L’information visuelle: Shannon et Attneave

Plan:

- La théorie de la communication de Shannon
- La théorie de la perception visuelle de Attneave
- Le principe de Helmholtz et la mise en œuvre computationnelle (Desolneux, Moisan, M.)

Sources:

C.E. Shannon: A mathematical theory of communication

F. Atteneave: Some informational aspects of visual

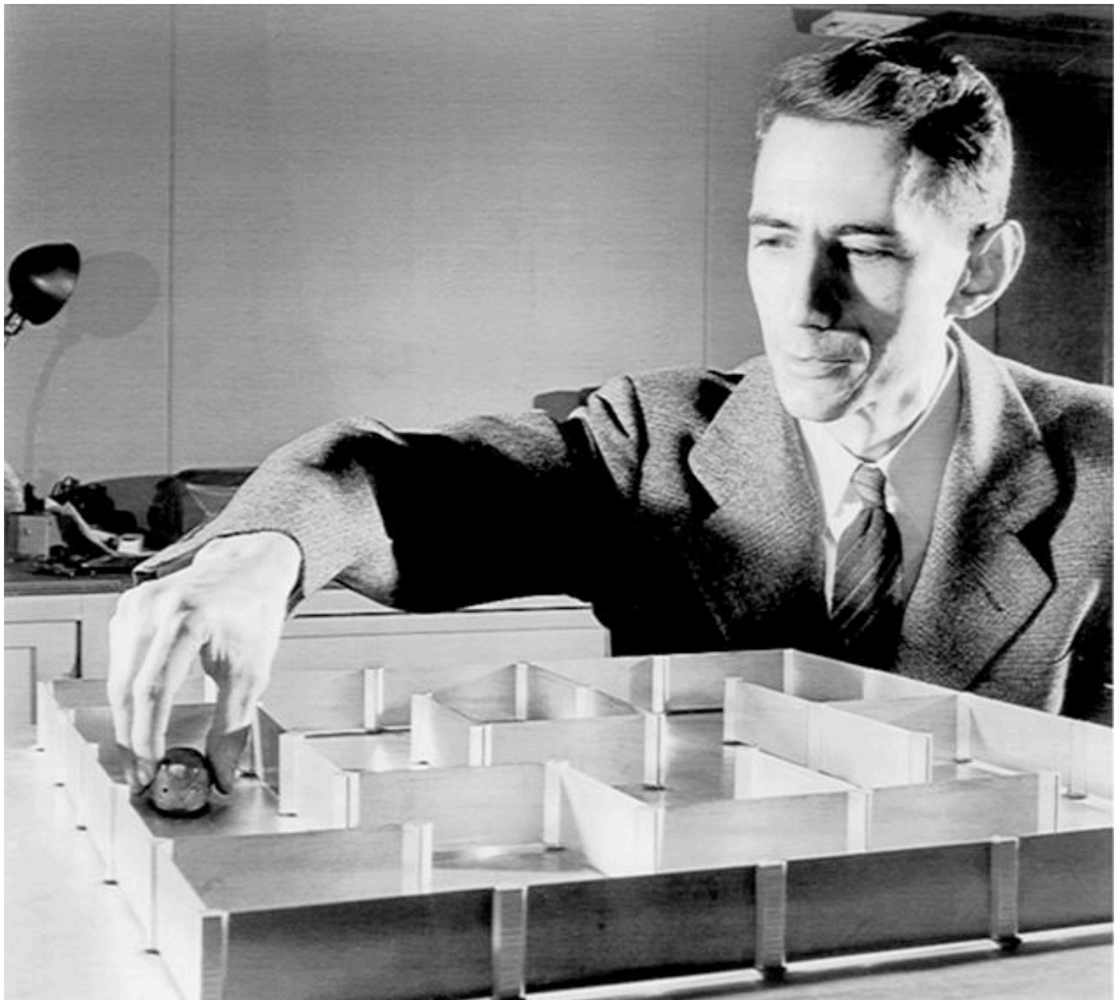
A. Desolneux, L. Moisan, JMM: Frome Gestalt theory to image analysis: a probabilistic approach

R. Grompone: LSD, (Line segment detector), on line demo,

http://www.ipol.im/

A Mathematical Theory of Communication

By C. E. SHANNON



The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design.

If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely. As was pointed out by Hartley the most natural choice is the logarithmic function. Although this definition must be generalized considerably when we consider the influence of the statistics of the message and when we have a continuous range of messages, we will in all cases use an essentially logarithmic measure.

