Measuring the visual salience of alignments by their non-accidentalness

Samy Blusseau, A. Carboni, Alejandro Maiche, Jean-Michel Morel, Rafael Grompone von Gioi

Vision Research xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Measuring the visual salience of alignments by their non-accidentalness

S. Blusseau [a,*], A. Carboni [b], A. Maiche [b], J.M. Morel [a], R. Grompone von Gioi [a]

[a] CMLA, ENS Cachan, France
[b] CIBPsi, Facultad de Psicología, Universidad de la República, Uruguay

ARTICLE INFO

Article history:
Received 30 January 2015
Received in revised form 24 August 2015
Accepted 26 August 2015
Available online xxxx

Keywords:
Contour integration
Good continuation
Detection
Ideal observer
Non-accidentalness
a contrario theory

ABSTRACT

Quantitative approaches are part of the understanding of contour integration and the Gestalt law of good continuation. The present study introduces a new quantitative approach based on the *a contrario* theory, which formalizes the non-accidentalness principle for good continuation. This model yields an ideal observer algorithm, able to detect non-accidental alignments in Gabor patterns. More precisely, this *parameterless* algorithm associates with each candidate percept a measure, the Number of False Alarms (NFA), quantifying its degree of masking. To evaluate the approach, we compared this ideal observer with the human attentive performance on three experiments of straight contours detection in arrays of Gabor patches. The experiments showed a strong correlation between the detectability of the target stimuli and their degree of non-accidentalness, as measured by our model. What is more, the algorithm's detection curves were very similar to the ones of human subjects. This fact seems to validate our proposed measurement method as a convenient way to predict the visibility of alignments. This framework could be generalized to other Gestalts.

1. Introduction

The question of how vision integrates a set of elements into the contour of a shape was raised by the Gestaltists. They identified several conditions that favored the emergence of such percepts (Wertheimer, 1923; Metzger, 1975; Kanizsa, 1980; Wagemans et al., 2012; Wagemans et al., 2012). Among them is the well known grouping principle (or *Gestalt*) of good continuation, that refers to the perceptual grouping of elements forming smooth curves.

This principle has been further investigated in psychophysics, with a recurring use of arrays of Gabor patches as stimuli (see Hess, Hayes, & Field, 2003; Hess & Field, 1999; Hess, May, & Dumoulin, 2014 for a review). The experiments described in Field, Hayes, and Hess (1993) opened a long standing research line dealing with various aspects of the *association field* hypothesis. For example, some studies (Ledgeway, Hess, & Geisler, 2005; Vancleef & Wagemans, 2013) compared the detection of paths formed by Gabor patches oriented parallel to the contour (*snakes*), to the detection of *ladders*, in which all the elements are orthogonal to the path, and *ropes* (for Ledgeway et al., 2005), where elements are obliquely oriented with respect to the path. In Nygård, Van Looy, and Wagemans (2009), the effect of orientation jitter, analogous to the "α" parameter of Field et al. (1993), was measured on the detection and identification of everyday objects' contours. The combination of the association field grouping effect and other potential cues, such as symmetry (Machilsen, Pauwels, & Wagemans, 2009; Sassi, Demeyer, & Wagemans, 2014), closure of contours as well as the surface enclosed in contours (Kovacs & Julesz, 1993; Machilsen & Wagemans, 2011), was also studied. As far as the detection of snakes is concerned, a result that is common to all the previously mentioned experiments is that the visual system is better at detecting *smooth* paths rather than jagged ones, and all the more so as they are formed by elements that are roughly parallel to the local tangent of the contour.

An important step towards the understanding of the mechanisms underlying these observations is the definition of a model that could *explain* and *predict* the quantitative results obtained experimentally. Can we predict, for example, as observed in Field et al. (1993), that a path composed of 12 Gabor patches embedded in an array of 244 other elements randomly oriented, is likely to be detected when the absolute angle β between two consecutive edges is less than 60°, and without orientation jitter ($\alpha = 0°$)? What about the decay of the detection performance for a fixed β when $\alpha = 15°$ and 30°? What if the path was composed of, say, only 8 elements? Computational models designed to make such predictions have already been proposed (Yen & Finkel, 1998; Ernst et al., 2012). The framework described in Yen and Finkel (1998) consists in a cortex-inspired network, in which the

* Corresponding author.
E-mail address: samy.blusseau@cmla.ens-cachan.fr (S. Blusseau).

http://dx.doi.org/10.1016/j.visres.2015.08.014
0042-6989/

interaction between cells is ruled by a mathematical formulation of the association field. In accordance with the experimental results of Field et al. (1993), the latter favors the synchronization between cells responding to close, co-circular stimulations, as well as among straight ladders configurations. The resulting algorithm contains 11 parameters that were tuned by optimizing the detection results on a few images. It is compared to the psychophysical data of Field et al. (1993); Kovacs and Julesz, 1993; Kapadia, Ito, Gilbert, and Westheimer, 1995, showing an interesting match between the perceptual results and the model's detections. Equally interesting is the Bayesian approach of Ernst et al. (2012). By defining a generative model of contours, the authors could compare human subjects to an ideal observer in their own experiment of contour detection. They used the psychophysical data they collected as a reference, and tuned the parameters of their algorithm by optimizing the correlation between the subjects and the model's responses. Although the obtained *"optimal observer"* still performed significantly better than the average of subjects, the model managed to mimic the dynamics of the subjects' responses, and to reproduce some observed phenomena, such as the early detection of longer contours. Another quantitative approach to contour detection in noise is Wilder, Feldman, and Singh (2015), which does not propose an ideal observer but a Bayesian prediction of the detectability of contours relative to their regularity.

The present paper introduces a new approach to build a non-parametric predictive model of good continuation, based on the *non-accidentalness principle*. This concept has emerged at the crossroad of human and computer vision research. Indeed, faced with the quantitative questions raised by the qualitative results found by the Gestalt school (Wertheimer, 1923; Kanizsa, 1980; Wagemans et al., 2012), vision scientists like Witkin and Tenenbaum remarked that *"the appearance of spatiotemporal coherence or regularity is so unlikely to arise by the chance interaction of independent entities that such regular structure, when observed, almost certainly denotes some underlying unified cause or process"* Witkin and Tenenbaum, 1983, p. 481. This idea that perception relies on non-accidental relationships to segment an observed scene into meaningful structures, is usually called the *non-accidentalness principle*. Although it was particularly well addressed in Witkin and Tenenbaum (1983), previous or contemporary formulations exist. It has inspired computer vision since the 1980s, when researchers relied on the so called *non-accidental features* to perform automatic scene interpretation (Stevens, 1980; Stevens, 1981; Kanade, 1981; Binford, 1981; Lowe & Binford, 1981; Witkin, 1982; Lowe et al., 1985; Lowe, 1987), and it was crucial in the success of the scale-space filtering method (Witkin, 1983) which led to actual breakthroughs in computer vision (Lowe, 1999; Lowe, 2004). What is more, the concept of non-accidental features also sparked the interest of psychophysicists, who wanted to test their relevance in perception (Wagemans, 1992; Wagemans, 1993; Van Lier, van der Helm, & Leeuwenberg, 1994; Feldman, 2007). The contribution of Feldman (2007) is actually beyond the experimental measure of the particular status of non-accidental properties in perception. It draws a parallel with their special status in the *minimal model* theory (Feldman, 1997a; Feldman, 2003). The latter theory was then featured a Bayesian structure (Feldman, 2009), in an attempt to bind together non-accidentalness, the likelihood principle and the simplicity principle at a perceptual level. A different approach of these questions is exposed in Van Lier et al. (1994); van der Helm, 2000; van der Helm, 2011, where an extensive theory formalizes the simplicity principle and shows, among many other things, its consistence with non-accidentalness, and in particular with Rock's *rejection of coincidence* (Rock, 1983).

Another mathematical formulation of non-accidentalness was introduced in Desolneux, Moisan, and Morel (2008), although the authors called *Helmholtz principle* what is more commonly called non-accidentalness. This probabilistic approach is known in computer vision and image processing as the *a contrario* theory. The expression *a contrario* (Latin for "by or from contraries") refers to the fact that the theory focuses the *probabilistic* modeling on *noise*, understood as the absence of signal. The model of the signal itself is a deterministic description of an ideal structure (in our case, a perfect alignment). When an *a contrario* detection algorithm analyzes a candidate feature, it evaluates the expected number of features with lower or equal error *relative to the ideal structure*, that could happen by chance in noise. In a way, the *a contrario* theory may remind of an aspect of the signal detection theory, as *"a computational framework that describes how to extract a signal from noise"* (Gold & Watanabe, 2010, p. 1). In signal detection theory, a detection in absence of signal is called a *false alarm*. Similarly, the expected number of accidental detections in noise, associated to a given feature by an *a contrario* algorithm, is called *Number of False Alarms* (NFA). The lower this number, the less likely it is that the detection of this feature be a false alarm. Thus, if the NFA is low enough, the feature is termed non-accidental and therefore significant. This methodology will be precisely defined and explained in more details in Section 6.

As a probabilistic formalization rooted in information theory, the *a contrario* framework, applied to the good continuation problem, may be compared to Bayesian approaches (Feldman, 1997b; Feldman, 2001; Feldman & Singh, 2005; Feldman, 2009; Wilder et al., 2015; Ernst et al., 2012). A Bayesian formulation requires prior and conditional distributions along with their parameters, for each hypothesis to be tested. In some problems, these are well known, or can be defined by natural assumptions; this leads to optimal predictions. In others, the optimality is undermined by the absence of sufficient knowledge. The *a contrario* methodology restrains its assumptions to a *deterministic* description of the ideal sought structure and a *probabilistic* background (or *a contrario*) hypothesis H_0, which models the absence of the relevant pattern by a maximal entropy distribution. The Bayesian approach is preferable when all the distributions and their parameters are well known, as it is able to take full advantage of this information. In some cases where we lack such knowledge, the *a contrario* methodology offers a good alternative, as it may avoid some heuristic assumptions. In our case, we don't have, *a priori*, perfect knowledge of what a salient alignment is for human perception. (Note that this is different from the knowledge of how the stimuli were generated for our particular experiments.) For example, Feldman (2009) discusses the problem of deciding whether a set of Gabor patches forms a smooth chain or not; in this setting, we don't know what should be the variance of a distribution on turning angles under the good continuation hypothesis. An *a contrario* approach would precisely avoid to set this parameter, since it would model smooth chains by a deterministic description of perfect good continuation, namely a zero value for turning angles. Then, without a probabilistic model for the smooth chain hypothesis, it is not possible to compute a likelihood ratio but, as described earlier, the *a contrario* theory provides a different strategy to decide if the null hypothesis should be rejected.

A whole family of unsupervised algorithms in computer vision are based on this theory. The framework has been used to detect line segment and curves (Desolneux et al., 2008; Grompone von Gioi, Jakubowicz, Morel, & Randall, 2012; Pătrăucean et al., 2012) as well as to perform shape matching and identification (Musé, Sur, Cao, & Gousseau, 2003; Cao et al., 1948), image segmentation (Cardelino et al., 2009), clustering (Tepper et al., 2011), and mirror symmetry detection (Pătrăucean, Grompone von Gioi, & Ovsjanikov, 2013).

The *a contrario* theory was confronted to human perception in Desolneux, Moisan, and Morel (2003). Although it was not explicitly presented as such, this work somehow tried to investigate the

Please cite this article in press as: Blusseau, S., et al. Measuring the visual salience of alignments by their non-accidentalness. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.08.014

role of non-accidentalness in visual perception. It already suggested that this approach might give an interpretation and quite accurate predictions of the perceptual thresholds. Its most remarkable insight is to translate several detection-affecting parameters into one unique measure of non-accidentalness, the NFA, which correlates well with the subjects detection performance. Yet the psychophysical protocols in Desolneux et al. (2003) had several limitations that also limited the strength of its conclusions.

