Equation 1 can be implemented by the network of Figure 2. The parameters $\mathbf{c}, \mathbf{t}, \mathbf{W}$ are searched for during learning by minimizing the error functional defined as

$$H[f^*] = H_{\mathbf{c},\mathbf{t},\mathbf{W}} = \sum_{i=1}^{N} (\Delta_i)^2,$$

where

$$\Delta_i \equiv y_i - f^*(\mathbf{x}) = y_i - \sum_{\alpha=1}^{n} c_\alpha G(\|\mathbf{x}_i - \mathbf{t}_\alpha\|_{\mathbf{W}}^2).$$

Iterative methods of the gradient descent type can be used for the minimization of H. An even simpler method that does not require calculation of derivatives is to look for random changes (controlled in appropriate ways) in the parameter values that reduce the error (cf. [14,2]). The interpretation of the network of Figure 2 is the following. The centers of the basis functions are similar to prototypes, since they are points in the multidimensional input space. Each unit computes a (weighted) distance of the inputs from its center and applies to it the radial function. In the case of the Gaussian, a unit will be the most active when the input exactly matches its center. The output of the network is a linear superposition of the activities of all the basis functions, plus direct, weighted connections from the inputs (the linear terms of $p(\mathbf{x})$) and from a constant input (the constant term). Notice that in the limit case of the basis functions approximating delta functions, the system becomes equivalent to a look-up table holding the examples.

2 Example: simulated experiments in hyperacuity

In the preceding section we have proposed that hyperacuity may consists of tasks learned by subjects from a few examples (mostly) in the psychophysical lab, exploiting modules (in the cortex?) that perform learning from examples, that is multivariate function approximation from sparse data. This hypothesis, if pushed to its extreme version, may represent a rather general framework for psychophysical modeling. To justify this proposal, we have conducted a series of simulated psychophysical experiments, in which a HyperBF module has been trained to perform several different hyperacuity tasks. The details of the experiments are described below.

2.1 Simulation details

The input to the module was an array of "photoreceptors" whose activity corresponded to the input image, blurred by the eye's optics. There were eight receptors, positioned randomly on a

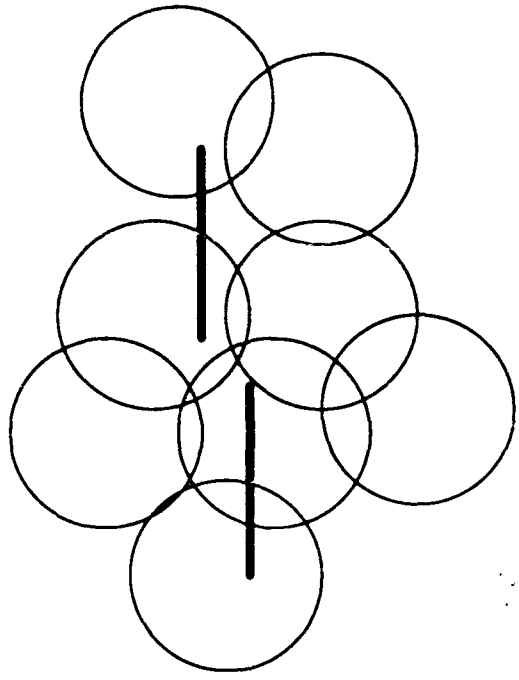

Figure 3: An illustration of the vernier acuity task: the subject has to tell whether the upper bar is to the left or to the right of the lower one. Human subjects (and the HyperBF simulation) perform this task at hyperacuity levels, that is, the minimum discernible horizontal displacement of the two bars is much smaller than the average distance between adjacent photoreceptors. The photoreceptor mosaic is shown superimposed on the stimulus. Each cone is shown as a circle that represents the Gaussian spread of a point source shining at the corresponding retinal location. This spread is due to the low-pass characteristics of the optics of the eye. Our simulation does not require positioning the "receptors" at precisely defined locations.

loose 4×2 grid (see Figure 3). Each of the receptors calculated its response by integrating the input over a region of the "retina" shaped as a Gaussian, with two space dimensions ($\sigma = 30''$) and one time dimension ($\sigma = 0.5$ units). The space dimensions spanned the entire $180'' \times 360''$ patch of the "retina", while the time dimension had an extent of ± 1 unit. The 8-component vector of receptor outputs constituted the input to the HyperBF module, which was trained to produce an output of $+1$ for one sense of the input vernier displacement, and -1 for the other.