The choice of a logarithmic base corresponds to the choice of a unit for measuring information. If the base 2 is used the resulting units may be called binary digits, or more briefly *bits,* a word suggested by J. W. Tukey. A device with two stable positions, such as a relay or a flip-flop circuit, can store one bit of information. N such devices can store N bits, since the total number of possible states is 2^N and $\log_2 2^N = N$.

3. The Series of Approximations to English

To give a visual idea of how this series of processes approaches a language, typical sequences in the approximations to English have been constructed and are given below. In all cases we have assumed a 27-symbol "alphabet," the 26 letters and a space.

1. Zero-order approximation (symbols independent and equiprobable).

   XFOML RXKHRJFFJUJ ZLPWCFWKCYJ FFJEYVKCQSGHYD QPAAMKBZAACIBZL-HJQD.

2. First-order approximation (symbols independent but with frequencies of English text).

   OCRO HLI RGWR NMIELWIS EU LL NBNESEBYA TH EEI ALHENHTTPA OOBTTVA NAH BRL.

3. Second-order approximation (digram structure as in English).

   ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TU-COOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE.

4. Third-order approximation (trigram structure as in English).

   IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PONDENOME OF DEMONS-TURES OF THE REPTAGIN IS REGOACTIONA OF CRE.

5. First-order word approximation. Rather than continue with tetragram, . . . , n-gram structure it is easier and better to jump at this point to word units. Here words are chosen independently but with their appropriate frequencies.

   REPRESENTING AND SPEEDILY IS AN GOOD APT OR COME CAN DIFFERENT NAT-URAL HERE HE THE A IN CAME THE TO OF TO EXPERT GRAY COME TO FURNISHES THE LINE MESSAGE HAD BE THESE.

6. Second-order word approximation. The word transition probabilities are correct but no further structure is included.

   THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT THE CHAR-ACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE TIME OF WHO EVER TOLD THE PROBLEM FOR AN UNEXPECTED.

Theorem 2: The only H satisfying the three above assumptions is of the form:

$$H = -K \sum_{i=1}^{n} p_i \log p_i$$

It is easily shown that

$$H(x,y) \leq H(x) + H(y)$$

with equality only if the events are independent (i.e., $p(i,j) = p(i)p(j)$). The uncertainty of a joint event is less than or equal to the sum of the individual uncertainties.

$$H(x,y) = H(x) + H_x(y).$$

The uncertainty (or entropy) of the joint event x,y is the uncertainty of x plus the uncertainty of y when x is known.

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = H(X) + H(Y) - H(X,Y).$$

The first defining expression has already been defined as the amount of information sent less the uncertainty of what was sent. The second measures the amount received less the part of this which is due to noise. The third is the sum of the two amounts less the joint entropy and therefore in a sense is the number of bits per second common to the two. Thus, all three expressions have a certain intuitive significance. $H(Y \mid X)$ est donc interprétée comme une mesure de la quantité de bruit, sans que l'on ait pour cela besoin de modéliser le bruit par lui-même. Rappelons que $I(X;Y)$ est une formule symétrique en X et Y, que l'on appelle aussi information mutuelle, et qui est nulle si et seulement si X et Y sont indépendantes.

Psychological Review
Vol 61, No. 3, 1954

SOME INFORMATIONAL ASPECTS OF VISUAL PERCEPTION

FRED ATTNEAVE

The Nature of Redundancy in Visual Stimulation: A Demonstration

Consider the very simple situation presented in Fig. 1. With a modicum of effort, the reader may be able to see this as an ink bottle on the corner of a desk. Let us suppose that the background is a uniformly white wall, that the desk is a uniform brown, and that the bottle is completely black. The visual stimulation from these objects is highly redundant in the sense that portions of the field are highly predictable from other portions. In order to demonstrate this fact and its perceptual significance, we may employ a variant of the "guessing game" technique with which Shannon (17) has studied the