In particular, Desolneux et al. (2003) did not compare *directly* human subjects to an algorithm. In contrast, we took inspiration from the innovative work by (Fleuret, Li, Dubout, Wampler, Yantis, & Geman, 2011) where human and machine performed the same visual categorization task, "side by side", on each stimulus. In a nutshell, the algorithm became an artificial subject, as it was also the case in Ernst et al. (2012). Similarly, our experiments compare human and machine vision in a simple perceptual task: the detection of jittered straight contours among a set of randomly oriented Gabor patches (Fig. 1). Following the definition of Geisler (2011, p. 2), "an ideal observer is a hypothetical device that performs a given task at the optimal level possible, given the available information and any specified constraints." In this sense, our *a contrario* detection algorithm is an ideal observer under the constraint that no relevant structure should be detected in noise.

The non-accidentalness principle can be interpreted as imposing minimal reliability conditions to percepts. A configuration that could have arisen by chance does not provide reliable information and should be rejected. Natural selection evolved reliable perception mechanisms which may obey a similar sound requirement, that no detection should occur in noise (Attneave, 1954). Undoubtedly, this minimal reliability condition is not the only factor involved in the salience of a particular structure. Nevertheless, it is interesting to evaluate whether this factor alone can predict the visibility of a specific Gestalt to some extent. The purpose of this paper is precisely to study to what extent the *a contrario* formulation of non-accidental alignments can account for their salience in a large set of stimuli. We shall illustrate that this theory permits to summarize into one single measure, the NFA, the many experimental parameters affecting this visibility. By addressing the case of alignments, we hope to point out the key features that make the approach applicable to other Gestalts. Indeed, given a grouping law described qualitatively, the *a contrario* framework proposes a way to evaluate quantitatively how strongly the configurations of elements must stick to this grouping law, to be distinguished from a general, accidental configuration.

In Sections 2,2,3,4,5 we present our experimental set up, consisting in three versions of a straight contours detection task, and expose the results obtained by 32 subjects. After describing our model and detailing a parameterless detection algorithm in Section 6, we characterize our set of stimuli by their level of non-accidentalness, that we measure by the NFA. This enables us to evaluate, in Sections 7 and 8, how non-accidentalness correlates with detection performance, and if an *a contrario* algorithm mimics, and therefore can predict, the subjects behavior on average. Finally, Section 9 summarizes our conclusions.

2. Methods

This section presents the techniques that are common to all the experiments described in the paper.

2.1. Stimuli

In all the following experiments, subjects were shown arrays of Gabor patches of 496×496 pixels, similar to those represented in the first row of Fig. 1. They were generated with the software GERT (v1.1) (Demeyer & Machilsen, 2011). An array contained a certain number N of patches, each of which represented a symmetrical Gabor function, characterized by its position in the image and its orientation θ. The family of Gabor functions we used is described by

$$\forall (x,y) \in \mathbb{R}^2,$$
$$G_{f,\theta,\sigma}(x,y) = \frac{1}{2}\left(1 + e^{-\frac{x^2+y^2}{2\sigma^2}} \cos\left(2\pi f (y\cos\theta - x\sin\theta)\right)\right) \quad (1)$$

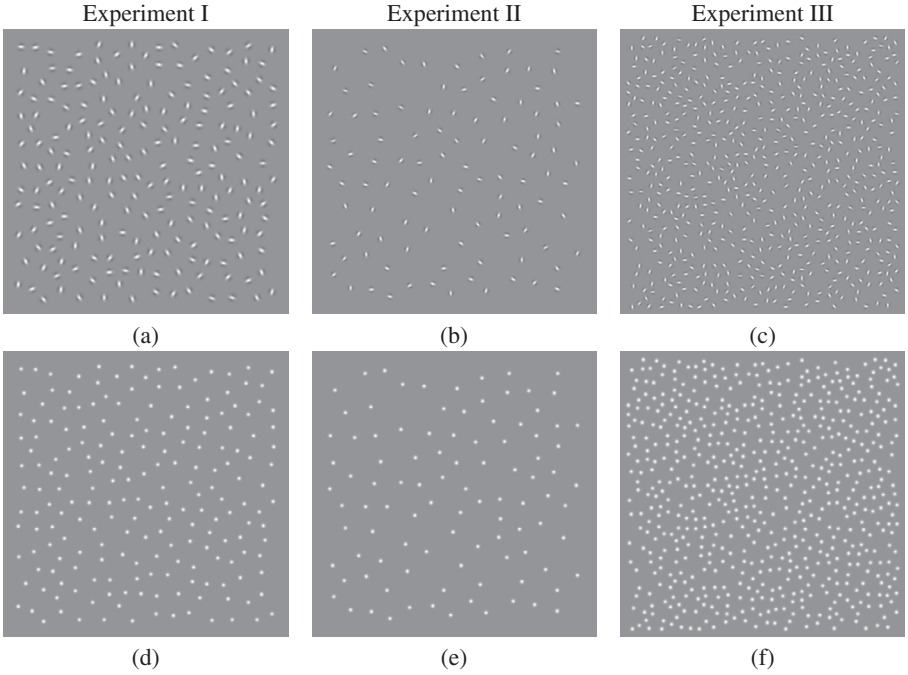


Fig. 1. First row: three arrays of N Gabor patches with long, moderately jittered straight contours; $N = 200$ in (a), $N = 100$ in (b) and $N = 600$ in (c). Images (d), (e) and (f) show dots representing the coordinates of the Gabor elements of images (a), (b) and (c), respectively.

Please cite this article in press as: Blusseau, S., et al. Measuring the visual salience of alignments by their non-accidentalness. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.08.014

f being the spatial frequency, σ the space constant set to $\sigma = \frac{1}{4f}$ and θ an angle in $[0°, 360°]$. Note that this function is centered on the origin, and can then be placed at the desired position in the image.

Each array included n aligned patches, regularly spaced by a distance r_a. Their orientations were chosen as follows: given the line's direction θ_l, the n angles were randomly and independently sampled according to a uniform distribution in $[\theta_l - \alpha, \theta_l + \alpha]$, with $\alpha \in [0°, 90°]$. This is equivalent to adding uniform noise, also called *angular jitter*, to orientations equal to θ_l. The alignment position and direction θ_l were selected randomly and uniformly in each image. In all the paper, these aligned patches will be referred as *target alignment* or *target elements*. The $N - n$ non-aligned elements, called *background elements*, were randomly placed but respected a minimal distance r_b from each other and from the aligned patches. Their orientations were randomly and independently sampled according to a uniform distribution in $[0, 180°)$. Distances r_a and r_b were set as functions of N to fulfill two requirements. First, r_b was tuned so that all N elements fit into the image and fill it homogeneously, avoiding clusters and empty regions. Second, we chose r_a larger than r_b to make the alignment almost impossible to detect from the coordinates of the elements only, that is to say from the proximity and width constancy cues only (Fig. 1(d)). More precisely, we set $r_b = 1.3 r_{hex}$ and $r_a = 2 r_{hex}$, where $r_{hex} \stackrel{\text{def}}{=} \frac{456}{\sqrt{2\sqrt{3}N}}$ is, in pixels, the maximal radius for N discs to fit in a 456×456 pixels square without intersecting (the number 456 corresponds to the image side minus two margins of 20 pixels, one at each border). Then the difficulty to detect the aligned Gabors was essentially ruled by their number n, and by the angular precision α. The larger n and the smaller α, the more conspicuous the alignment.

2.2. Procedure

This work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Informed consent was obtained for experimentation with human subjects.

Each experimental session counted 126 trials divided into three blocks of 42, with pauses after the first and the second block. In each trial, the subject was shown on a screen an array containing N Gabor patches, n of which were aligned ($n \in \{4, \dots, 9\}$) and affected by an angular jitter of fixed intensity $\alpha \in \{9°, 15°, 22.5°, 30°, 45°, 90°\}$. Four trials were conducted for each couple of conditions (n, α) with $\alpha \leqslant 45°$, and one trial per maximal jitter condition $(n, 90°)$, in a random order. The subjects knew that every stimulus contained an alignment, but did not know where, how long and how jittered it was. They were asked to click on an element they perceived as part of the alignment, and to give their best guess in case they did not see any clear alignment. We wanted the task to be *attentive*. The stimulus remained on the screen until an answer was given. The elapsed time between the presentation of the stimulus and the click is what we call the reaction time, and was measured at each trial. Subjects were advised to spend no more than 20 s per stimulus. They were free to follow this qualitative suggestion or not and no time information was provided during the sessions. The coordinates of the click were recorded as well. After the subject's click, a transition gray image was displayed for 500 ms and then the next stimulus appeared.

To help understand the task, subjects were first shown two training sequences of 5 and 10 trials, with the corresponding target alignments at the end of each sequence.

2.3. Apparatus

The monitor properties, the distance to the screen and the illumination conditions could vary from one subject to another. Indeed, various computers were used at different places along the sessions. However we can estimate the average distance between the subjects' eyes and the monitor to be about 70 cm, and the average size of a pixel to be approximately 0.02 cm × 0.02 cm, so that the 496×496 images subtented about $8° \times 8°$ of visual angle in average. The experiments were set up to work on a web browser, and they were taken by selected subjects under our supervision. They are accessible at http://bit.ly/na_alignments. Our bet was that, for the scope of our study, some variability in the viewing conditions would not have significant inter-subject effects. What differed between the experiments described hereafter, was the number N and the size of the Gabor patches.

3. Experiment I ($N = 200$)

A previous version of this experiment was presented in Blusseau, Carboni, Maiche, Morel, and Grompone von Gioi (2014).

3.1. Subjects

Twelve subjects, five women and seven men, with age between 20 and 40 years old, took the experiment (accessible at http://bit.ly/ac_alignments). All the subjects were naive to the purpose of the experiment and had normal or corrected to normal vision.

3.2. Stimuli

The stimuli of this experiment were those described in Section 2.1, with $N = 200$ Gabor patches per image, and a spatial frequency $f = 0.077$ cycles per pixel (see Fig. 1(a) for an example). Four sequences of 126 images were pre-computed, according to the categorization detailed in Section 2.2, and for each of the twelve sessions one sequence was picked randomly by the computer program. One subject was shown sequence 1, four subjects saw sequence 2, sequence 3 was seen by two subjects and sequence 4 by five subjects.

3.3. Results

Each click made by a subject was associated to the nearest Gabor element in the image, and counted as a valid detection when that element belonged to the target alignment. This count permitted to define the detection rate. Fig. 2 displays the results of Experiment I (represented by a solid line), as well as those of Experiments II and III, described in the next sections. Fig. 2(a) and (b) plot the detection rate as a function of the jitter level α and the number n of aligned elements, whereas (c) and (d) plot the reaction times as a function of the same parameters (all the trials were counted; there was no attempt to remove outliers). The error bars give 95% confidence; they are defined as $\left[\bar{x} - 2\frac{\sigma}{\sqrt{q}}, \bar{x} + 2\frac{\sigma}{\sqrt{q}}\right]$, where $\bar{x}, \sigma$ and q are respectively the mean, standard deviation and number of trials of the corresponding condition. In Table 1 are reported detection rates and reaction times achieved by each subject of Experiment I.