The performance of the module was estimated by measuring the absolute error, that is, the distance between the actual output (which could be any number between -1 and $+1$; for a proof see [6]) and the desired output (± 1; see Figure 4). Without going into the details, we point out that the absolute output error is a good analog of acuity threshold, since the two are related monotonically.

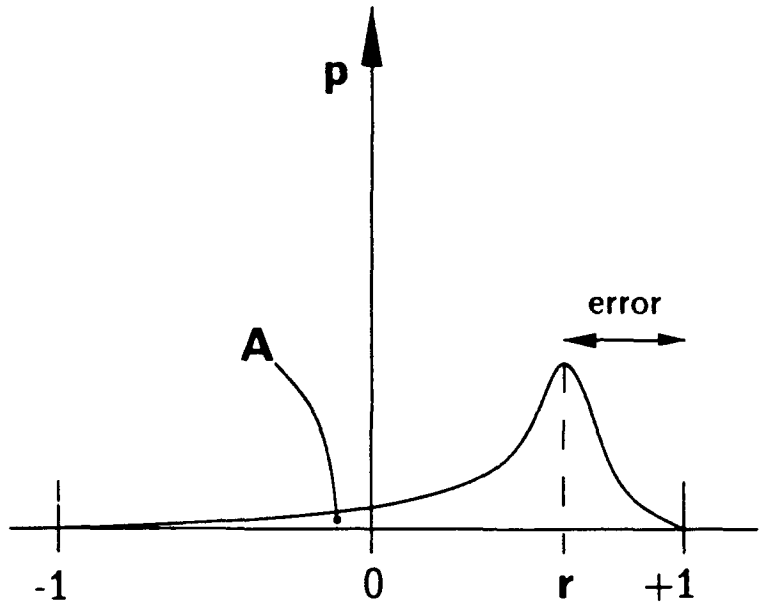

Figure 4: The relationship between the performance index used in the simulations — the absolute output *error* of the HyperBF module — and the acuity threshold. The probability density of the output is shown as a distribution centered, say, at $r > 0$, whose tail extends across 0 to the other half of the ± 1 range of possible values. The area A under the tail of the distribution indicates the probability of erroneous response, given the statistics represented by the mean and standard deviation of *error* (the two parameters we have measured in the simulations). The acuity threshold, in turn, can be related to the probability of erroneous response through probit analysis.

2.2 Replication of the basic psychophysical findings for the vernier task

The HyperBF module coupled to the input mechanism described above successfully replicated, after a training phase typically consisting of about 50 "examples", the following four basic findings of the psychophysics of hyperacuity in human subjects:

- The equivalent acuity threshold was significantly lower than the spacing of the receptors in the simulated retina ([10,22]; Figure 5).
- The threshold improved with increasing vertical separation of the two segments comprising the vernier stimulus ([24]; Figure 6). We note that in human subjects this improvement reverts with further increase in the vertical separation; this phenomenon was also replicated by the model.

- The threshold deteriorated with increasing orientation difference between training and testing trials. This deterioration was more pronounced for shorter stimuli ([21]; Figure 7).

- Performance remained at hyperacuity levels when the stimuli moved across the retina, and was the highest when the velocity of the stimulus translation was the same during training and testing ([23]; Figure 8).

Importantly, the hyperacuity-level performance was independent of the precise location of the receptors. At the same time, different quasi-random receptor mosaics yielded different thresholds, sometimes by as much as a factor of two. A similar range of hyperacuity thresholds is observed in human subjects, even at full acuity and perfectly normal eyes.

2.3 Comparison among line vernier, three-point bisection and dot vernier tasks

The next experiment compared the performance of an HyperBF module in the vernier task with that in another hyperacuity task, the three-point bisection. The stimulus in the bisection task consists of three dots, arranged in a vertical line, at an approximately even spacing. The subject has to determine whether the middle dot is above or below the midpoint of the segment formed by the other two dots. The HyperBF module learned this hyperacuity task just as easily as it did in the line vernier case.