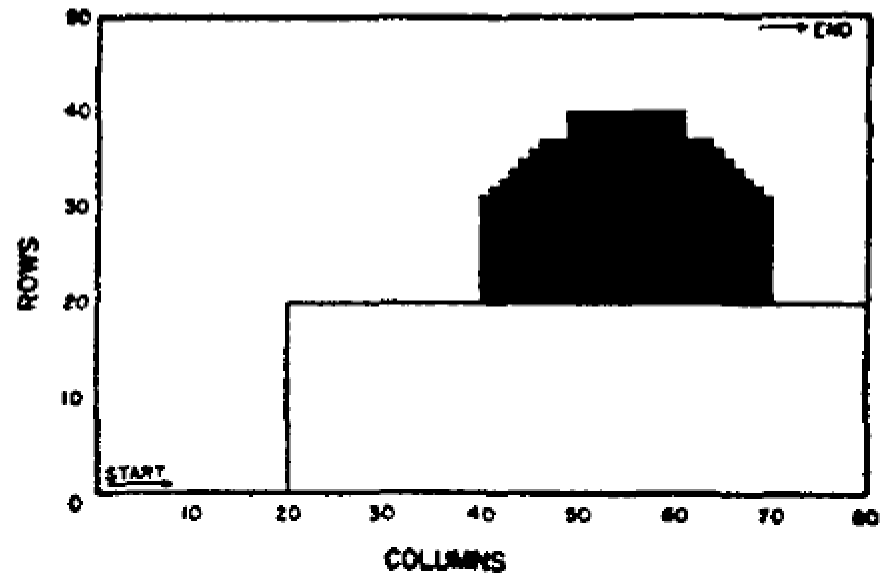


Fig. 1. Illustration of redundant visual stimulation

redundancy of printed English. We may divide the picture into arbitrarily small elements which we "transmit" to a subject (*S*) in a cumulative sequence, having him guess at the color of each successive element until he is correct. This method of analysis resembles the scanning process used in television and facsimile systems, and accomplishes the like purpose of transforming two spatial dimensions into a single sequence in time. We are in no way supposing or assuming, however, that perception normally involves any such scanning process. If the picture is divided into 50 rows and 80 columns, as indicated, our *S* will guess at each of 4,000 cells as many times as necessary to determine which of the three colors it has. If his error score is significantly less than chance $[2/3 \times 4{,}000 + 1/2(2/3 \times 4{,}000) = 4{,}000]$, it is evident that the picture is to some degree redundant. Actually, he may be expected to guess his way through Fig. 1 with only 15 or 20 errors. It is fairly apparent that the technique described, in its present form, is limited in applicability to simple and somewhat contrived situations.

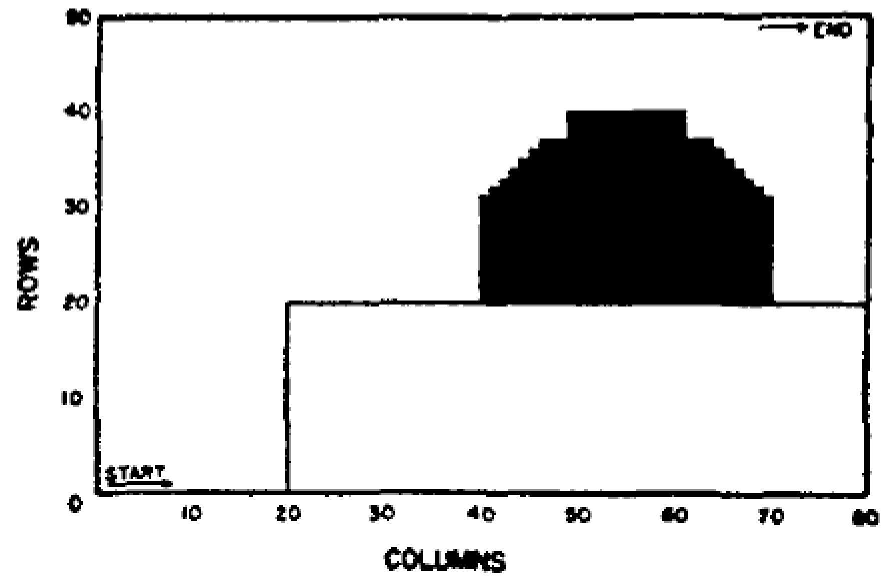


FIG. 1. Illustration of redundant visual stimulation

After a few errors at the beginning of the first row, he will discover that the next cell is "always" white, and predict accordingly. This prediction will be correct as far as Column 20, but on 21 it will be wrong. After a few more errors he will learn that "brown" is his best prediction, as in fact it is to the end of the row. Chances are good that the subject will assume the second row to be exactly like the first, in which case he will guess it with no errors; otherwise he may make an error or two at the beginning, or at the edge of the "table," as before. He is almost certain to be entirely correct on Row 3, and on subsequent rows through 20. On Row 21, however, it is equally certain that he will erroneously predict a transition from white to brown on Column 21, where the *corner* of the table is passed.