3.4. Discussion

Subjects 8 and 11 achieved higher rates than the rest of the subjects, but they also dedicated more time to the task. The results plotted in Fig. 2 are consistent with those obtained in previous studies on the influence of orientation jitter on contour detection (Field et al., 1993; Ledgeway et al., 2005; Nygård et al., 2009). The more jittered the alignments are, the lower the average detection performance, and the longer the subjects took in looking for them (Figs. 2(a) and (c)). It was also harder

Please cite this article in press as: Blusseau, S., et al. Measuring the visual salience of alignments by their non-accidentalness. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.08.014

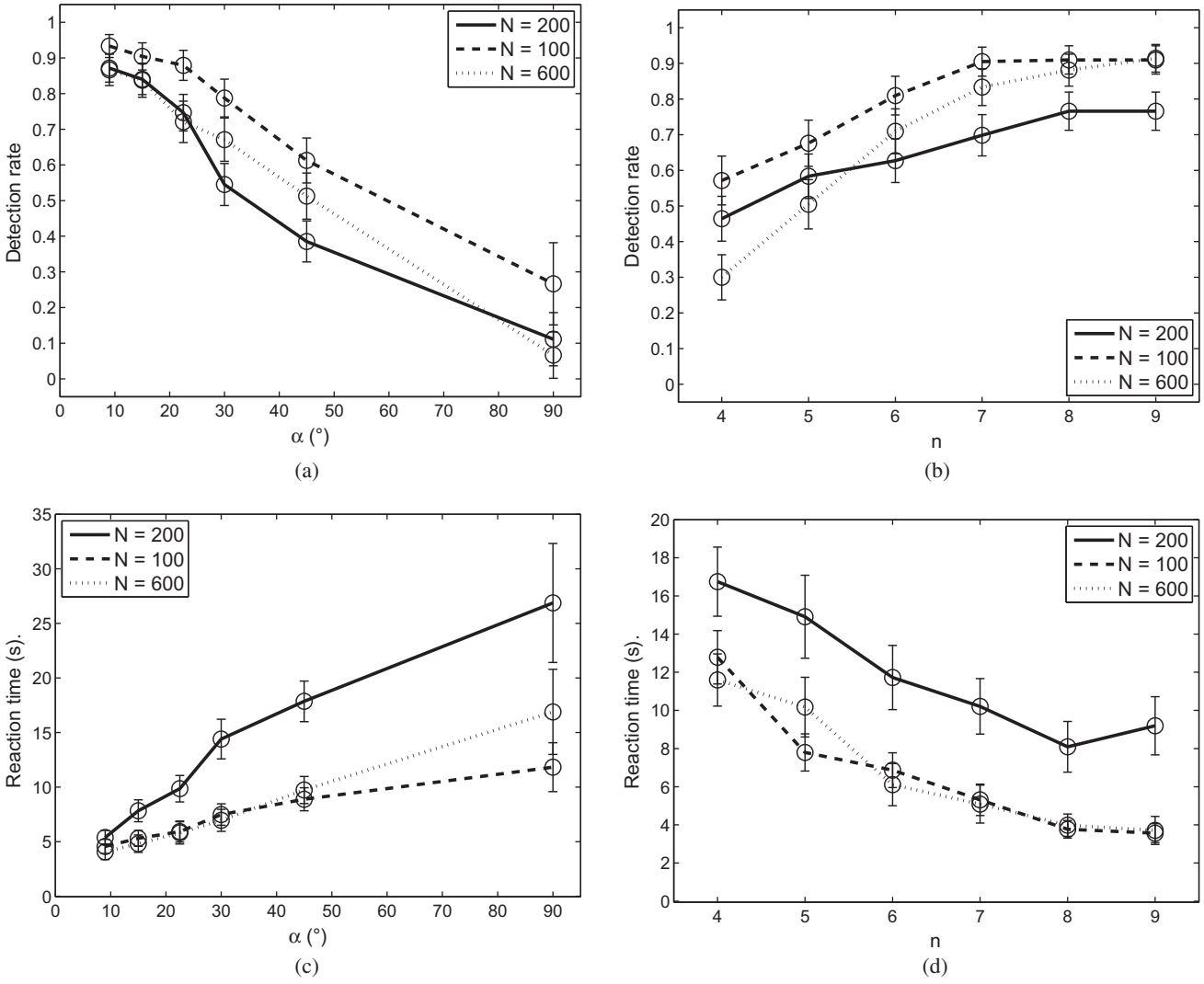


Fig. 2. Results of the Experiments I–III (see also Tables 1–3). The subjects' detection rates (top row) and reaction times (bottom row) are plotted as functions of the jitter intensity α (left hand column) and the number n of aligned elements (right hand column). The error bars give 95% confidence; they are defined as $\left[\overline{x} - 2\frac{\sigma}{\sqrt{q}}, \overline{x} + 2\frac{\sigma}{\sqrt{q}}\right]$, where $\overline{x}, \sigma$ and q are respectively the mean, standard deviation and number of trials of the corresponding condition.

Table 1
Individual detection rates and average reaction time per trial, for Experiment I.

Subject	1	2	3	4	5	6	7	8	9	10	11	12	all
Det. rate	0.63	0.65	0.65	0.56	0.64	0.52	0.67	0.80	0.61	0.64	0.80	0.63	0.65
Aver. r. t. (s)	7.7	9.5	8.4	12.9	14.1	7.5	9.3	20.9	7.9	8.2	21.0	14.4	11.8

to detect short alignments than long ones (Fig. 2(b) and (d)). It is worth noting that subjects still achieved about 10% detection rate in the maximal jitter condition (8 detections out of 72 trials), while chance level is $\frac{1}{6}\sum_{k=4}^{9}\frac{k}{N} = 3.3\%$ for $N = 200$. Three factors could explain these 8 clicks on one of the target elements in the maximal jitter conditions. The first one is pure chance, when the subject did not see any alignment and clicked on a random element, that turned out to belong to the target. This kind of event seems to have occurred in one of the eight mentioned cases. Indeed, the corresponding alignment appears to be impossible to detect, and the clicked point does not seem to belong to any other relevant structure (see Fig. 3, first row). The second factor is the overlapping of the target and a random structure that looks like a jittered alignment. From the observation of the concerned stimuli and the subjects' answers, it seems that this scenario may explain 4 out of the 8 detections in noise (see two of these in Fig. 3, second row). The three remaining clicks seem to be actual detections of a part of the target nine-elements-alignment. Indeed, they occurred on the same stimulus, showing an agreement of three subjects on the same linear structure (Fig. 3, third row). The latter case is the most interesting, since it shows that random orientations are not always enough to mask the alignment.

The *a contrario* detection theory allows the definition of parameterless algorithms that automatically adapt to changes in the size of the data. In order to demonstrate this property and compare it with perception, it was natural to set up experiments with stimuli containing less (see Section 4) and and more (see Section 5) patches than in Experiment I. This is the motivation for Experiments II and III.

Please cite this article in press as: Blusseau, S., et al. Measuring the visual salience of alignments by their non-accidentalness. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.08.014

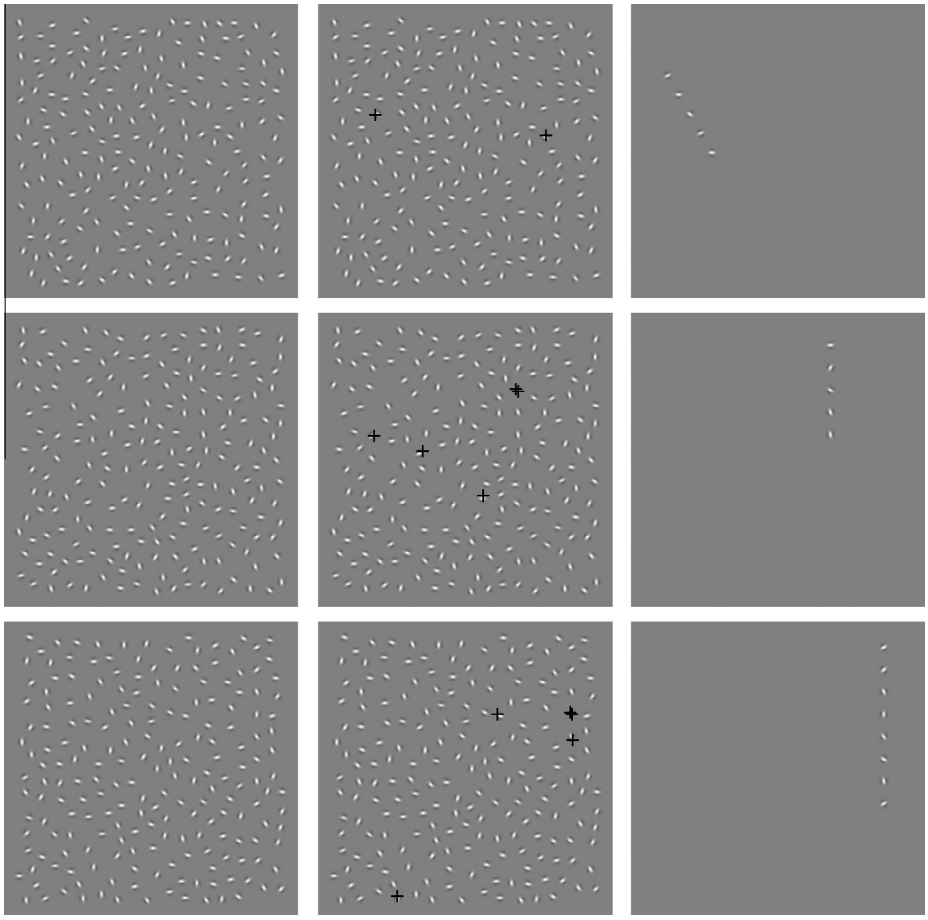


Fig. 3. Three examples of detections in the maximal jitter condition in Experiment I. Each row represents an example. The left hand column shows the original stimuli, the central column displays the subjects' clicked points, and in the right-hand column are represented the target alignments. First row: one subject seems to have clicked completely randomly on the target alignment. Second row: the two subjects who clicked on the target probably saw a different random structure. Third row: three subjects seem to agree on the detection of a subset of the target, composed of nine elements.

4. Experiment II ($N = 100$)

4.1. Subjects

Ten subjects participated in Experiment II (accessible at http://bit.ly/ac_alignments_100elts). They were three women and seven men, with age between 20 and 60 years old, and had normal or corrected to normal vision. Contrary to Experiment I, here one of the subjects is co-author of the present paper (subject 9 in Table 2); the others were naive to the purpose of the experiment.

4.2. Stimuli

The stimuli for this experiment differed from Experiment I by two parameters: first, they counted $N = 100$ Gabor patches each; second, the spatial frequency was set to $f = 0.12$ cycles per pixel, that is to say approximately 1.5 times the frequency used to build the stimuli of Experiment I. Fig. 1(a) is an example of this kind of stimulus. In images with half as many elements but with same size as in Experiment I, the patches are more distant from each other. By increasing the spatial frequency f and making Gabor patches look thinner than in Experiment I, we wanted to accentuate the sensation of empty spaces between neighboring elements. Only one pre-computed sequence of 126 images was used for this experiment, because we wanted to compare all the ten subjects on the very same stimuli.

4.3. Results and discussion

In Experiment II the subjects obtained globally better detection rates than those of Experiment I (see Tables 1 and 2 and Fig. 2). This is not surprising since in this last experiment, the stimuli were less crowded than in the previous one. As a consequence, the set of stimuli presented less difficulty on average. Here we observe almost 30% for the detection rate in the maximal jitter condition (detection in 16 out of 60 trials), the chance rate being approximately 6.5%. Although the hypothesis formulated in Section 3.4 still hold to account for this observation, it seems that in most of the 16 cases the subjects *did* detect the target alignment, thanks to other cues such as the perfect alignment of the elements' coordinates, and their perfectly regular spacing. Since the proportion of aligned elements relative to the total number of elements was

Table 2
Individual detection rates and average reaction time per trial, for Experiment II.

Subject	1	2	3	4	5	6	7	8	9	10	all
Det. rate	0.75	0.83	0.79	0.78	0.89	0.79	0.66	0.84	0.87	0.77	0.80
Aver. r. t. (s)	5.2	5.0	13.1	5.5	7.0	7.9	6.8	5.9	2.6	7.8	6.7

Please cite this article in press as: Blusseau, S., et al. Measuring the visual salience of alignments by their non-accidentalness. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.08.014

Table 3
Individual detection rates and average reaction time per trial, for Experiment III.

Subject	1	2	3	4	5	6	7	8	9	10	all
Det. rate	0.70	0.68	0.62	0.67	0.72	0.79	0.68	0.71	0.67	0.66	0.69
Aver. r. t. (s)	7.8	6.1	4.4	5.1	3.7	11.4	7.9	7.5	4.8	9.2	6.8

higher than in Experiment I, it is not surprising that the masking was sometimes less efficient here.