Another experiment made a comparison between the line vernier task and a similar one in which each of the line segments has been replaced by two dots (situated at its endpoints). The network learned this task, as it did previously in the line vernier and the bisection cases. The comparison between the two vernier tasks appears in Figure 9. The better performance of the HyperBF module in the dot vernier task for small X-offsets parallels a recent surprising finding with human subjects (M. Fahle, personal communication).

2.4 Replication of the decrease of vernier threshold with practice

A major characteristic of human performance in hyperacuity tasks is the gradual and constant improvement of the threshold, which continues, albeit at a slow rate, after ten thousand trials ([9]; see the appendix). We have replicated this phenomenon by endowing the model with a learning mechanism that we call "incremental learning" (see also [3]) and that consists of two phases. First, gradual improvement was obtained by letting the model perform a local random search in the space of HyperBF center coordinates. Second, when the model's performance on a new input was markedly inadequate (in comparison with recent history), that input was adjoined to the model as an additional center (prototype). In the appendix we discuss how

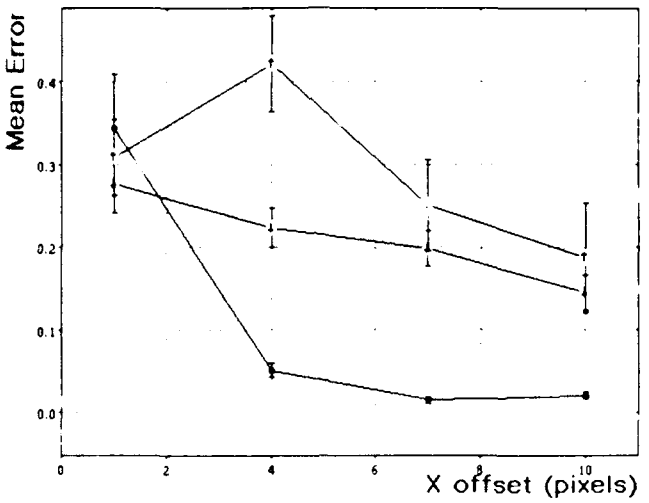

Figure 5: Mean error of the synthesized module vs. X-offset of the vernier stimulus. The module was trained to output 1 for left offset and -1 for right offset. Consequently, error of 0.1 corresponds to high performance (bars in this and other figure denote ± 1 standard error of the mean) The values of X-offset along the abscissa are the lower bounds of an octave range (e.g., 4 pixels means that the offsets were uniformly distributed between 4 and 8 pixels; in all our simulations the scale was 10 pixels to 30″). The three curves correspond to three training/testing combinations. In the first one (•), the same X-offset range was used both for training and testing. In the other two combinations ($\downarrow$ and $\uparrow$), the testing range was one-half and twice as large as the training range, respectively. Note that X-offsets which yielded high performance (mean error smaller than 0.05) are much smaller than the photoreceptor spacing (6″, compared to about 30″).

the incremental learning algorithm can be naturally extended to work even without explicit examples, that is without feedback, for appropriate tasks.

The algorithm for adjusting the positions of the existing centers was as follows. For each new input, the system made between 10 and 100 random changes in the value of a randomly chosen coordinate of a center (the amplitude of the change was about ten percent of coordinate value). After each change the error for that particular input was recalculated. If the new error was lower (and, with a small probability, if the error increased), the change was incorporated into the system, otherwise the change was reversed (cf. [2]).[1] If at any stage during the simulated

[1]The probability of keeping a change that led to a higher error could be decreased with time, as in the

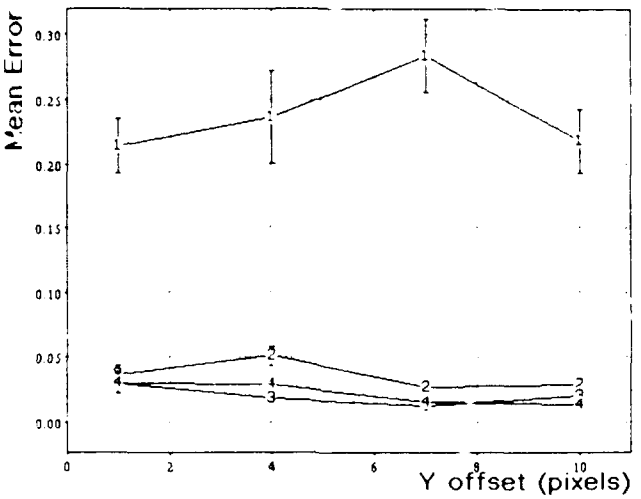

Figure 6: Mean error of the synthesized module vs. Y-offset of the vernier stimulus, by X-offset. The four curves correspond to four values of X-offset. Once the X-offset is high enough to guarantee good performance (curves 2, 3 and 4), increasing the Y-offset improves the performance level, as it does in human subjects.