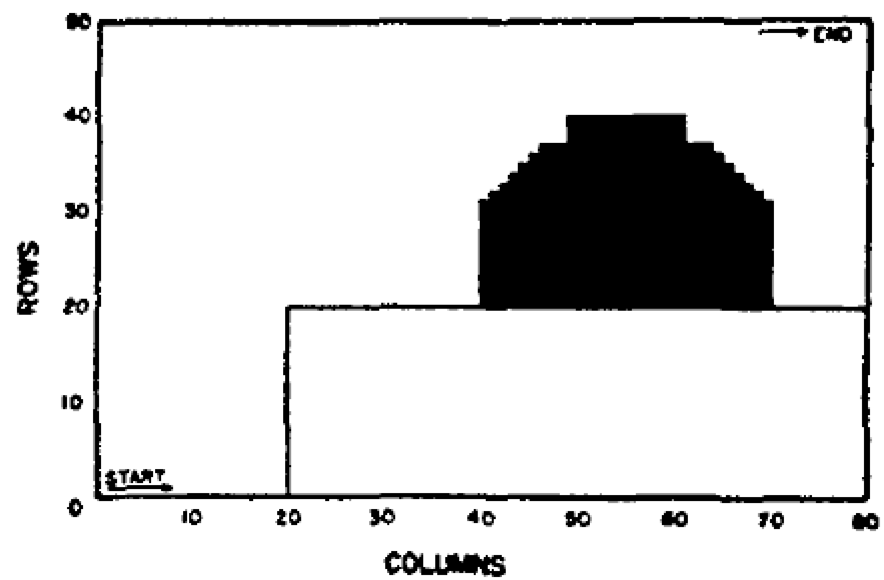


FIG. 1. Illustration of redundant visual stimulation

[4] The reader may be comforted to know that six subjects have actually been run on the task described. Their errors, which ranged in number from 13 to 26, were distributed as suggested above, with a single interesting exception: 4 of the 6 *S*s assumed on Row 1 that the brown area would be located symmetrically within the field, and guessed "white" on Column 61. By the use of Shannon's formulas (17) it was estimated that the field contains between 34 (lower limit) and 156 (upper limit) bits of information, in contrast to a possible maximum of 6,340 bits. The redundancy is thus calculated to be between 97.5 and 99.5 per cent.

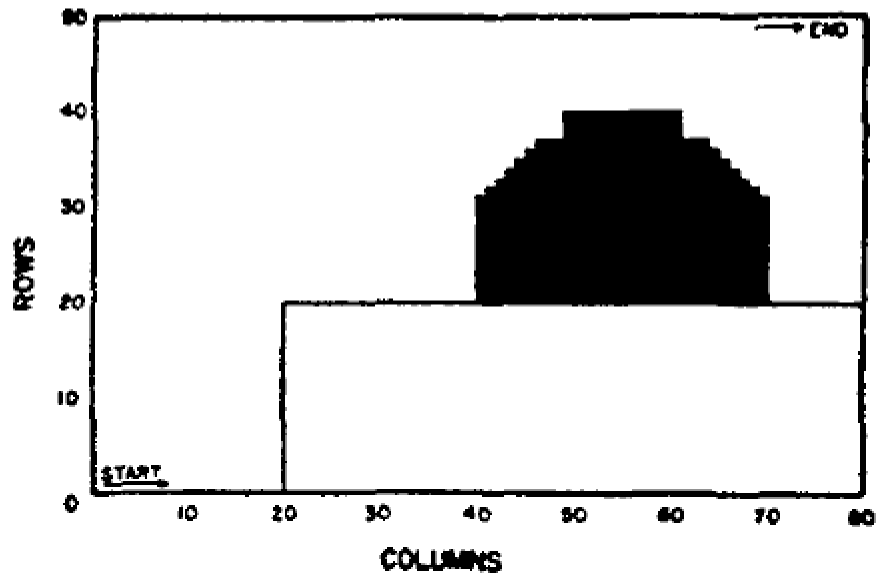


FIG. 1. Illustration of redundant visual stimulation

Our subject's behavior to this point demonstrates two principles which may be discussed before we follow him through the remainder of his predictions. It is evident that redundant visual stimulation results from either (*a*) an area of homogeneous color ("color" is used in the broad sense here, and includes brightness), or (*b*) a contour of homogeneous direction or slope. In other words, information is concentrated along contours (i.e., regions where color changes abruptly),[2] and is further concentrated at those points on a contour at which its direction changes most rapidly (i.e., at angles or peaks of curvature).