5. Experiment III ($N = 600$)

5.1. Subjects

As in the previous experiment, we had ten participants in Experiment III (accessible at http://bit.ly/ac_alignments_600elts), three women and seven men. They were between 20 and 60 years old, and had normal or corrected to normal vision. Except for one of them, who is co-author of the present paper (the same person as in Experiment II and subject 5 in Table 3), the subjects were naive to the purpose of the experiment.

5.2. Stimuli

The stimuli were of the same kind as the one of Fig. 1(c). They counted $N = 600$ Gabor patches each and the spatial frequency was set to $f = 0.12$ cycles per pixel, like in Experiment II. This time, the spatial frequency was chosen to avoid overlapping between neighbors. Only one pre-computed sequence of 126 images was used for this experiment, like in Experiment II.

5.3. Results and discussion

Like in Sections 3.4 and 4.3, we shall discuss the detection performance of almost 7% for the maximal jitter condition (4 out of 60 trials), above chance level (approximately 1%). Here the masking effect seems to have been efficient in 5 out of 6 stimuli with maximally jittered alignments. Indeed, the observed performance is mainly due to one stimulus, in which three subjects seem to have detected a subset of the nine elements forming the target alignment. In the remaining case it is more difficult to guess if the subject perceived the target alignment or another structure overlapping with it.

Since this experiment is the one with the most crowded stimuli, we would expect the detection rates to be lower than in all the previously presented experiments. This holds for the comparison between Experiments II and III, but subjects of Experiment I, in which there were $N = 200$ Gabor patches per stimulus, achieved a generally lower performance than those of Experiment III (see Table 1, 3 and Fig. 2). This could be related to the difference of spatial frequency defining the Gabor functions. In Experiment I, this frequency was lower and thus the patches' central blobs looked thicker than in Experiments II and III. This may have created a more efficient crowding effect, explaining lower detection rates in the former experiments, despite a smaller number of patches per image compared to Experiment III.

6. Model and algorithm

We propose to introduce briefly the intuitive ideas and formal concepts behind the *a contrario* approach. For further details, we refer the reader to Desolneux et al. (2008).

An *a contrario* method requires two models. On the one hand, a geometric model, which is deterministic and, on the other hand, a probabilistic model.

The geometric model defines an ideal structure x^* along with a function $d_{x^*}(\cdot)$ measuring a deviation from it. For example, in our case, the ideal structure is a set of aligned Gabor patches with perfectly aligned orientations. The measure of a deviation from this ideal configuration, is detailed in Section 6.1 but we can already provide a hint of it: in Fig. 5(e) for example, the smaller the angles α_i, the closer the depicted chain to a perfect alignment.

The probabilistic model, also called *a contrario* or *background* model, is the statistical hypothesis H_0 of the absence of relevant structure – the so called *"null hypothesis"*. It represents the most general assumption on the data. Consistently with Attneave's principle stating that we do not perceive any structure in white noise (Attneave, 1954), an *a contrario* model generally gives a maximal entropy to the position of the building blocks of the percept. For example if these building blocks are oriented, their orientations must be random, independent, and uniformly distributed for each block. By Attneave's principle the emergence of a percept in a realization of such a model should be unlikely, and in any case purely accidental. Therefore this background model is used to test the significance of an observation, as follows. Let X be a random variable consistent with H_0. Given an observed feature x with a deviation $d_{x^*}(x)$ from the ideal structure, its relevance is measured by the probability $\mathbb{P}_{H_0}(d_{x^*}(X) \leqslant d_{x^*}(x))$ to be at least as close to x^* under H_0. Small deviations yield small probabilities, and are thus rare events in the background model.

The purpose of the *a contrario* framework is precisely to help decide when an observed deviation is small enough to be considered non-accidental. The occurrence of an event does not only depend on its probability. It also depends on the number of observations. For example, at the roulette game, the odds of a given number n between 0 and 36 is $1/37 \approx 0.027$. With a single bet on that given n, it would be quite lucky to win. On the contrary, if a player keeps betting 1000 times on that same number, it would be surprising not to win at least once. This intuition finds a formalization in the Bonferroni correction method to test statistical significance of p-values, as it is well explained in Mumford and Desolneux (2010, pp. 114–118).

The *a contrario* methodology uses the same kind of correction term: to evaluate the non-accidentalness of an event, one should take into account the whole set of observations that were necessary to come across that particular event. For example, when looking for alignments in an array of Gabor patches, we define *a priori* a family of sets of patches, that will be tested as candidates to be alignments. This is what is usually called the *family of tests*, noted T, and N_T denotes the number of tests in T. Then, given a tested candidate x that yielded the deviation $d_{x^*}(x)$, denote by $X_1, X_2, \ldots, X_{N_T}$, the random issues to the N_T tests, under the background model H_0. The question is now how common or, on the contrary, how surprising it would be to observe a deviation smaller than $d_{x^*}(x)$ among these N_T tests. An answer is to compute the expected number of variables X_i that verify $d_{x^*}(X_i) \leqslant d_{x^*}(x)$. This quantity is called *Number of False Alarms* associated to x, noted $\mathrm{NFA}(x)$. Since all the X_is follow H_0, we get

$$\mathrm{NFA}(x) \stackrel{\text{def}}{=} \mathbb{E}\left[\sum_{i\in T} \mathbb{1}_{d_{x^*}(X_i)\leqslant d_{x^*}(x)}\right] = \sum_{i\in T} \mathbb{P}[d_{x^*}(X_i) \leqslant d_{x^*}(x)] = N_T \times \mathbb{P}[d_{x^*}(X) \leqslant d_{x^*}(x)] \quad (2)$$

Please cite this article in press as: Blusseau, S., et al. Measuring the visual salience of alignments by their non-accidentalness. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.08.014

where X also follows H_0. NFA(x) represents the number of deviations to the ideal model smaller than $d_{x^*}(x)$ that are expected to happen by accident in a set of N_T random tests. Since we are looking for a non-accidental small deviation, accidental ones are considered as "false alarms".

By definition, if NFA(x) $\geqslant 1$, we expect to observe at least one deviation smaller than, or equal to $d_{x^*}(x)$, in random data consistent with H_0. In other words, it would not be surprising that such an event happen by accident, and the observation of x should not lead to reject H_0. On the contrary, NFA(x) < 1 means that a deviation smaller than $d_{x^*}(x)$ is rather unexpected under H_0, and may indicate the presence of a relevant pattern. Thus, small NFAs characterize non-accidentalness and large ones accidental features, with a transition zone around NFA(x) $= 1$.

In the following subsections, we detail the geometric and *a contrario* models for our study.

6.1. The geometric model

For a given a tuple $(g_1, \dots, g_n)$ of n Gabor patches, we consider the variables $\alpha_1, \alpha_{2L}, \alpha_{2R}, \dots, \alpha_{(n-1)L}, \alpha_{(n-1)R}, \alpha_n$, as illustrated in Fig. 5(e). Variable α_{iL} is the absolute angle between the orientation of Gabor patch i and the line joining it to the patch $i-1$, while α_{iR} is the same thing changing $i-1$ for $i+1$. Since the first and last patches in the tuple have no previous and next elements respectively, we simply note $\alpha_1 \stackrel{\text{def}}{=} \alpha_{1R}$ and $\alpha_n \stackrel{\text{def}}{=} \alpha_{nL}$. Then we define $\omega_1 \stackrel{\text{def}}{=} \alpha_1, \omega_n \stackrel{\text{def}}{=} \alpha_n$ and for $i = 2, \dots, n-1, \omega_i \stackrel{\text{def}}{=} \max(\alpha_{iL}, \alpha_{iR})$; see Fig. 5(f).

The ideal alignment that will be our reference structure in the present study, is such a tuple of Gabor patches for which all angles ω_is are equal to 0°. In words, it is a set of aligned patches whose orientations are the same as their line's orientation.

Then we measure the deviation of a general tuple from an ideal alignment by its maximum angle $\omega^* \stackrel{\text{def}}{=} \max\{\omega_1, \dots, \omega_n\}$.

6.2. The a contrario model

Now we need to define the *a contrario* model for our stimuli. Fig. 4 shows three arrays of Gabor patches with *only background elements* and thus no significant alignment. Formally, we model an array of N Gabor patches by a set $\mathbf{g} = \{(x_i, \theta_i)\}_{i=1\dots N}$, where $x_i \in [0,1]^2$ represents the coordinates of patch number i, and $\theta_i \in \mathbb{R}$ its orientation (in this section the variables representing angles will be expressed in radians). We note $\mathbf{x} \stackrel{\text{def}}{=} \{x_1, \dots, x_N\}$ and, for any three points $x, y, z \in \mathbf{x}$, ϕ_{xyz} is the angle between vectors $\vec{xy}$ and $\vec{yz}$. Finally, the set of the 6 nearest neighbors of a point $x \in \mathbf{x}$ is noted $\mathcal{N}_6(x)$ (see Fig. 5(a) and (b)). Then we define an *a contrario* array of Gabor patches as a random set $\mathbf{G} = \{(X_i, \Theta_i)\}_{i=1\dots N}$ verifying two properties:

1. The random variables $\Theta_1, \dots, \Theta_N$ are independent and uniformly distributed in $[0, 2\pi)$.
2. For any $X \in \mathbf{X}$ and $Y \in \mathcal{N}_6(X)$ the angle $\Phi^*_{XY} \stackrel{\text{def}}{=} \min_{Z \in \mathcal{N}_6(Y)} |\Phi_{XYZ}|$ is uniformly distributed in $[0, \frac{\pi}{6})$ (see Fig. 5(d)), and is independent from $\Phi^*_{X'Y'}$ for any $(X', Y') \neq (X, Y)$.

Let's clarify the relation between this definition and the background stimuli of Fig. 4. Property 1 corresponds exactly to the rule we used to define the background elements' orientations in our arrays of patches (see Section 2.1). The relevance of Property 2 needs more justification. As explained in Section 2.1, we built each stimulus so that N elements fit in it, that two elements were not too close to each other (being distant of at least r_b), and that there were no empty regions in the image. These requirements are somehow a converse to those of the well known problem consisting in filling a region with as many spheres as possible, given their radius r. Indeed, in our case we know the number of discs to fit into the square image, and we want to set their radius, noted r_b, in order to get the most homogeneous layout. The most compact way to fit discs in a given region is to lay them on a hexagonal lattice. That is why we set $r_b = 1.3 r_{hex}$ (see Section 2.1), the value $2r_{hex}$ corresponding to placing the elements on an exact lattice, while we wanted to allow some randomness in their coordinates. The resulting stimuli look like hexagonal lattices affected by some noise, in which Property 2 approximately holds (see Fig. 5(a)).

6.3. Description of the algorithm

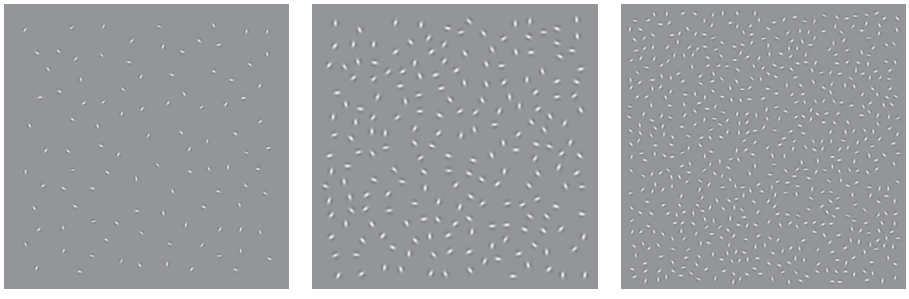
Input: The input of the detection algorithm is a set $\mathbf{g} = \{(x_i, \theta_i)\}_{i=1\dots N}$, that models a Gabor array composed of N patches (see Section 6.2).

Step 1: List the N distances of each point x_i to its nearest neighbor, and define d_{avg} as the mean value of this list.