experiment the current input was too distant from any of the existing HyperBF centers, that input was adjoined to the model as a new center (cf. learning by example acquisition in the CLF model of object recognition ([8, 9]; see also [18]). The performance of the resulting algorithm that combined adjustment of existing centers with recruitment of new centers is shown in Figure 10.

3 Conclusions

3.1 Discussion

The skeleton model described in the preceding sections is specific enough to be put to a psychophysical test. One possible way to do so is to test the prediction of the model regarding generalization of performance from a well-practiced to an unfamiliar range of inputs. Consider, for concreteness' sake, the vernier acuity task. If the human visual system relies on a memory-based mechanism such as HyperBF interpolation to solve this problem, a drop in performance (that is, an increase in the error rate) is expected when the range of the stimuli is suddenly

simulated annealing approach to optimization [12] (this feature, however, appeared to be unnecessary for our purposes).

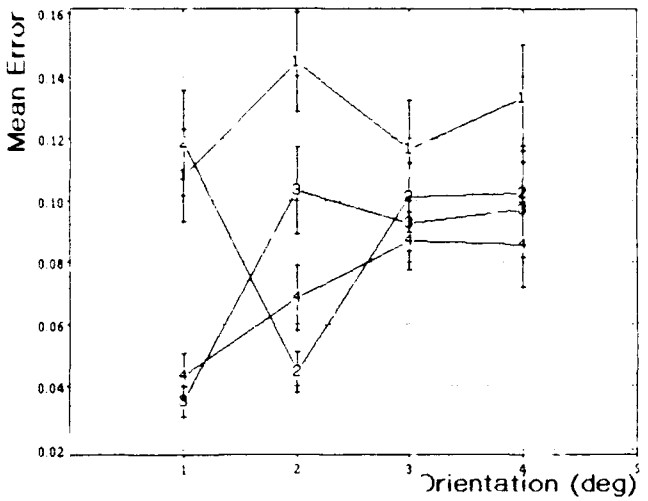

Figure 7: Mean error of the synthesized module vs. orientation of the stimulus (shown along the abscissa as the lower bound of a 1-octave range, in degrees), by stimulus length. X-offset was between 4 and 8 pixels (12″ to 24″), Y-offset was 1 pixel (3″). The four curves correspond to four values of segment length, from 10 to 40 pixels (30″ to 80″). In general performance is seen to deteriorate with increased orientation range.

changed (e.g., if the verniers are made smaller by a factor of two or more in comparison with their values during training). The same prediction holds for a change to a different hyperacuity task (say from the top left stimulus in Figure 1 to the bottom right one). If regression analysis is used to obtain an estimate of the psychometric function from error rates, such a change in the stimulus range would cause a decrease in the coefficient of determination of the regression, or in related measures of the goodness of fit. Moreover, the subsequent recovery of performance should be slower if no feedback is provided after the change (even though some learning appears to be possible even without explicit feedback; see the appendix). There are preliminary indications that both these phenomena indeed happen in practice [9].

No such response to a change in the stimulus range should be found if the visual system has a built-in scale invariance mechanism. Different versions of scale-invariant models of early visual processing have been offered in the past (e.g., [20]). For our present purpose, a simple scheme, in which invariance is achieved through simultaneous processing of the input at several levels of resolution (corresponding to several overlapping grids of "ganglion" cells of different size and spacing), would suffice. In such a case, the system could be prepared in advance, say, to a reduction in the input scale (up to a certain limit), simply because the small-scale grid