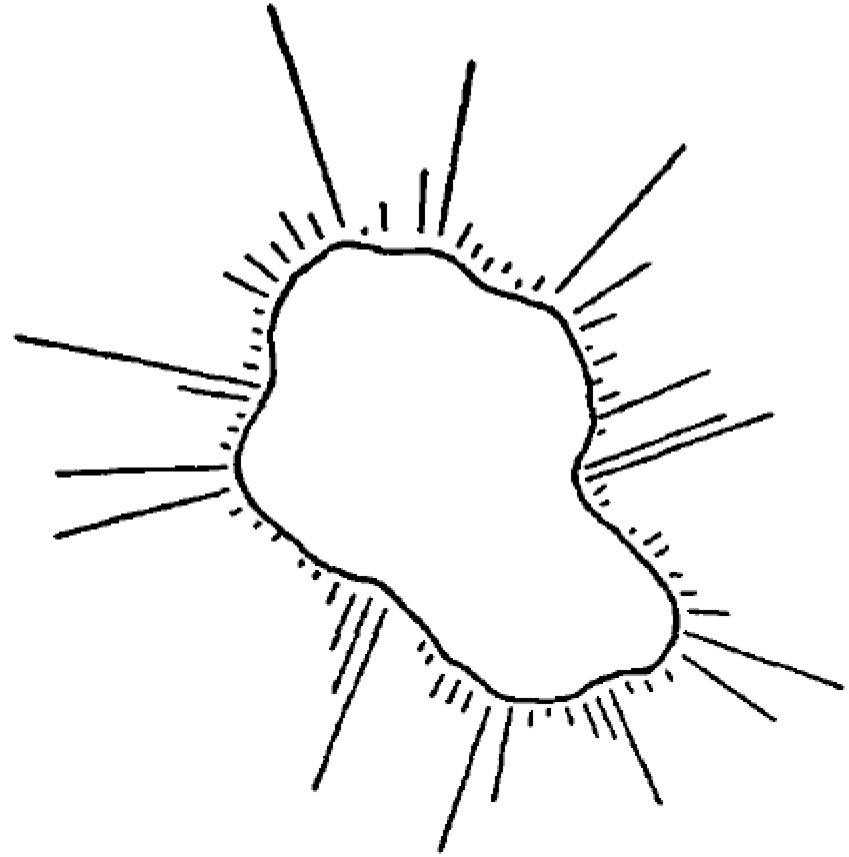


FIG. 2. Subjects attempted to approximate the closed figure shown above with a pattern of 10 dots. Radiating bars indicate the relative frequency with which various portions of the outline were represented by dots chosen.

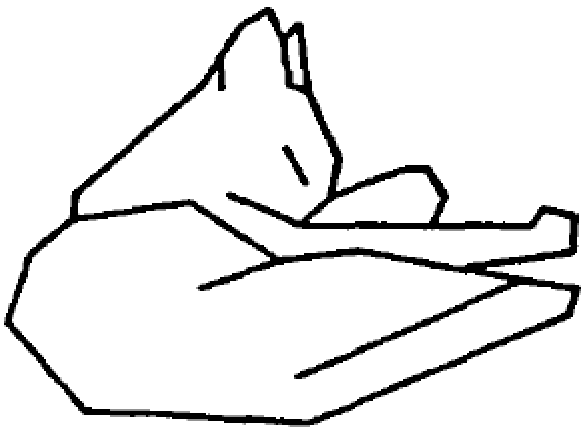


FIG. 3. Drawing made by abstracting 38 points of maximum curvature from the contours of a sleeping cat, and connecting these points appropriately with a straightedge.

It should be fairlv evident bv now that many of the gestalt principles of perceptual organization pertain essentially to information distribution. The *good gestalt* is a figure with some high degree of internal redundancy. That the grouping laws of *similarity*, *good continuation*, and *common fate* all refer to conditions which reduce uncertainty is clear enough after the preceding discussion, and we shall presently see that *proximity* may be conceptualized in a like manner. It is not surprising that the perceptual machinery should "group" those portions of its input which share the same information: any system handling redundant information in an efficient manner would necessarily do something of the sort. Musatti (20) came very close to the present point when he suggested that a single principle of *homogeneity* might subsume Wertheimer's laws as special cases. All of our hypothetical *S*'s extrapolations have involved some variety of homogeneity (or invariance), either of color, of slope, or of pattern.

A troublesome question arises in this connection: where does perception leave off and inductive reasoning begin?

