Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Vision Research 44 (2004) 479–482
Vision Research
www.elsevier.com/locate/visres

Reply to Gosselin and Schyns

Richard F. Murray [a,*], Jason M. Gold [b]

[a] Center for Perceptual Systems, University of Texas at Austin, 1 University Station A8000, Austin, TX 78712-0187, Austin, USA
[b] Department of Psychology, Indiana University, 1101 East 10th Street, Bloomington, IN 47405, USA

Received 29 September 2003

Abstract

We discuss Gosselin and Schyns' (2003) reply to our criticisms and constructive suggestions concerning the bubbles method Murray and Gold (2003). We find that their reply does not mollify our concerns, and we still believe that reverse correlation will generally be preferable to the bubbles method until further developments (a) demonstrate more clearly what the bubbles method actually measures and (b) introduce a type of windowing noise that is less likely to disrupt observers' strategies.

1. Our LAM analysis of bubbles is too limited to be of any value

Gosselin and Schyns' reply does not alleviate our concerns about the bubbles method in its present form. Here we discuss their most important claims, and explain why we do not think that they adequately address our criticisms and constructive suggestions.

Gosselin and Schyns argue that "the LAM is not sufficiently general in scope to impose any prescriptive standards on the conduct of research in visual categorization", and hence that our LAM-based analysis is not useful. Certainly the LAM is an incomplete model of human performance, but this sweeping judgement is far too dismissive. As we said in our article, the LAM is a useful first-order approximation that captures many aspects of human performance, and serves as a starting point for more complex models. Furthermore, many nonlinear models are locally linear, which means that a linear analysis is often adequate in psychophysical tasks where the stimuli cover only a narrow range (Ahumada, 1987). Even in high-level tasks like face and letter identification, an analysis of observers' performance in terms of templates and internal noise can lead to robust and surprising results (e.g., Gold, Bennett, & Sekuler, 1999; Solomon & Pelli, 1994; Tjan, Braje, Legge, & Kersten, 1995).

Furthermore, the LAM and our LAM-based analysis are not as restrictive as Gosselin and Schyns imply. They claim that our analysis can only be applied to tasks in which the bubbles window small spatial regions, rather than windowing regions in some more abstract representation of the stimulus, such as a scale space. This is simply not true. We described the bubbles method in terms of spatial bubbles, because this is the approach that Gosselin and Schyns have used in almost all their work, but our analysis used a very general and abstract description of the bubbles method. Our analysis can be used whenever the observer's decision variable is regarded as a linear, Gaussian-noise contaminated function of the stimulus in *some* representation, however abstract. (In particular, the representation could be related to the photometric representation by the arbitrarily complex morphing operations mentioned by Gosselin and Schyns.)

In any case, our LAM-based analysis is the *only* rigorous analysis of the bubbles method to date. Gosselin and Schyns criticize our approach as being too simplistic, but as we made clear, we believe that it is a limited but useful first step in putting their method on a rigorous footing. For instance, we showed that their RAP law is valid for linear observers, and this led us to note that in general, for nonlinear observers, it is *not* valid—surely a useful contribution, as they themselves often make use of their RAP law, but have never discussed either its justification or its domain of validity.

* Corresponding author. Address: Department of Psychology, University of Pennsylvania, 3815 Walnut Street, Philadelphia, PA 19104-6196, USA. Tel.: +1-215-898-7300; fax: +1-215-746-6848.
E-mail address: murray@psych.upenn.edu (R.F. Murray).

0042-6989/$ - see front matter
doi:10.1016/j.visres.2003.10.008

Furthermore, we did suggest that for anyone interested in using the bubbles method, it is important to extend our LAM analysis to more complex models. Thus our credo is not the younger, naïve Candide's "*tout est pour le mieux*", but his later conclusion after long experience that "*il faut cultiver notre jardin*" [1]—in this case, a garden of bubbles.

2. A bubbles image completely reveals an observer's template

Gosselin and Schyns accept our proof that for a LAM observer, the expected value of a bubbles image is given by our equation (5):

$$E[B] = u + v \cdot b * b * (T \circ (I_X - I_Y))$$

They argue, nevertheless, that a bubbles image completely recovers a LAM observer's template in all interesting cases. They suggest that one can recover the template by using a single-pixel bubble, so that the double-convolution disappears, and then by dividing the bubbles image pointwise by the ideal template, $I_X - I_Y$, leaving a term proportional to the template T except at points where $I_X - I_Y = 0$, at which locations the division is undefined.