Step 2: Define an oriented graph $\gamma = (\mathbf{x}, \mathbf{e})$, $\mathbf{x}$ being the set of vertices of γ, and $\mathbf{e}$ the set of edges, defined as follows: $(x_i, x_j) \in \mathbf{e}$ if and only if x_j is one of the 6 nearest neighbors of x_i and $d(x_i, x_j) \leqslant 2d_{avg}$, where $d(x_i, x_j)$ is the Euclidean distance between the two points. We will denote by $\mathcal{N}_\gamma(x)$ the set of neighbors of x in γ; see Fig. 5(a).

Step 3: Set $\mathcal{C} = \varnothing$. Then for each point $x \in \mathbf{X}$, and for each neighbor $y \in \mathcal{N}_\gamma(x)$, initialize a chain $c = (x, y)$, and add c to $\mathcal{C}$. This initialization step is illustrated in Fig. 5(a) and (b).

Step 4: Expand each started chain c trying to keep it as rectilinear as possible. More precisely, denote by x and y the penultimate and last points of c. Set $z^* = \arg\min_{z \in \mathcal{N}_\gamma(y)} |\phi_{xyz}|$ the neighbor of y that minimizes $|\phi_{xyz}|$. Following the notations of Section 6.2, $|\phi_{xyz^*}| = \phi^*_{xy}$. If z^* is not already in c then add z^* at the end of c, add c to $\mathcal{C}$ and carry on expanding c as long as it contains less than $\sqrt{N}$ points. By this process, represented in Fig. 5(b–d), we build all



(a) (b) (c)

Fig. 4. Three arrays of Gabor patches containing only background elements and illustrating the *a contrario* model.

Please cite this article in press as: Blusseau, S., et al. Measuring the visual salience of alignments by their non-accidentalness. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.08.014

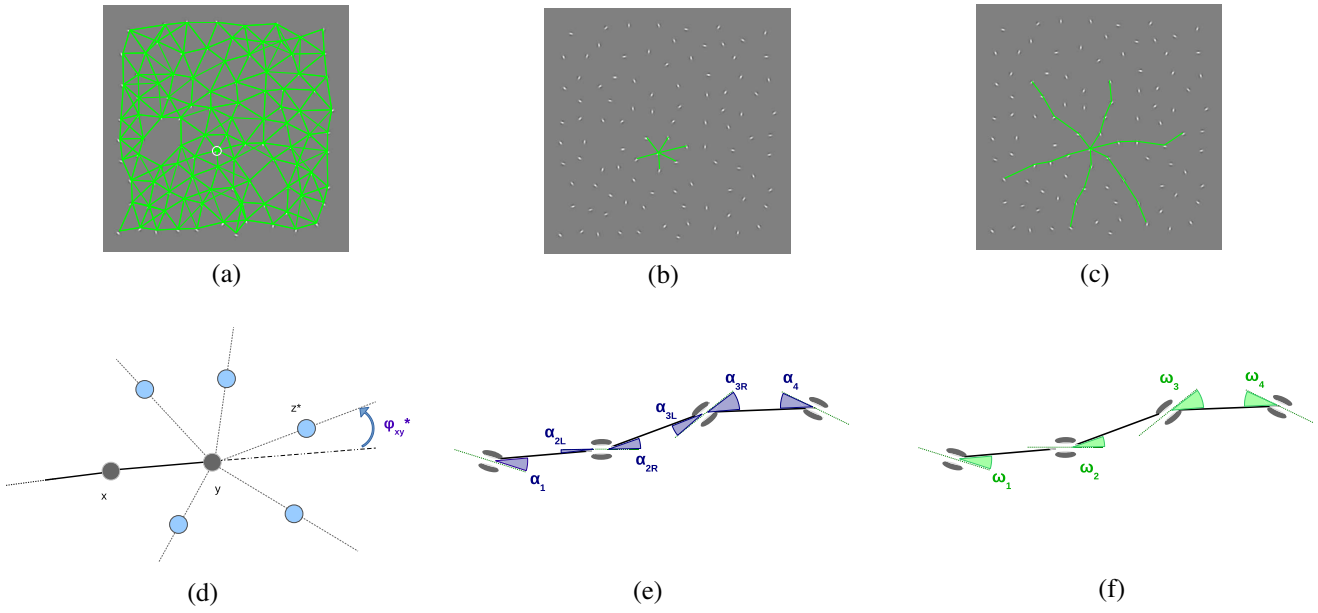


Fig. 5. (a) On this array of 100 Gabor patches, we represent the associated graph γ (the orientations of the edges were omitted). In this graph, most of the points are linked to their 6 nearest neighbors, except when a neighbor is too remote, as it often happens for elements that are close to the image border. For a point $x \in \mathbf{x}$ (surrounded in white in (a)), a chain (x, y) is started for each $y \in \mathcal{N}_\gamma(x)$ in (b), and expanded in (c) into the most rectilinear possible chain. Picture (c) shows the result after 3 iterations of the expansion process. (d) Given two neighbors x and y, we look for the neighbor z of y that minimizes the absolute angle between $\overline{xy}$ and $\overline{yz}$. (e) and (f) The variables the algorithm measures when analyzing a chain containing $n = 4$ points.

the chains to be tested as candidate alignments. Then the number of tests is approximately

$$N_T \stackrel{\text{def}}{=} 6 \times N \times \sqrt{N}. \tag{3}$$

N_T is actually an overestimation of the number of tests, since not all nodes in γ have six neighbors.

Step 5: Compute the Number of False Alarms (NFA) of each chain of $\mathcal{C}$ containing at least three points, as follows. For a given chain $c = (z_1, \ldots, z_n)$ with $n \geqslant 3$, consider the variables $\omega_1, \ldots, \omega_n$, as defined in Section 6.1 and illustrated in Fig. 5(e) and (f). Under the assumption that $\mathbf{g}$ is a realization of an *a contrario* array of Gabor patches, the orientations $\omega_1, \omega_2, \ldots, \omega_n$ are samples of n independent random variables $\Omega_1, \Omega_2, \ldots, \Omega_n$. Ω_1 and Ω_n are uniformly distributed in $[0, \frac{\pi}{2}]$, and a short development shows that $\Omega_2, \ldots, \Omega_{n-1}$ have a cumulative distribution function $F(\omega) = \mathbb{P}(\Omega_2 \leqslant \omega) = \ldots = \mathbb{P}(\Omega_{n-1} \leqslant \omega)$, defined by

$$F(\omega) = \begin{cases} \frac{12\omega^2}{\pi^2} & \text{if } 0 \leqslant \omega < \frac{\pi}{12} \\ \frac{12}{\pi^2}\left(\frac{\pi}{6}\omega - \frac{\pi^2}{12^2}\right) & \text{if } \frac{\pi}{12} \leqslant \omega < \frac{5\pi}{12} \\ \frac{12}{\pi^2}\left(\omega^2 - \frac{2\pi\omega}{3} + \frac{\pi^2}{6}\right) & \text{if } \frac{5\pi}{12} \leqslant \omega \leqslant \frac{\pi}{2}. \end{cases} \tag{4}$$

Note that for a random patch Z_i in a chain, the probability law of $\max(\alpha_{iL}, \alpha_{iR})$ depends on the angle $\Phi^*_{Z_{i-1}Z_i}$. The cumulative distribution function F is obtained by integrating over all possible values for $\Phi^*_{Z_{i-1}Z_i}$, according to the law hypothesized in the *a contrario* model. Noting $\omega^* \stackrel{\text{def}}{=} \max\{\omega_1, \ldots, \omega_n\}$ and $\Omega^* \stackrel{\text{def}}{=} \max\{\Omega_1, \ldots, \Omega_n\}$, the probability that all Ω_i be less than the observed ω^* is

$$\mathbb{P}(\Omega^* \leqslant \omega^*) = \left(\frac{\pi\omega^*}{2}\right)^2 \times F(\omega^*)^{n-2}. \tag{5}$$

As usual in the *a contrario* theory, we get a natural definition of the NFA of chain c:

$$\begin{aligned} \text{NFA}(c) &\stackrel{\text{def}}{=} N_T \times \mathbb{P}(\Omega^* \leqslant \omega^*) \\ &= N_T \times \left(\tfrac{\pi\omega^*}{2}\right)^2 \times F(\omega^*)^{n-2}. \end{aligned} \tag{6}$$

Output: The algorithm returns the lowest NFA and the chain that achieves it.

6.4. Property of the algorithm

As explained at the beginning of Section 6, a small NFA characterizes an unexpected event, while common events have a large NFA. The NFA as defined in the previous section is a positive real number, upper bounded by N_T. The longer the chain and the smaller the angles ω_i, the smaller the NFA. Consequently, the algorithm is built to detect chains with little direction change and Gabor patches roughly locally tangent to their chain. The following proposition justifies the interpretation of the NFA as a measure of non-accidentalness.

Proposition 1. Let $\varepsilon > 0$, and $\mathbf{G} = \{(X_i, \Theta_i)\}_{i=1\ldots N}$ an *a contrario* array of N Gabor patches. Then the expected number of chains c such that $\text{NFA}(c) < \varepsilon$, is less than ε.

The proof of Proposition 1 is standard (Desolneux et al., 2008). One of its implications is that there is on average less than one chain with NFA lower than 1 in an *a contrario* stimulus. The bottom row of Fig. 6 shows that no such meaningful structure was found in some stimuli containing only background elements. Since our stimuli are close to the *a contrario* model, a target alignment with NFA < 1 is likely to be detected by the algorithm because it is unlikely that other chains in the background be assigned a lower NFA. Conversely, this is not guaranteed anymore for a target alignment with NFA $\geqslant 1$.

7. Subjects compared to the algorithm

In this section we compare the subjects' detection performance to the algorithm described in Section 6.3. The question is whether a rigorous mathematical model of non-accidental alignments could match the average human detections, and there is no better way to confront the theory to the subjects than by building an artificial observer on that theory.

(a) NFA = 3.10^{-9} (b) NFA = 5.10^{-4} (c) NFA = 9.10^{-6}

(d) NFA = 11.4 (e) NFA = 4.9 (f) NFA = 42.8

Fig. 6. Detections by the algorithm in the arrays of Figs. 1 (first row) and 4 (second row). The detected chains in the second row do not look like "good" alignments, and are actually difficult to see. This is consistent with the values of the NFA that are all larger than 1 and thus not considered as significant.

Recall that each click made by a subject is associated to the nearest Gabor element in the image, and counted as a valid detection when that element belongs to the target. The same is done for the algorithm, by selecting the most central element among the returned ones (that is to say, the closest to the barycenter of the returned elements). This count permits to define the detection rate for the algorithm as well.

In Fig. 7, each of the first three rows compares the subjects of one experiment to the algorithm. The fourth row presents the same data averaged over all three experiments. The plots of columns A and B show the detection rate as a function of the jitter level and the number of aligned elements respectively, whereas columns C and D display respectively the detection rate and the reaction time as functions of $\log_{10}(\mathrm{NFA})$ of the target alignment. In these two latter sets of plots, we grouped the data into ten bins with equal number of trials. We did not defined a reaction time for the algorithm, thus only the subjects' curves appear in column D.