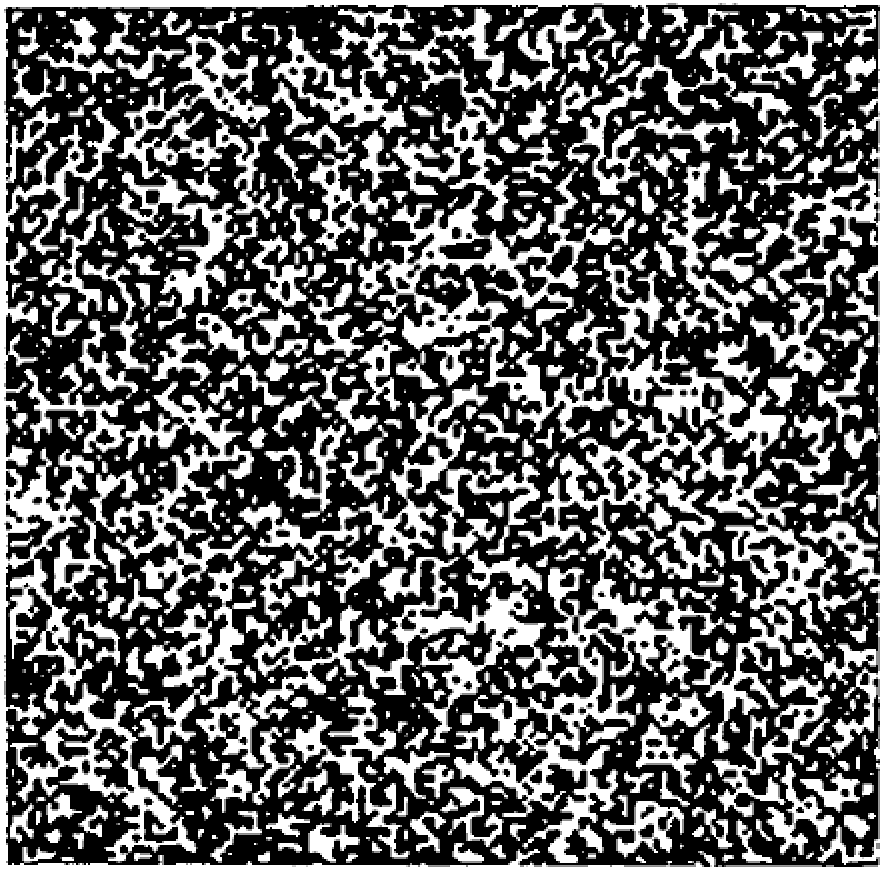


FIG. 4 A "random field" consisting of 19,600 cells. The state of each cell (black vs. white) was determined independently with a p of .50.

Further, it may be argued on neurological grounds that the human brain could not possibly utilize all the information provided by states of stimulation which were not highly redundant. According to Polyak's (14) estimate, the retina contains not less than four million cones. At any given instant each of these cones may be in either of two states: firing or not firing. Thus the retina as a whole might be in any one of about $2^{4,000,000}$ or $10^{1,200,000}$ states, each representing a different configuration of visual stimulation. Now, if by some unspecified mechanism each of these states were to evoke a different unitary response, and if a unitary response consists merely of the firing of a single unique neuron, then $10^{1,200,000}$ of such response-neurons would be required. The fantastic magnitude of this figure becomes somewhat apparent when one calculates that only about 10^{54} neurons could be packed into a cubic light year. The fact that the

