COMPARING HUMAN AND MACHINE DETECTION THRESHOLDS

An a contrario model for non-accidentalness.
Samy Blusseau 1, José Lezama 1,2, Rafael Grompone von Gioi 1,2, Jean-Michel Morel 1, Gregory Randall 2

1: CMLA, ENS Cachan, France 2: IIE, Universidad de la Republica, Uruguay

The mathematical theory of an a contrario detection formalizes the non-accidentalness principle [2] and attempts to predict ideal perception thresholds. Thus, it is natural to reconsider from a computational perspective, classic and new psychophysical experiments evaluating the human perception performance. To this aim, we chose the psychophysical experiments by Wagemans et al. [3] where subjects are presented with Gabor-rendered outlines of real world objects. In these experiments, orientation jitter was added to the elements with the aim of determining its effect on human object detection performance. Using the a contrario theory, the human detection thresholds can be compared rationally to the algorithmic ones. To allow a broader experimentation, we built an online web facility where users can perform object detection experiments, and compare their detection curves to the ones predicted analytically by the computational model.

Background

Contours detection

From Wagemans et al’s experiment [3], we kept the Gabor-rendering of shapes and their masking by adding orientation jitter on contours. In this first attempt to predict detection thresholds with an a contrario model, we focused on straight contours for their simplicity.

A contrario model

The non-accidentalness principle states that, among a set of potential structures, only the configurations that would rarely appear by chance are perceptually relevant. The "a contrario" model translates this principle in a mathematical language, as follows: a configuration is perceptually meaningful when its expectation in noise is less than 1. This means that in average, only one false detection would be made in a noise image.

We define an upper bound of this expectation of an event in noise, and call it "Number of False Alarms", or NFA.

Human detection

Protocole

This experiment is accessible on the web at bit.ly/aligned_gabors. During a session of the experiment the subject sees 50 images. More precisely:

- 5 training stimuli (the first 5 images)
- 30 images are randomly sampled from the database according to the following probabilities: 25 % for negative stimuli (all elements have random orientations), 35 % for positive stimuli (some elements have constrained orientations).
- A Yes/No question for each stimulus: the subject has to answer whether he sees or not a straight line; his response time is measured but no time limitation is imposed.

Stimuli Database

The database is large enough to avoid repetitions (more than 14 000 images), and was generated with GERT (v1.1) [1]. Each image contains N = 200 Gabor elements, not too close from each other.

The positive stimuli (containing a straight line) vary according to:

- the straight line’s length : from 3 to 10 aligned elements
- noise level: the added orientation jitter belongs to an interval [−θ, θ] where θ ∈ {0°, 15°, 22.5°, 30°, 45°, 60°, 75°, 90°}
- the position of the segment’s center : 25 positions covering the image’s area
- the slope of the segment, defined by the angle α ∈ [−10°, −45°, −30°, −15°, 0°, 15°, 30°, 45°, 60°, 90°] with the horizontal axis

Given a pair (a,b) of Gabor elements, we define i as the expected number of elements in the stripe of length d and width δ pixels, knowing that the average distance between two neighbours is xavg, that is:

\[ i = \frac{x_{avg}}{d + \delta} + 1 \]

Then, for a precision p ∈ [0; 1], k(p) is the actual number of elements that are in the green stripe and whose orientation is parallel to (a,b) with precision p.

On the left hand illustration, the one in the middle shows a “full” stripe in which one element is not parallel to (a,b) with precision p ; in the third one, only 3 elements are in the stripe, all with same orientation under precision p.

The binomial test \( B(p, n, p) = \sum_{i=p}^{n} \binom{n}{i} (1-p)^{n-i} p^i \) can be computed for each pair and any precision p. In the algorithms, each pair is tested with 5 precisions: p = 0.05, p = 0.1, p = 0.2, p = 0.5, p = 0.95.

For an image containing N Gabor elements, the total number of tests is

\[ NFA = \text{number of pairs} \times \text{number of tested precisions} \times \text{number of bins} \]

and for a given pair (a,b), the significance of the corresponding straight line is given by its NFA:

\[ SFEK(a,b) = NFA \times \min_{p=0.05, 0.1, 0.2, 0.5, 0.95} B(p, n, k(p)) \]

The algorithm detects the structure having the lowest NFA if it is less than 1.

Results and discussion

Examples of machine detection

Machine detection

Grouping laws

Orientation similarity and width constancy

- 5 out of 4 Gabor elements sharing orientation (a,b) with precision p ∈ [0; 1]

Width constancy: \( d_1 = d_2 = d_3 = \frac{d}{2} \)

References