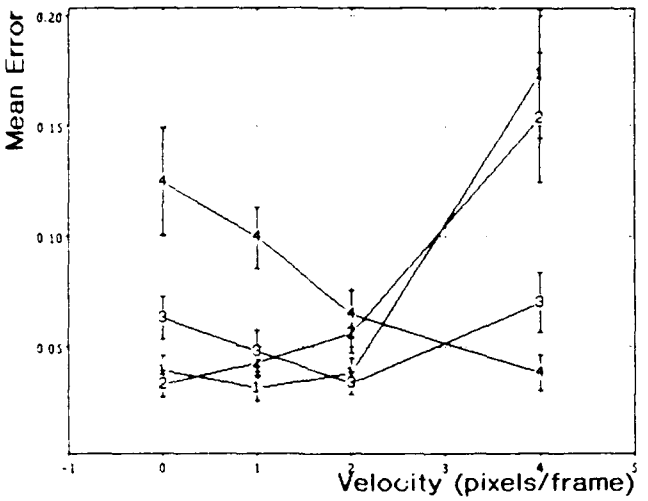

Figure 8: Mean error of the synthesized module vs. velocity of the stimulus during testing. X-offset was between 4 and 8 pixels (12″ to 16″), Y-offset was 1 pixel (3″). The four curves correspond to four values of velocity during training (same set of 4 values as the testing velocities). In general, performance deteriorates with increased testing velocity, but to a lesser extent if the training velocity was relatively high as well.

would exhibit, after the reduction, a pattern of activity isomorphic to the pattern evoked by the large-scale input in the large-scale grid (see Figure 11). Finally, we remark that the mechanism underlying scale invariance (if any) could be probed by blurring the input to the extent that the small-scale grids, but not the large-scale ones, are affected.

In general, we expect that the cortex performs suitable pre-processing to provide approximate invariance to certain basic transformations, without the need for explicit learning. Translation, in addition to scale, is another obvious candidate transformation for which invariance could be built in. The bare version of our network, described here, would not generalize from one patch of the retina to another (though this may not be fully necessary; cf. [15]). It seems likely that translation invariance, at least up to a certain extent, should be provided by mechanisms preceding the learning stage (possibly related to the "focus of attention" idea). It is possible that preprocessing mechanisms could also provide invariance to the specific stimulus type by computing the equivalent of "place tokens". This would enable the system to generalize automatically (without the need for examples) from, say, line stimuli to, say, dot stimuli, but would, of course, void to some extent the significance of our model. In any case, the input to a learning model such as the one we have outlined should not be raw photoreceptor

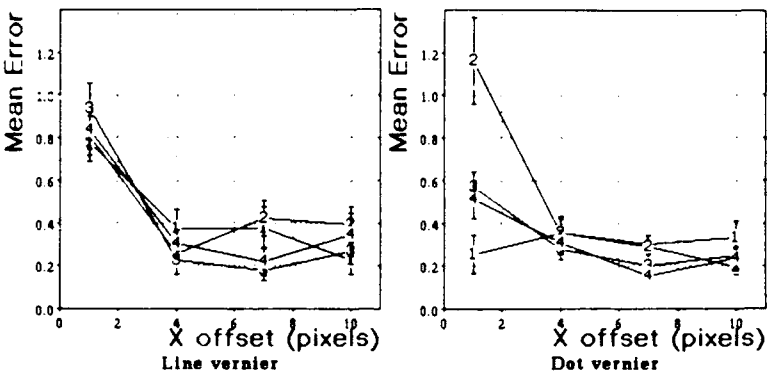

Figure 9: Mean error of the synthesized module vs. X-offset of the vernier stimulus, by Y-offset. *Left:* line vernier stimulus. *Right:* dot vernier stimulus. Note better performance in the latter case for small X-offsets.

activities, as in our simulations, but rather pre-processed photoreceptor activities. The type of preprocessing in human vision and the associated pseudo-invariances it supports are an experimental question of great interest. Of course, any lack of generalization would be support for the model. Experimental demonstration of tranfer of learning with respect to translation and scale, would not represent in our view a major problem for the model, though it would require a more complex preprocessing than the one we have simulated. Transfer of learning from one type of stimulus to another (see Figure 1) would be a more serious blow to the spirit of our model and therefore a more critical test of its validity.

3.2 Summary

The specific implication of this work is that human-like performance in different hyperacuity tasks can be obtained by modules synthesized "on the fly" from a few examples of that task. In view of the results reported above, we conjecture that the module responsible for hyperacuity-level performance is synthesized in a demand-driven fashion, when the task is first performed by the subject.