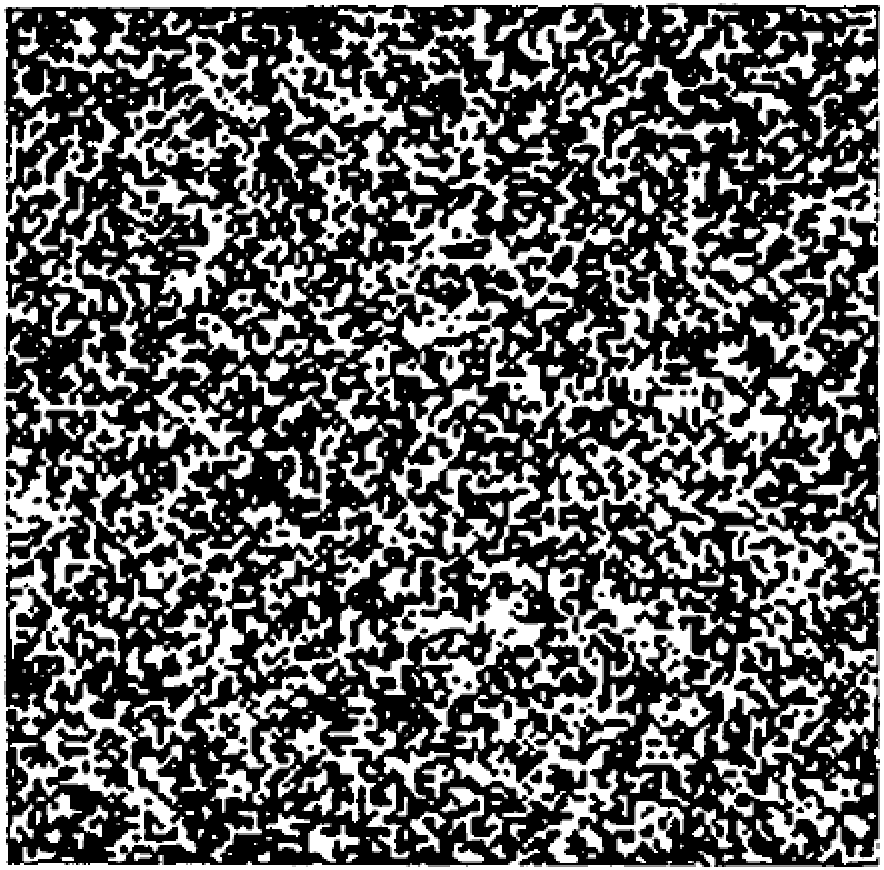


FIG. 4 A "random field" consisting of 19,600 cells. The state of each cell (black vs. white) was determined independently with a p of .50.

In an effort to get some notion of what such a random field would be like, Fig. 4 was constructed. Each of the $140^2 = 19{,}600$ small cells of the figure was either filled or not filled according to the value of a number obtained from a conversion of Snedecor's (18) table of random numbers from decimal to binary.[6] If the figure is viewed from a distance such that the angle subtended by a cell is of the order of the "minimum separable" (about 1′), it illustrates roughly how a small por-

[6] This laborious task was carried out by Airmen 1/C W. H. Price and E. F. Chiburis. Unfortunately, a slight distortion of the relative sizes of black and white cells was introduced in the photographic copying process. The figure was constructed not only for demonstration purposes, but also to serve as a source of random patterns for experimental use. It may also be used wherever a table of random binary numbers is needed, facilitating, for example, the selection of random "draws" from a binomial distribution.

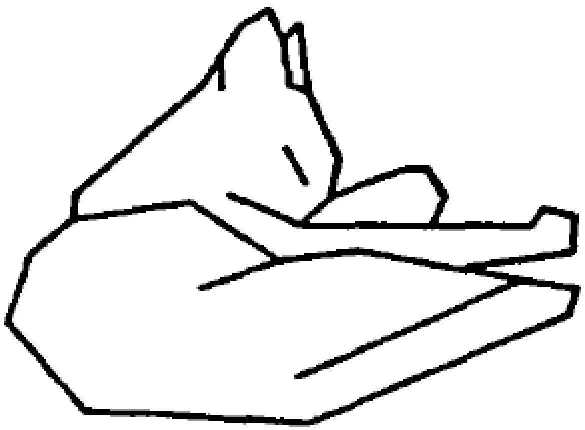


FIG. 3. *Drawing made by abstracting 38 points of maximum curvature from the contours of a sleeping cat, and connecting these points appropriately with a straightedge.*

It appears, then, that when some portion of the visual field contains a quantity of information grossly in excess of the observer's perceptual capacity, he treats those components of information which do not have redundant representation somewhat as a statistician treats "error variance," averaging out particulars and abstracting certain statistical homogeneities. Such an averaging process was involved in drawing the cat for Fig. 3. It was said earlier that the points of the drawing corresponded to places of maximum curvature on the contour of the cat, but this was not strictly correct; if the principle had been followed rigidly, it would have been necessary to represent the ends of individual hairs by points. In observing a cat, however, one does not ordinarily perceive its hairs as individual entities; instead one perceives that the cat is *furry*.