Tables 4–6 show the results of a statistical analysis of the agreement between the subjects and the algorithm's behaviors, as well as among subjects. To compare the answers of two observers (whether two subjects or a subject and the algorithm), we adapted an idea exposed in Ernst et al. (2012). Suppose two observers i and j were presented with the *same* set of n_c stimuli, and achieved respectively c_i and c_j correct detections in it. We want to measure the similarity in the distributions of their correct answers, beyond their detection rates. For example, imagine both observers detected $c_i = c_j = 5$ alignments out of $n_c = 10$ trial. They may have answered in exact opposite ways, if each observer detected in the trials where the other failed. In that case, we would like to point out how different these behaviors are, despite the identical detection rates. Conversely, if $c_i = 7$, $c_j = 4$ with observer i having answered correctly in the 4 trials where observer j succeeded as well, then observer i did better than j in 3 out of the 10 trials, but both observers agreed on the other 7 trials. Thus they show quite similar behaviors, no matter the difference in their detection rates. Formally, let $s_{i,1},\dots,s_{i,n_c}$ and $s_{j,1},\dots,s_{j,n_c}$ be the binaries answers (1 for a detection, 0 for a missed trial) given by observers i and j to the same n_c stimuli. Then $k = \sum_{l=1}^{n_c} s_{i,l}s_{j,l} + (1-s_{i,l})(1-s_{j,l})$ denotes the number of trials in which the observers agreed, i.e. both detected or both failed. We compare the observed number k to its random counterpart $K = \sum_{l=1}^{n_c} S_{i,l}S_{j,l} + (1-S_{i,l})(1-S_{j,l})$, assuming that the c_i and c_j detections are distributed randomly and independently (formally, the random vectors $(S_{i,1},\dots,S_{i,n_c})$ and $(S_{j,1},\dots,S_{j,n_c})$ are independent and uniformly distributed over $\{v \in \{0,1\}^{n_c}, \sum_{l=1}^{n_c} v_l = c_i\}$ and $\{v \in \{0,1\}^{n_c}, \sum_{l=1}^{n_c} v_l = c_j\}$, respectively). Under this basic assumption, we get the following law for K:

$$\mathbb{P}(K=k|n_c,c_i,c_j) = \begin{cases} \binom{c_i}{\frac{1}{2}(k-(n_c-\sigma_{ij}))}\binom{n_c-c_i}{\frac{1}{2}(n_c-\delta_{ij}-k)} \Big/ \binom{n_c}{c_j} & \text{if } 0 \leqslant k-|n_c-\sigma_{ij}| = 0 \bmod (2) \text{ and } k \leqslant n_c-|\delta_{ij}| \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where $\sigma_{ij} = c_i + c_j$ and $\delta_{ij} = c_i - c_j$. We denote by $p_k = \mathbb{P}(K \geqslant k|n_c,c_i,c_j) = \sum_{l=k}^{n_c} \mathbb{P}(K=l|n_c,c_i,c_j)$ the probability of observing at least k identical answers in the responses of the two observers. The number p_k is a p-value measuring if the observed k is significantly greater than expected under the basic assumption and, consequently, if there is significant agreement. In Tables 4–6 we report, for each subject, the values (k, p_k) resulting from the comparison with the algorithm on the 126 trials seen by the

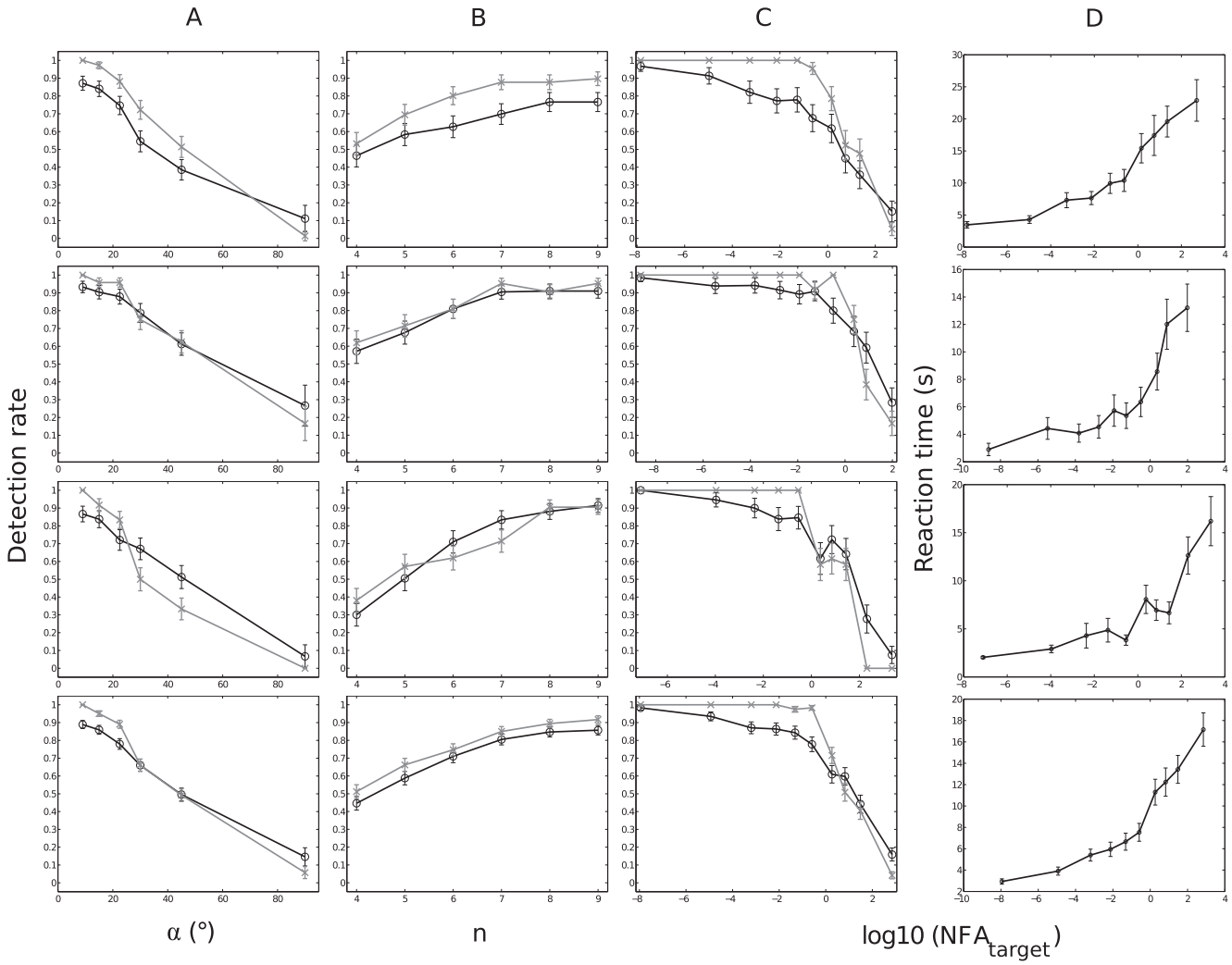


Fig. 7. Comparison of the subjects of Experiments I–III to the *a contrario* detection algorithm. The solid black lines represent the subjects, whereas the gray solid lines represent the algorithm's results. The first, second and third row display the results of experiments I, II and III respectively, whereas the fourth row shows the same data averaged over all three experiments. The detection rates are plotted as functions of the jitter intensity α (column A), the number n of aligned elements (B) and the $\log_{10}$ of the target's NFA (C). The latter is also the x variable in column D's plots, the y axis representing the reaction time, in seconds.

subject. The same figures are given for each pair of subjects who saw the same sequence of images.

8. Discussion

The first three columns of Fig. 7 exhibit a strong similarity of shape and values between the subjects' detection curves and the algorithm's ones. The match is obviously not perfect, especially in the comparison with the subjects of Experiment I, in which the algorithm performs generally better than the average of the subjects. The four plots of column A show that the algorithm tends to be more accurate than the average of the subjects on moderately jittered stimuli, whereas subjects are generally better than the algorithm at detecting more jittered alignments, especially in the maximal jitter condition. Column C provides an additional interpretation of the differences between the subjects and the algorithm. In these graphics all the stimuli are ordered by the $\log_{10}(\mathrm{NFA})$ of their target alignment. They show that the algorithm achieves better detection rates in stimuli with lower NFA (non-accidental alignments), and gets worse than the average of the subjects beyond a certain threshold, always larger than 1 (in the plot, $\log_{10}(\mathrm{NFA}) = 0$).

A possible explanation of the discrepancies between the average of the subjects and the algorithm is the following. On the one hand, the algorithm may never miss the alignment that is most distinguishable from noise in a stimulus, because of the exhaustive search it carries out. This is what we expect from an ideal observer under the specified constraints. Subjects are less exhaustive and may lack attention sometimes (recall that they had only a minimal training). Besides, the algorithm takes as input the exact coordinates and orientations of the patches, whereas the subjects may have a lower visual accuracy. On the other hand, although we built the stimuli so as to avoid the detection of the alignments in absence of the orientation cue (see Section 2.1), it seems that in some cases the background elements could not completely mask the perfect rectilinearity and the regular spacing of the target elements. For example, with a little effort, one can find the hidden alignment by looking only at Fig. 1(e). According to the observed performance in the maximal jitter condition (Sections 3–5), and the impressions that some of the subjects reported, we may infer that the regularity information was exploited by the subjects in some cases. On the contrary, the *a contrario* model described in Section 6.2 does not take into account this information. When the subjects are using information not available to the ideal observer, it is not surprising that they get better detection rates.

Please cite this article in press as: Blusseau, S., et al. Measuring the visual salience of alignments by their non-accidentalness. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.08.014

Table 4
The values (k, p_k) are reported for Experiment I. The variable k is the number of identical responses given by two observers in a set of 126 trials, and the p-value p_k measures how significant it is to observe at least k identical responses. Recall that in Experiment I not all the subjects saw the same sequences of stimuli. Therefore, the cells are filled only for pairs of observers presented with the same stimuli.

Subject	1	2	3	4	5	6	7	8	9	10	11	12	Algorithm
1	–	–	–	–	–	–	–	–	–	–	(105, 1.2e−12)	–	(95, 3.5e−06)
2	–	–	–	–	(99, 2.3e−08)	(97, 1.3e−09)	(94, 6.8e−06)	–	–	–	–	(102, 4.9e−10)	(101, 1.8e−08)
3	–	–	–	–	–	–	–	(105, 6.3e−12)	(97, 3.9e−08)	(101, 9.4e−10)	–	–	(93, 4.9e−05)
4	–	–	–	–	–	–	–	–	–	–	–	–	(82, 0.0014)
5	–	–	–	–	–	(98, 4.3e−10)	(93, 1.3e−05)	–	–	–	–	(99, 1.5e−08)	(98, 3.7e−07)
6	–	–	–	–	–	–	(89, 5.6e−06)	–	–	–	–	(97, 1.8e−09)	(90, 2.7e−07)
7	–	–	–	–	–	–	–	–	–	–	–	(88, 0.00053)	(95, 3e−05)
8	–	–	–	–	–	–	–	–	(98, 4.7e−09)	(100, 7.5e−09)	–	–	(108, 1.2e−07)
9	–	–	–	–	–	–	–	–	–	(100, 1.1e−09)	–	–	(100, 7.1e−11)
10	–	–	–	–	–	–	–	–	–	–	–	–	(96, 1.1e−06)
11	–	–	–	–	–	–	–	–	–	–	–	–	(104, 1.5e−05)
12	–	–	–	–	–	–	–	–	–	–	–	–	(99, 6.7e−08)

Table 5
The values (k, p_k) for Experiment II. The variable k is the number of identical responses given by two observers in a set of 126 trials, and the p-value p_k measures how significant it is to observe at least k identical responses.

Subject	1	2	3	4	5	6	7	8	9	10	Algorithm
1	–	(110, 2.3e−11)	(96, 0.00043)	(100, 5.3e−06)	(98, 0.0013)	(106, 5.9e−09)	(93, 2.4e−05)	(108, 9.2e−10)	(102, 1e−05)	(99, 1.1e−05)	(100, 2.9e−05)
2	–	–	(100, 0.00063)	(104, 3.3e−06)	(104, 0.003)	(102, 9.2e−05)	(93, 5.5e−05)	(112, 5.9e−09)	(106, 0.00016)	(95, 0.009)	(104, 8.6e−05)
3	–	–	–	(104, 7.3e−07)	(98, 0.04)	(102, 1.8e−05)	(91, 0.00026)	(102, 0.0002)	(106, 8.7e−06)	(95, 0.0029)	(104, 1.1e−05)
4	–	–	–	–	(98, 0.015)	(106, 5.7e−08)	(95, 5.2e−06)	(104, 6.6e−06)	(106, 1.6e−06)	(103, 6.8e−07)	(98, 0.0014)
5	–	–	–	–	–	(104, 0.00019)	(91, 0.00042)	(104, 0.01)	(112, 2.1e−05)	(101, 0.00052)	(104, 0.003)
6	–	–	–	–	–	–	(91, 0.00026)	(108, 1.4e−07)	(112, 6.8e−10)	(107, 8.4e−09)	(102, 9.2e−05)
7	–	–	–	–	–	–	–	(91, 0.00041)	(95, 3.9e−06)	(92, 8.3e−05)	(89, 0.0019)
8	–	–	–	–	–	–	–	–	(112, 2.5e−07)	(97, 0.0036)	(102, 0.0015)
9	–	–	–	–	–	–	–	–	–	(101, 0.00031)	(110, 9.5e−07)
10	–	–	–	–	–	–	–	–	–	–	(99, 0.00037)

Table 6
The values (k, p_k) for Experiment III. The variable k is the number of identical responses given by two observers in a set of 126 trials, and the p-value p_k measures how significant it is to observe at least k identical responses.