More generally, one may apply the same line of reasoning to other visual tasks studied by psychophysicists. To this effect, it is important that the technique we have used for learning can be implemented as a simple biologically plausible network [19]. Furthermore, this approach has recently been demonstrated as effective in modeling central aspects of human performance in three-dimensional object recognition [17,6]. It remains to be seen whether the above framework

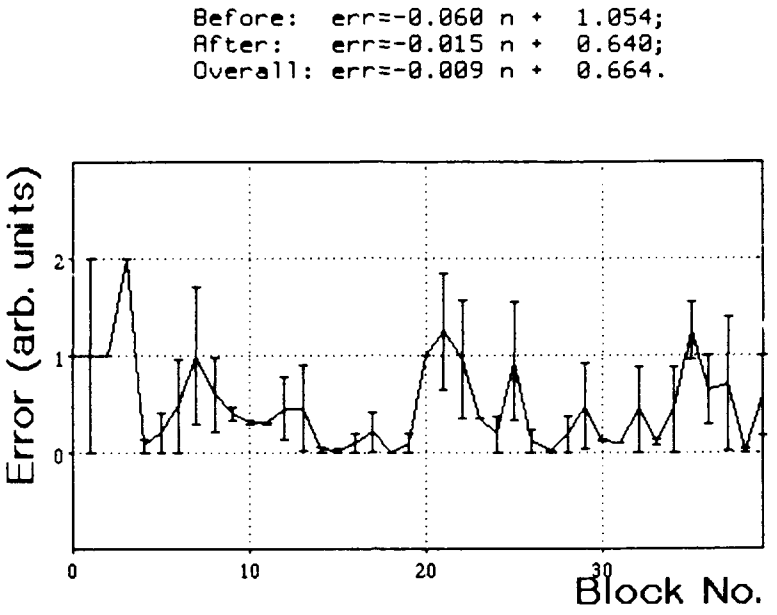

Figure 10: Replication of the gradual improvement in vernier acuity with practice by the random search technique described in section 2.4. The best linear fit to the data set has a slope of -0.009. In other words, the HyperBF module exhibited continuous learning, just as the human subjects did.

would prove useful in unifying the existing diverse theoretical approaches to the modeling of visual perception, and of brain function in general.

Appendix: Learning modes of the HyperBF scheme

Incremental learning and bootstrapping in the absence of feedback

The HyperBF module must be allowed to improve its performance throughout the testing stage, with and without feedback. This can be achieved by using the algorithm that we described in the body of the paper: centers are added when the model performance is inadequate. Coefficients are modified and - possibly on a slower time scale - centers are moved. Performance is easily measured in the presence of feedback, in terms of the error between the predicted value and the correct one. If no feedback is available, it is still possible to estimate performance if the new input is not too far away from the existing centers, so that the network can classify it correctly,