We have two initial objections that are important to note, but that can be met. First, the term u must first be subtracted from the bubbles image for this scheme to work. It can be calculated easily; see our Appendix (Murray & Gold, 2003). Second, no bubbles experiment has ever been carried out with single-pixel bubbles, and several of the reported experiments actually rely on using several sizes of bubbles, so it is not clear how usable this single-pixel scheme is in bubbles experiments as they are actually practised. However, even if a multi-pixel bubble is used, the effects of the double-convolution can be undone by deconvolution. The resulting signal-to-noise ratio will be low at high spatial frequencies, but nonetheless, with some effort this approach might be made to work.

Our main objection is that, contrary to Gosselin and Schyns' claims, even a cursory review of the reverse correlation literature shows that observers' responses are *often* influenced by stimulus locations where the ideal template is zero. Observers use irrelevant landmarks in vernier alignment tasks (Ahumada, 1996), illusory and occluded contours in shape discrimination tasks (Gold, Murray, Bennett, & Sekuler, 2000), irrelevant, uncued locations in attention tasks (Shimozaki, Eckstein, & Abbey, 2002), and uninformative pre- and post-stimulus intervals in detection tasks (Neri & Heeger, 2002). Furthermore, observers invariably use uninformative stimulus regions surrounding informative regions, probably because of spatial uncertainty. It is impossible to recover any of these regions of observers' templates with Gosselin and Schyns' scheme. Consequently, as we claimed, a bubbles image does not completely determine a LAM observer's template.

Moreover, a moment's reflection shows that Gosselin and Schyns' scheme for recovering a LAM observer's template from a bubbles image is ill-conditioned not only at stimulus locations where the ideal template is zero, but also where it is *near* zero. In a bubbles image, these two types of locations will generally have almost indistinguishably close values, because neither of them greatly help the observer to give a correct response. Dividing the bubbles image by near-zero locations in the ideal template will magnify the inevitable statistical noise enormously, and the resulting estimate of the template will be practically useless.

Thus, contrary to Gosselin and Schyn's claim, reverse correlation does recover much more information about LAM observers than the bubbles method does.

3. The bubbles method does not change observers' strategies

Gosselin and Schyns carry out a face identification experiment to compare the strategies that observers use in bubbles and reverse correlation experiments, and they conclude that the strategies are roughly similar. Their interpretation of this experiment is fatally flawed in two ways. First, Gosselin and Schyns assume that a LAM observer in their categorization task can have only a single template. In a categorization task with many possible stimuli and just two responses, a LAM observer is normally assumed to have a stored template for each possible stimulus (e.g., Peterson, Birdsall, & Fox, 1954; Tjan et al., 1995). In this case, the calculation of the classification image is difficult, and must be done separately for each possible stimulus-response pair (Watson, 1998). Thus Gosselin and Schyns miscalculate the classification image: one cannot simply sum the noise images within each response category, and take the difference of these sums, as they do.

Second, and more crucially, Gosselin and Schyns make a basic logical error. We claim that, in many tasks, the bubbles method will change observers' strategies, and we have shown that this is demonstrably true in at least one task (the fat–thin task). Thus we have shown that, in general, one must be concerned that the bubbles method may change observers' strategies. Gosselin and Schyns show that in one task the bubbles method (perhaps) does not change observers' strategies, but from this they cannot conclude that in general, it does not change observers' strategies. In fact, our single counterexample (the fat–thin task) shows that it sometimes *does*. Unfortunately, they do not discuss our experiment,

[1] We must cultivate our garden.

so we do not know why they think it is inconclusive. (They do not discuss our experiment with respect to the question of whether a bubbles image completely recovers an observer's template because they believe that there are too many zero-valued pixels in the ideal template for the fat–thin task. However, this objection has nothing to do with the question of whether the bubbles method changes observers' strategies.)

Finally, Gosselin and Schyns discuss the three theoretical reasons that we gave to support our claim that the bubbles method is more likely to change observers' strategies than reverse correlation.

(a) We argued that the obliteration of large, randomly chosen parts of the stimulus is more likely to make observers change their strategies from trial to trial, than is adding Gaussian white noise. They argue that bubbles are more disruptive than Gaussian noise only if the bubbles are too large. Our response is that (i) bubbles experiments to date have used a small number ($\sim$25) of large bubbles, so our objection was more than theoretical, and (ii) the idea of using many tiny bubbles is very similar to our suggestion that one should use multiplicative Gaussian white noise, rather than randomly placed bubbles, and we agree that this would be a useful modification to the bubbles method.

(b) We noted that many psychophysical and physiological experiments have shown that observers must contend with internal Gaussian noise, even when there is no external noise, and we suggested that moderate amounts of external noise are therefore unlikely to drastically change observers' strategies. They point out that it is also the case the parts of objects are often occluded, as in a bubbles experiment. However, occlusion seems usually to occur in the form of large, contiguous segments of objects being occluded by other objects. Do we normally identify faces through 25 small randomly placed holes in occluding surfaces? To us, the analogy seems more than a little strained.