Subject	1	2	3	4	5	6	7	8	9	10	Algorithm
1	–	(108, 1.9e−13)	(110, 6.1e−17)	(109, 2.3e−14)	(103, 1.5e−09)	(105, 3.9e−10)	(108, 1.9e−13)	(106, 1.7e−11)	(104, 4e−11)	(103, 1.1e−10)	(98, 1.6e−07)
2	–	–	(100, 6.9e−10)	(111, 4e−16)	(105, 3.9e−11)	(99, 3.5e−07)	(108, 1.1e−13)	(102, 2.1e−09)	(102, 4.2e−10)	(103, 7.9e−11)	(96, 8.3e−07)
3	–	–	–	(107, 2e−14)	(101, 1.8e−10)	(99, 1.1e−09)	(100, 6.9e−10)	(100, 7.6e−10)	(96, 6.9e−08)	(105, 4.7e−13)	(92, 4.8e−06)
4	–	–	–	–	(112, 1.4e−16)	(104, 1.8e−10)	(109, 1.4e−14)	(109, 3.3e−14)	(109, 7.9e−15)	(106, 7.8e−13)	(103, 1.5e−10)
5	–	–	–	–	–	(102, 1.6e−07)	(111, 1.8e−15)	(103, 3.3e−09)	(99, 4.5e−08)	(100, 8.8e−09)	(101, 9.1e−09)
6	–	–	–	–	–	–	(103, 1.8e−09)	(105, 1.6e−09)	(97, 1.1e−06)	(96, 1.9e−06)	(101, 2.8e−08)
7	–	–	–	–	–	–	–	(106, 7.2e−12)	(98, 5.7e−08)	(97, 1.3e−07)	(102, 7.6e−10)
8	–	–	–	–	–	–	–	–	(106, 2.8e−12)	(103, 1.5e−10)	(96, 2e−06)
9	–	–	–	–	–	–	–	–	–	(101, 7.5e−10)	(98, 5.7e−08)
10	–	–	–	–	–	–	–	–	–	–	(89, 0.0002)

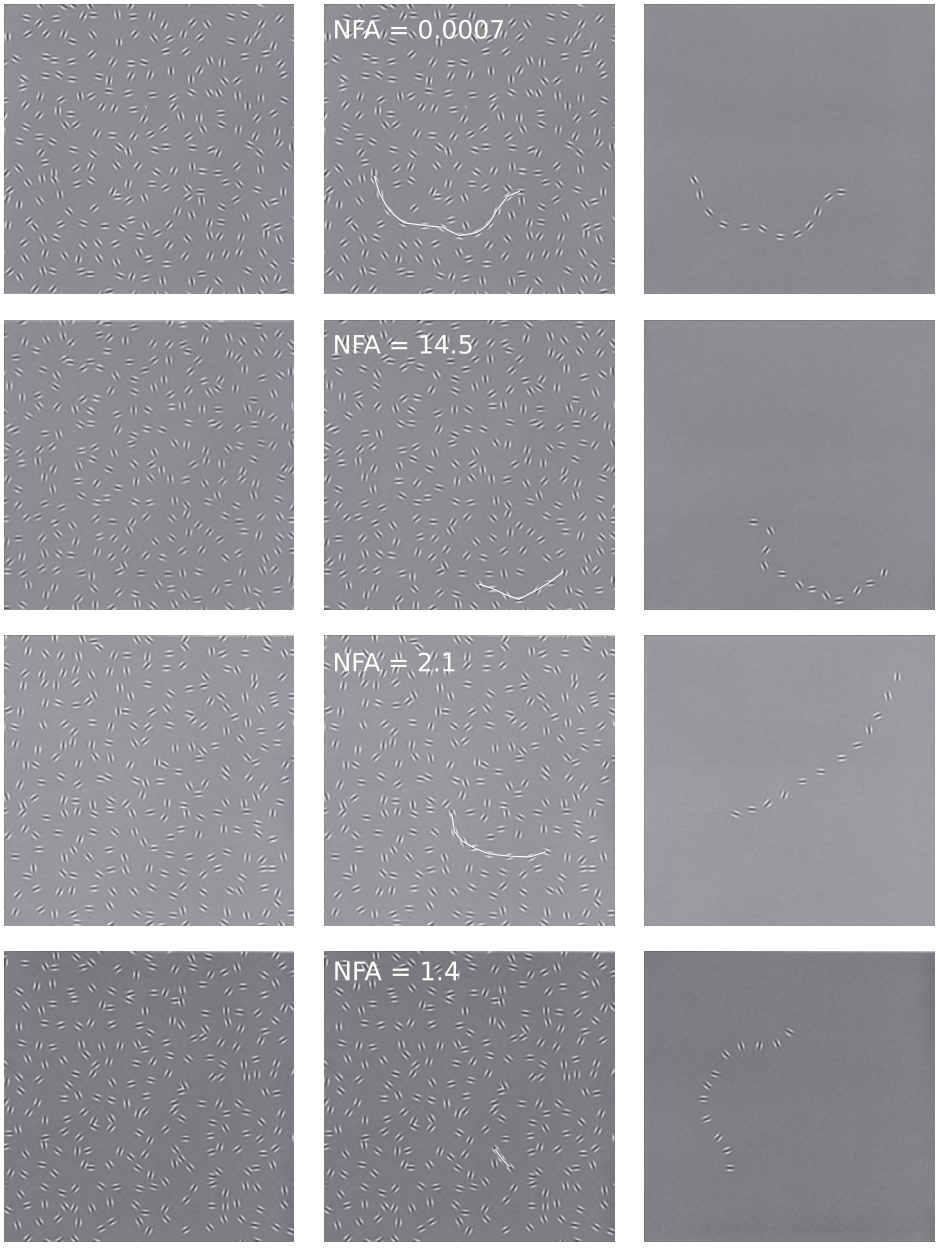


Fig. 8. Some examples of contour detection by a slightly modified version of our algorithm. In the left hand are images extracted from Field et al. (1993) (from top to bottom row: Figs. 3, 6, 10 and 8). The coordinates and orientations of the patches were computed by smoothing the image, selecting the pixels darker than 0.6 times its mean value, and by calculating the center and main direction of each resulting connected component. As described in Field et al. (1993), in this kind of arrays, the Gabor patches are approximately set on a square grid. This is why we changed the number of visited neighbors by the algorithm from 6 to 4, and changed the number of tests N_T accordingly. The middle column displays the chain with smallest NFA, found by our algorithm. In the right hand column we report the corresponding target path, extracted from Field et al. (1993) as well.

Nevertheless, the algorithm does perform quite similarly to the subjects. As Tables 4–6 show, its responses agree significantly with those of the subjects ($p \leqslant 0.003$). This agreement is actually comparable to the inter-subjects one. We have thus a statistical confirmation of the good match observed in Fig. 7.

Furthermore, this similarity is not the result of the optimization of a set of parameters to get the best possible fit. It is the direct output of one unique non-parametric algorithm, that was compared to human subjects on every single trial. In short, while a model fit to the subjects results would be easily obtained by tuning the parameters of a parametric algorithm, here the fit between algorithm and subjects is absolute.

Another observation that must be pointed out is that the NFA of the target alignments is an accurate one-dimension measure of the stimuli's difficulty. Indeed, in Fig. 7, columns C and D show decreasing detection rates and increasing reaction times with respect to the NFA. According to Proposition 1, points with $\log_{10}(\text{NFA})$ lower than -2 in these plots represent stimuli in which the target alignment is expected to occur less than once every hundred *a contrario* array of Gabor patches. At the other end of the NFA axis, on the right of $\log_{10}(\text{NFA}) = 2$ for example, stand the images whose hidden alignment is an event that could occur, on average, more than a hundred times in a single *a contrario* Gabor array, and is thus indistinguishable from noise. Consequently, what columns C and D show in Fig. 7, is a strong correlation between detectability and non-accidentalness.

Moreover, these results raise the hope that the *a contrario* framework might become a useful tool for other psychophysical

Please cite this article in press as: Blusseau, S., et al. Measuring the visual salience of alignments by their non-accidentalness. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.08.014

studies of the non-accidentalness principle in perceptual grouping. It allows the translation of many parameters used to build stimuli, into one interpretable measure that permits comparing all the stimuli with each other. In Fig. 7, columns C and D are an example of such flexibility, since they present the detection rates and reaction times as a function of the NFA, over a wide range of stimuli ruled by three parameters. In the present study we focused on the detection of straight contours in order to start with a model easy to define and to explain. But the same technique could be adapted to the study of symmetry, motion, or more general contour integration.

To illustrate this, we made a minor modification in our algorithm permitting to measure the non-accidentalness of stimuli proposed in Field et al. (1993) (Fig. 8). In this experiment, for each image extracted from Field et al. (1993), the coordinates and orientations of the patches were computed by smoothing the image, selecting the pixels darker than 0.6 times its mean value, and by calculating the center and main direction of each resulting connected component. As described in Field et al. (1993), in this kind of arrays, the Gabor patches are approximately set on a square grid. This is why we changed the number of visited neighbors by the algorithm from 6 to 4. We changed the number of tests N_T accordingly. The detection results again show a good fit between NFA and detectability of curves.

The limitations of the present work also require a discussion. First, in the experimental protocol most aspects of the viewing conditions were not normalized. What is more, the difference in the Gabor patches' spatial frequency between Experiment I on one side, and Experiments II and III on the other side, may have induced effects that we did not expect and that were not the scope of our study. Also, the fact that the subjects were only minimally trained may add variability among their response criteria and overall reduce the detection rate in comparison to the algorithm. However, our intention was not to uncover new psychophysical phenomena, but to present a new interpretation of well known effects. In fact, despite the just mentioned limitations, the behavior observed in our experiments is consistent with previous studies on contour integration and the association field (Field et al., 1993; Ledgeway et al., 2005; Ernst et al., 2012; Nygård et al., 2009).

Comparing our results to those of the latter studies may also be questionable, though, beyond the fact that we focused on straight contours. Indeed, contrary to the referred experiments, in ours, attention played an important part in the perceptual process, since the subjects were presented with the stimuli without time limitation. Thus, the good fit between the ideal observer's and the subjects' performance is explainable, as the subjects had all time to find the good continuations. Reproducing our experiments according to a preattentive paradigm could be the object of a future work.

9. Conclusion

This paper has presented an attempt to interpret and quantify a specific class of perceptual grouping, thanks to a mathematical model of the non-accidentalness principle. We exposed a way to apply the *a contrario* theory to implement a parameterless algorithm, adapted to the detection of a fixed Gestalt in a masking background. The strong correlation between the subjects' detection performance and the NFA, as well as the match between the algorithm and the subjects' curves, seem to argue in favor of the non-accidentalness principle as a way to interpret and predict the perceptual grouping of jittered alignments. We believe that the presented method could be adapted to address other questions of quantitative Gestalt, with the NFA as independent variable in psychophysical experiments.

Acknowledgments

Work partly founded by Centre National d'Etudes Spatiales (CNES, MISS Project), European Research Council (advanced grant Twelve Labours), Office of Naval research (ONR grant N00014–14-1–0023), DGA Stéréo project, ANR-DGA (project ANR-12-ASTR-0035), FUI (project Plein Phare) and Institut Universitaire de France.

References

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review, 61*(3), 183–193. http://dx.doi.org/10.1037/h0054663.