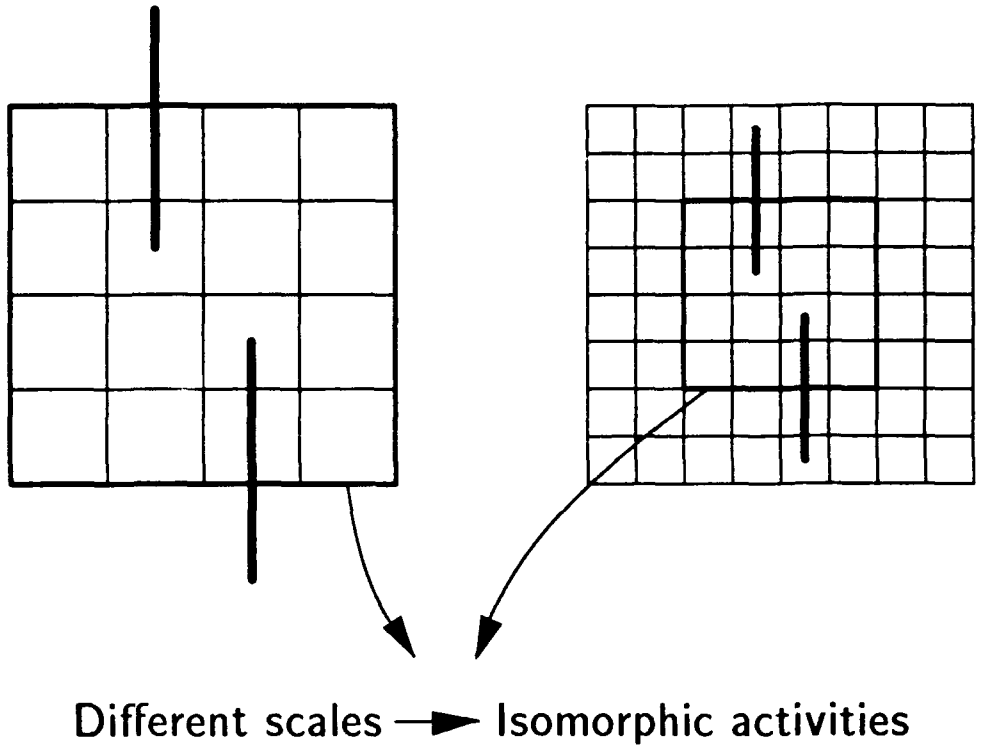

Figure 11: Scale invariance in hyperacuity tasks can be achieved in principle through simultaneous processing of the input at several levels of resolution, corresponding to several overlapping grids of "ganglion" cells of different size and spacing.

even if not very reliably. Thus, a small modification of the scheme makes it to work in the absence of feedback, under certain conditions. Imagine that a few examples of the hyperacuity task are given with feedback, that is with the correct classification. Subsequently, new stimuli are given without feedback. If these stimuli are sufficiently similar to the original examples, the network may be able to classify them correctly and then incorporate them as new centers (i.e., templates), effectively bootstrapping the learning process.

Notice that such incremental learning tasks are not uncommon. In particular, hyperacuity is often tested within the paradigm of adapting the size of the offset to the subject's performance, therefore decreasing it slowly during the test. Under these conditions, the offset in each trial is never less than half the offset of the previous trial. According to our simulations, the network described earlier can generalize rather well to offsets of half the size (but not to offsets of say, four times the training size). The incremental learning algorithm described in the main text may be extended in the following way. In the absence of feedback the network attempts to

classify a new stimulus and to use it as an example for incremental learning, provided that the classification is sufficiently reliable (that is, provided there is at least one unit in the network which is sufficiently active, indicating that the new stimulus is sufficiently close to one of the existing centers).

Learning algorithms: details

The basic mechanism of learning in HyperBF networks is the computation of the optimal set of coefficients $\mathbf{c}_\alpha$ which relate the network's output vector to the vector whose components are the activities of the individual basis function units. Finding the matrix of coefficients amounts to the solution of a linear system, provided that the number of input/output examples is the same as the number of basis functions. If there are more examples than basis functions, the resulting overconstrained system can be solved by pseudoinverse methods. A one-shot method of this type does not appear to be biologically plausible. However, an equivalent result may be achieved, for the case of $\mathbf{c}_\alpha$, by gradient descent that can be implemented through a Hebbian mechanism (see [18]).

In the overconstrained case repositioning the HyperBF centers $\mathbf{t}_\alpha$ through gradient descent can also improve the module's performance (see also section 2.4). To cite a concrete example, we have trained a 20-center network with 50 vernier examples, achieving mean error of 0.67 ± 0.07. After 20 steps of gradient descent, the error dropped to 0.045 ± 0.006.

In a more realistic situation, the HyperBF module should be allowed to improve its performance not only during specially designated training trials, but also throughout the testing stage as it is the case for our incremental learning algorithm described in the main text. Our simulation is based on a random search method described in section 2.4, and on augmenting the HyperBF module with a Widrow-Hoff learning mechanism (see [25]), in which the coefficients $\mathbf{c}_\alpha$ are modified according to the following formula:

$$\mathbf{c}^{t+1} = \gamma \mathbf{c}^t \left(\mathbf{f}^t - \hat{\mathbf{f}}^t \right) \mathbf{h}^t$$

where $\mathbf{f}^t$ and $\hat{\mathbf{f}}^t$ are the correct and the estimated output values at trial t, and $\mathbf{h}^t$ is the vector of intermediate-layer values (which are the activities of the basis units). In other words, the coefficients $\mathbf{c}_\alpha$ are modified by an amount proportional to the error made in the current trial. It has been shown [25,11] that the Widrow-Hoff mechanism is equivalent to an incremental computation of the appropriate pseudoinverse. In our simulations, mean error typically improved by 0.004 per trial for about 100 trials (as found by a linear regression of error on trial number), then became constant.[2]

[2]These figures varied with the coefficient γ of the Widrow-Hoff equation.