(c) We noted that noise masking functions are typically linear, indicating that observers' sampling (i.e., template) efficiency is not drastically altered by adding external white noise. In reply, Gosselin and Schyns report a new face identification experiment showing that the threshold contrast energy of a (complete, pre-windowed) stimulus declines linearly as a function of the number of bubbles. Our response is that, first, it would be helpful to see a careful explanation of *why* this implies that observers' strategies are constant as a function of the number of bubbles. Essentially, this result shows that the on-screen contrast energy is constant as a function of the number of bubbles, and it is certainly plausible that this is the signature of a constant strategy. However, to take just one possible problem, it is not clear how an observer's uncertainty concerning the number and position of bubbles will complicate this picture. We cited the linearity of noise masking functions in our argument, because it can be shown to correspond to constant sampling efficiency as a function of external noise power (Burgess, Wagner, Jennings, & Barlow, 1981). No such results have ever been derived to aid in the interpretation of Gosselin and Schyns' experiment. Their interpretation is plausible, and we are willing to believe it, but it must be shown to be correct.

Our second objection is that in all bubbles experiments to date, Gosselin and Schyns have used around 25 medium-sized bubbles, whereas in this experiment they use 400–700 very small bubbles. In our view, the reason why the bubbles method changes observers' strategies is that when small, randomly chosen parts of the stimulus are shown from trial to trial, the observer will use whichever part is available on any given trial. When the stimulus is shown through a very large number of very small bubbles, the situation changes entirely: on any given trial, one or more bubbles are very likely to fall in any reasonably large stimulus region, and there will be less incentive for the observer to change his strategy from trial to trial. In fact, in the theoretical limit of an extremely large number of extremely small bubbles (e.g., the size of a monitor phosphor molecule), changing the number of bubbles is tantamount to simply changing the stimulus contrast. Thus the validation experiment was carried out under very different conditions than all bubbles experiments to date, and we think that it gives little support to the bubbles method as it has actually been used. (Again, the notion of using many tiny bubbles is very similar to our suggestion of using multiplicative white Gaussian noise, and we do agree that with this new modification, the bubbles method is less likely to disrupt observers' strategies.)

Finally, and most crucially, Gosselin and Schyns make the same fatal logical error as before: we have shown that in at least one task (the fat–thin task) the bubbles method does disrupt observers' strategies, and even if Gosselin and Schyns' interpretation of their face identification experiment is correct, citing one task where the bubbles method does not disrupt observers' strategies does little to refute the conclusion that, in general, one must be concerned that using the bubbles method in studying a novel task will disrupt observers' strategies. Again, they do not discuss our experiment, so we do not know why they think it is inconclusive.

4. Conclusion

We still believe that the bubbles method could be a useful addition to the current library of system identification methods. However, the problems with the method in its present form are serious, and they cannot be argued away. As we outlined in our article, we believe that the correct approach is to develop the method further, by analyzing it rigorously in the context of more sophisticated models of vision, and by experimenting with different forms of visual noise to find a way of minimizing the tendency of the bubbles method to disrupt observers' strategies.

References

Ahumada, A. J. (1996). Perceptual classification images from vernier acuity masked by noise. *Perception, 25*, 18.

Ahumada, A. J. (1987). Putting the visual system noise back in the picture. *Journal of the Optical Society of America A, 4*(12), 2372–2378.

Burgess, A. E., Wagner, R. F., Jennings, R. J., & Barlow, H. B. (1981). Efficiency of human visual signal discrimination. *Science, 214*(4516), 93–94.

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Signal but not noise changes with perceptual learning. *Nature, 402*(6758), 176–178.

Gold, J. M., Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2000). Deriving behavioral receptive fields for visually completed contours. *Current Biology, 10*(11), 663–666.

Gosselin, F., & Schyns, P. G. (2003). No troubles with bubbles. *Vision Research, 44*.

Murray, R. F., & Gold, J. M. (2003). Troubles with bubbles. *Vision Research, 44*.

Neri, P., & Heeger, D. J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nature Neuroscience, 5*(8), 812–816.

Peterson, W. W., Birdsall, T. G., & Fox, W. C. (1954). The theory of signal detectability. *Transactions of the IRE Professional Group on Information Theory, 4*, 171–212.

Shimozaki, S., Eckstein, M. P. & Abbey, C. (2002). Uncertain humans in a structurally certain world: attentional leaking with 100% valid postcues as seen by classification images. Fall Vision Meeting, San Francisco, CA, October 24–27, 2002.

Solomon, J. A., & Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature, 369*(6479), 395–397.

Tjan, B. S., Braje, W. L., Legge, G. E., & Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research, 35*(21), 3053–3069.

Watson, A. B. (1998). Multi-category classification: template models and classification images [ARVO abstract #4197]. *Investigative Ophthalmology and Visual Science, 39*(4), S912.