Binford, T. O. (1981). Inferring surfaces from images. *Artificial Intelligence, 17*(1), 205–244. http://dx.doi.org/10.1016/0004-3702(81)90025-4.

Blusseau, S., Carboni, A., Maiche, A., Morel J., Grompone von Gioi R. (2014). A psychophysical evaluation of the a contrario detection theory, in: *Image Processing (ICIP), 2014 IEEE International Conference on, 2014* (pp. 1091–1095). doi: http://dx.doi.org/10.1109/ICIP.2014.7025217.

Cao F., Lisani J., Morel J. -M., Musé P., Sur F. (2008). A theory of shape identification, Vol. 1948 of Lecture Notes in Mathematics, Springer. http://dx.doi.org/10.1007/978-3-540-68481-7.

Cardelino J., Caselles V., Bertalmio M., Randall G. (2009). A contrario hierarchical image segmentation. In: *Image Processing (ICIP), 2009 16th IEEE International Conference on, 2009*, pp. 4041–4044. doi: http://dx.doi.org/10.1109/ICIP.2009.5413723.

Demeyer M., Machilsen B. (2011). The construction of perceptual grouping displays using GERT, Behavior Research Methods, online first 1–8.

Desolneux, A., Moisan, L., & Morel, J. M. (2003). Computational gestalts and perception thresholds. *Journal of Physiology – Paris, 97*, 311–324. http://dx.doi.org/10.1016/j.jphysparis.2003.09.006.

Desolneux, A., Moisan, L., & Morel, J. M. (2008). *From Gestalt Theory to Image Analysis, a Probabilistic Approach*. Springer.

Ernst, U. A., Mandon, S., Schinkel-Bielefeld, N., Neitzel, S. D., Kreiter, A. K., & Pawelzik, K. R. (2012). Optimality of human contour integration. *PLoS Computational Biology, 8*(5), e1002520. http://dx.doi.org/10.1371/journal.pcbi.1002520.

Feldman, J. (1997a). Regularity-based perceptual grouping. *Computational Intelligence, 13*(4), 582–623.

Feldman, J. (1997b). Curvilinearity, covariance, and regularity in perceptual groups. *Vision Research, 37*(20), 2835–2848. http://dx.doi.org/10.1016/S0042-6989(97)00096-5.

Feldman, J. (2001). Bayesian contour integration. *Attention, Perception, & Psychophysics, 63*(7), 1171–1182. http://dx.doi.org/10.3758/BF03194532.

Feldman, J. (2003). Perceptual grouping by selection of a logically minimal model. *International Journal of Computer Vision, 55*(1), 5–25.

Feldman, J. (2007). Formation of visual objects in the early computation of spatial relations. *Perception & Psychophysics, 69*(5), 816–827.

Feldman, J. (2009). Bayes and the simplicity principle in perception. *Psychological Review, 116*(4), 875.

Feldman, J., & Singh, M. (2005). Information along contours and object boundaries. *Psychological Review, 112*(1), 243–252. http://dx.doi.org/10.1037/0033-295X.112.1.243.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local association field. *Vision Research, 33*(2), 173–193. http://dx.doi.org/10.1016/0042-6989(93)90156-Q. URL http://www.sciencedirect.com/science/article/pii/004269899390156Q.

Fleuret F., Li T., Dubout C., Wampler, E.K., Yantis, S., Geman, D. (2011). Comparing machines and humans on a visual categorization test, PNAS 108(43). pp. 17621–17625. arXiv: http://www.pnas.org/content/108/43/17621.full.pdf+html.

Geisler, W.S. (2011). Contributions of ideal observer theory to vision research, Vision Research 51 (7), 771 – 781, vision Research 50th Anniversary Issue: Part 1. doi:http://dx.doi.org/10.1016/j.visres.2010.09.027. URL http://www.sciencedirect.com/science/article/pii/S0042698910004724.

Gold, J. I., & Watanabe, T. (2010). Perceptual learning. *Current Biology, 20*(2). http://dx.doi.org/10.1016/j.cub.2009.10.066.

Grompone von Gioi, R., Jakubowicz, J., Morel, J. M., Randall G. (2012). *LSD: a line segment detector, Image Processing On Line*. doi:http://dx.doi.org/10.5201/ipol.2012.gjmr-lsd.

Hess, R., & Field, D. (1999). Integration of contours: New insights. *Trends in Cognitive Sciences, 3*(12), 480–486.

Hess, R., Hayes, A., & Field, D. (2003). Contour integration and cortical processing. *Journal of Physiology-Paris, 97*(23), 105–119. neurogeometry and visual perception. http://dx.doi.org/10.1016/j.jphysparis.2003.09.013. URL http://www.sciencedirect.com/science/article/pii/S0928425703000561.

Hess, R. F., May, K. A., & Dumoulin, S. O. (2014). Contour integration: Psychophysical, neurophysiological and computational perspectives. In J. Wagemans (Ed.), *Oxford Handbook of Perceptual Organization*. USA: Oxford University Press.

Kanade, T. (1981). Recovery of the three-dimensional shape of an object from a single view. *Artificial Intelligence, 17*(13), 409–460. http://dx.doi.org/10.1016/0004-3702(81)90031-X. URL http://www.sciencedirect.com/science/article/pii/000437028190031X.

Kanizsa, G. (1980). Grammatica del vedere, Il Mulino.
Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron, 15*(4), 843–856. http://dx.doi.org/10.1016/0896-6273(95)90175-2. http://www.sciencedirect.com/science/article/pii/0896627395901752.
Kovacs, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: Effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Sciences, 90*(16), 7495–7497.
Ledgeway, T., Hess, R. F., & Geisler, W. S. (2005). Grouping local orientation and direction signals to extract spatial contours: Empirical tests of association field models of contour integration. *Vision Research, 45*(19), 2511–2522.
Lowe, D. G. (1985). *Perceptual organization and visual recognition* (Ph.D. thesis). Standford University.
Lowe, D. G. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence, 31*(3), 355–395. http://dx.doi.org/10.1016/0004-3702(87)90070-1.
Lowe, D. G. (1999). Object recognition from local scale-invariant features. In: *Computer vision. The proceedings of the seventh IEEE international conference on, Vol. 2, IEEE* (pp. 1150–1157).
Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision, 60*(2), 91–110. http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94.
Lowe, D. G., & Binford, T. O. (1981). The interpretation of three-dimensional structure from image curves. In *Proceedings of the 7th international joint conference on Artificial intelligence – Volume 2, IJCAI'81* (pp. 613–618). San Francisco, CA, USA: Morgan Kaufman Publishers Inc.. URL http://dl.acm.org/citation.cfm?id=1623264.1623267.
Machilsen, B., & Wagemans, J. (2011). Integration of contour and surface information in shape detection. *Vision Research, 51*, 179–186. http://dx.doi.org/10.1016/j.visres.2010.11.005.
Machilsen, B., Pauwels, M., & Wagemans, J. (2009). The role of vertical mirror symmetry in visual shape detection. *Journal of Vision, 9*(12), 11.
Metzger, W. (1975). *Gesetze des Sehens* (3rd ed.). Frankfurt am Main: Verlag Waldemar Kramer.
Mumford D., Desolneux A., *Pattern theory: the stochastic analysis of real-world signals*, AK Peters, 2010.
Musé, P., Sur, F., Cao, F., Gousseau, Y. (2003). Unsupervised thresholds for shape matching. In: *IEEE Int. Conf. on Image Processing, ICIP*, Barcelona, Spain.
Nygård, G., Van Looy, T., & Wagemans, J. (2009). The influence of orientation jitter and motion on contour saliency and object identification. *Vision Research, 49*, 2475–2484.
Pătrăucean V., Gurdjos, P., Grompone von Gioi, R. (2012). A parameterless line segment and elliptical arc detector with enhanced ellipse fitting. In: A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, C. Schmid (Eds.), *Computer Vision ECCV 2012, Lecture Notes in Computer Science*, Springer Berlin Heidelberg, pp. 572–585. doi:10.1007/978-3-642-33709-3_41. URL http://dx.doi.org/10.1007/978-3-642-33709-3_41.
Pătrăucean, V., Grompone von Gioi, R., Ovsjanikov, M. (2013). Detection of mirror-symmetric image patches, in: *CVPRW*, pp. 211–216.
Rock, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.
Sassi, M., Demeyer, M., & Wagemans, J. (2014). Peripheral contour grouping and saccade targeting: The role of mirror symmetry. *Symmetry, 6*(1), 1–22.
Stevens, K. A., (1980). *Surface perception from local analysis of texture and contour* (Ph.D. thesis). Massachusetts Institue of Technology.
Stevens, K. A. (1981). The visual interpretation of surface contours. *Artificial Intelligence, 17*(1), 47–73. http://dx.doi.org/10.1016/0004-3702(81)90020-5.
Tepper M., Musé P., Almansa, A. (2011). Meaningful clustered forest: an automatic and robust clustering algorithm, CoRR abs/1104.0651. URL http://dblp.uni-trier.de/db/journals/corr/corr1104.html#abs-1104-0651.
Vancleef, K., & Wagemans, J. (2013). Component processes in contour integration: A direct comparison between snakes and ladders in a detection and a shape discrimination task. *Vision Research, 92*, 39–46.
van der Helm, P. A. (2000). Simplicity versus likelihood in visual perception: From surprisals to precisals. *Psychological Bulletin, 126*(5), 770. http://dx.doi.org/10.1037/0033-2909.126.5.770.
van der Helm, P. A. (2011). Bayesian confusions surrounding simplicity and likelihood in perceptual organization. *Acta Psychologica, 138*(3), 337–346. http://dx.doi.org/10.1016/j.actpsy.2011.09.007. URL http://www.sciencedirect.com/science/article/pii/S0001691811001661.
Van Lier, R., van der Helm, P., & Leeuwenberg, E. (1994). Integrating global and local aspects of visual occlusion. *Perception-London, 23*. http://dx.doi.org/10.1068/p230883. 883–883.
Wagemans, J. (1992). Perceptual use of nonaccidental properties. *Canadian Journal of Psychology, 46*(2), 236–279.
Wagemans, J. (1993). Skewed symmetry: A nonaccidental property used to perceive visual forms. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 364–380.
Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin, 138*(6), 1172–1217. http://dx.doi.org/10.1037/a0029333.
Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., & van Leeuwen, C. (2012). A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. *Psychological Bulletin, 138*(6), 1218–1252. http://dx.doi.org/10.1037/a0029334.
Wertheimer, M. (1923). Untersuchungen zur lehre von der gestalt. ii. *Psychologische Forschung, 4*(1), 301–350. http://dx.doi.org/10.1007/BF00410640.
Wilder, J., Feldman, J., & Singh, M. (2015). Contour complexity and contour detection. *Journal of Vision, 15*(6), 6. http://dx.doi.org/10.1167/15.6.6. arXiv:/data/Journals/JOV/933934/i1534-7362-15-6-6.pdf.
Witkin, A.P. (1982). Intensity-based edge classification. In: *AAAI, Vol. 82*, (pp. 36–41). URL http://www.aaai.org/Library/AAAI/1982/aaai82-009.php.
Witkin, A. P. (1983). Scale-space filtering. In *Proceedings of the eighth international joint conference on artificial intelligence – Volume 2, IJCAI'83* (pp. 1019–1022). San Francisco, CA, USA: Morgan Kaufman Publishers Inc.. http://dl.acm.org/citation.cfm?id=1623516.1623607.
Witkin, A. P., & Tenenbaum, J. M. (1983). *Human and machine vision*. Academic Press. Ch. On the role of structure in vision, pp. 481–543.
Yen, S.-C., & Finkel, L. H. (1998). Extraction of perceptually salient contours by striate cortical networks. *Vision Research, 38*(5), 719–741. http://dx.doi.org/10.1016/S0042-6989(97)00197-1. http://www.sciencedirect.com/science/article/pii/S0042698997001971.

Please cite this article in press as: Blusseau, S., et al. Measuring the visual salience of alignments by their non-accidentalness. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.08.014