References

[1] H. B. Barlow. Reconstructing the visual image in space and time. *Nature*, 279:189–190, 1979.

[2] B. Caprile and F. Girosi. A non-deterministic minimization algorithm. A. I. Memo 1254, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, Sept. 1990.

[3] B. Caprile, F. Girosi, and T. Poggio, 1991. in preparation.

[4] C. R. Carlson and R. W. Klopfenstein. Spatial frequency model for hyperacuity. *Journal of the Optical Society of America*, A2:1747–1751, 1985.

[5] F. H. C. Crick, D. C. Marr, and T. Poggio. An information-processing approach to understanding the visual cortex. In F. Schmitt, editor, *The organization of the cerebral cortex*. MIT Press, Cambridge, MA, 1980.

[6] S. Edelman and T. Poggio. Bringing the Grandmother back into the picture: a memory-based view of object recognition. A.I. Memo No. 1181, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1990.

[7] S. Edelman and D. Weinshall. A self-organizing multiple-view representation of 3D objects. *Biological Cybernetics*, 64:209–219, 1991.

[8] S. Edelman and D. Weinshall. A self-organizing multiple-view representation of 3D objects. A.I. Memo No. 1146, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, August 1989.

[9] M. W. Fahle, S. Edelman, and T. Poggio. Learning of vernier acuity, 1991. in preparation.

[10] E. Hering. Ueber die Grenzen der Sehschaerfe. In *Bericht. Mathem.-Physikal. Klasse Saechs.*, page 16. Ges. Wissenschaften, Leipzig, 1899.

[11] A. Hurlbert and T. Poggio. Learning a color algorithm from examples. A.I. Memo No. 909, CBIP Paper 25, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Ma., 1987.

[12] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[13] S. A. Klein and D. M. Levi. Hyperacuity thresholds of 1 sec: theoretical predictions and empirical validation. *Journal of the Optical Society of America*, A2:1170–1190, 1985.

[14] S. Lin and B. W. Kernighan. An effective heuristic algorithm for the traveling salesman problem. *Operations Research*, 21:498–516, 1973.

[15] T. Nazir and J. K. O'Regan. Some results on translation invariance in the human visual system. *Spatial vision*, 5:81–100, 1990.

[16] T. Poggio. A theory of how the brain might work. Proc. Cold Spring Harbor meeting on Quantitative Biology and the Brain.

[17] T. Poggio and S. Edelman. A network that learns to recognize three-dimensional objects. *Nature*, 343:263–266, 1990.

[18] T. Poggio and F. Girosi. A theory of networks for approximation and learning. A.I. Memo No. 1140, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1989.

[19] T. Poggio and F. Girosi. Regularization algorithms for learning that are equivalent to multilayer networks. *Science*, 247:978–982, 1990.

[20] E. L. Schwartz. Local and global functional architecture in primate striate cortex: outline of a spatial mapping doctrine for perception. In D. Rose and V. G. Dobson, editors, *Models of the visual cortex*, pages 146–157. Wiley, New York, NY, 1985.

[21] R. Watt and F. W. Campbell. Vernier acuity: interactions between length effects and gaps when orientation effects are eliminated. *Spatial Vision*, 1:31–38, 1985.

[22] G. Westheimer. The spatial sense of the eye. *Invest. Ophthal. Vis. Sci.*, 18:893–912, 1979.

[23] G. Westheimer and S. P. McKee. Visual acuity in the presence of retinal image motion. *Journal of the Optical Society of America*, 65:847–850, 1975.

[24] G. Westheimer and S. P. McKee. Spatial configurations for visual hyperacuity. *Vision Research*, 17:941–947, 1977.

[25] B. Widrow and S. D. Stearns. *Adaptive signal processing*. Prentice Hall, Englewood Cliffs, NJ, 1985.

[26] H. R. Wilson. Responses of spatial mechanisms can explain hyperacuity. *Vision Research*, 26:453–469, 1986.