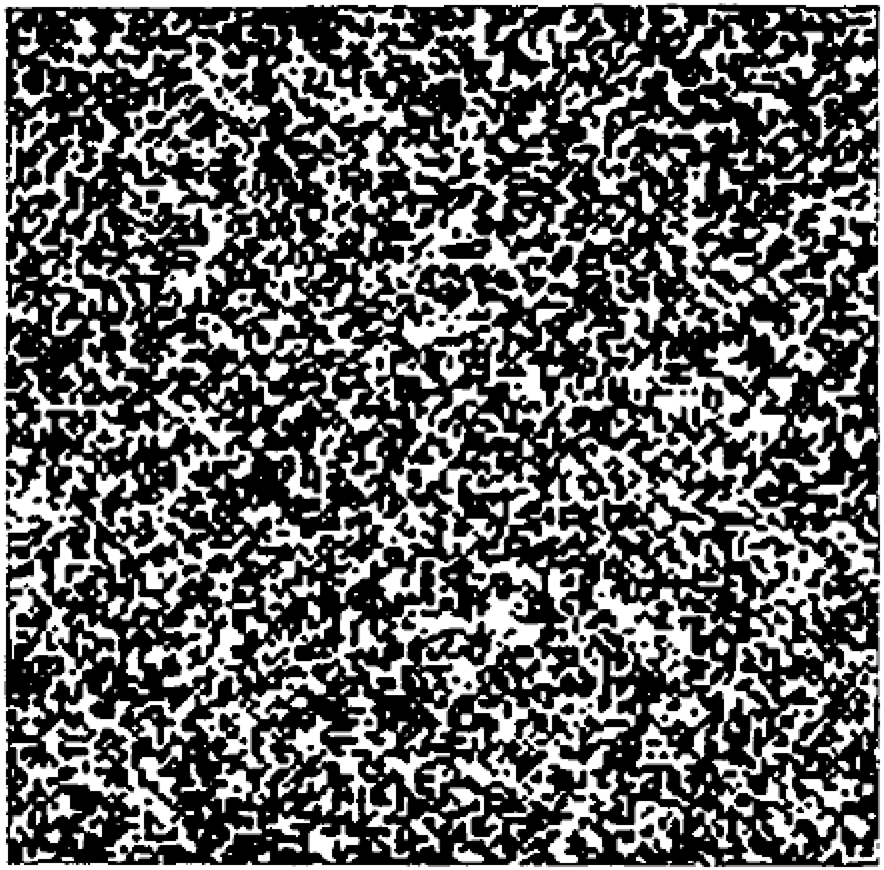


FIG. 4 A "random field" consisting of 19,600 cells. The state of each cell (black vs. white) was determined independently with a p of .50.

2. Likewise, an area of homogeneous texture may be described by specifying the statistical parameters which characterize the texture and the boundaries of the area over which these parameters are relatively invariant. Thus, if Fig. 4 represented a part of the upholstery of a sofa, it would probably be satisfactory simply to instruct the receiving mechanism to reproduce the texture by filling in cells of a certain size from any table of random numbers. It is true that this process would result in the complete loss of 19,600 bits of information; the essential point is that we are dealing here with a class of stimuli from which such a huge information loss is perceptually tolerable.

Since the points occupy a smaller range of alternative coordinates on the local axes than on arbitrary axes, less information is required for their specification. If the amount of information thus saved is greater than the amount needed to specify the positions of the local origins, a net saving will result. What is redundant in the present case is the *approximate location* of points in a cluster: this component is isolated out when a local origin is described

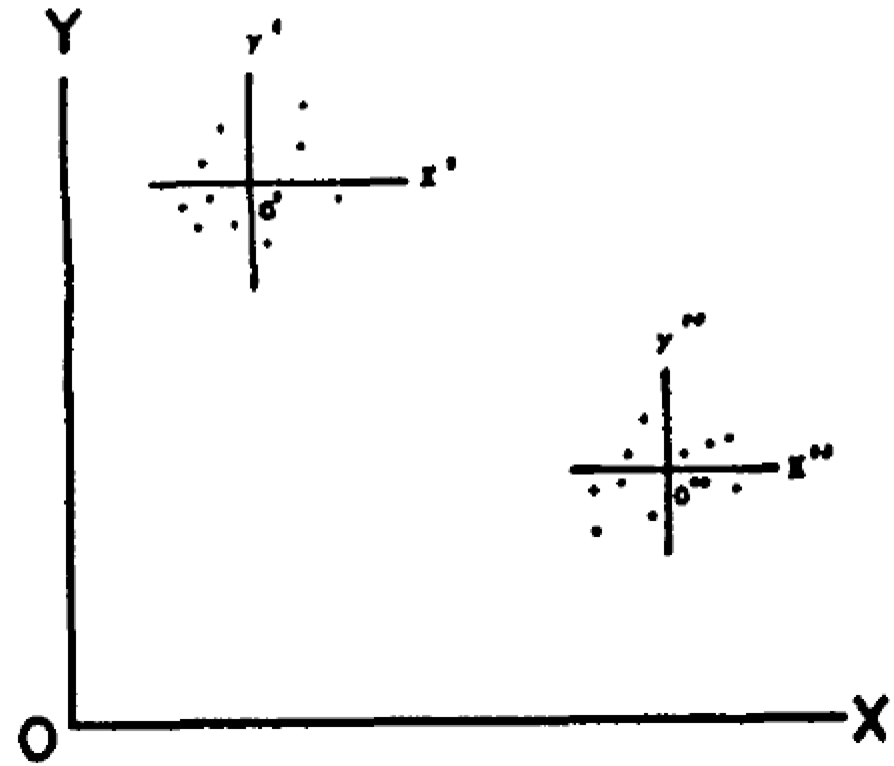
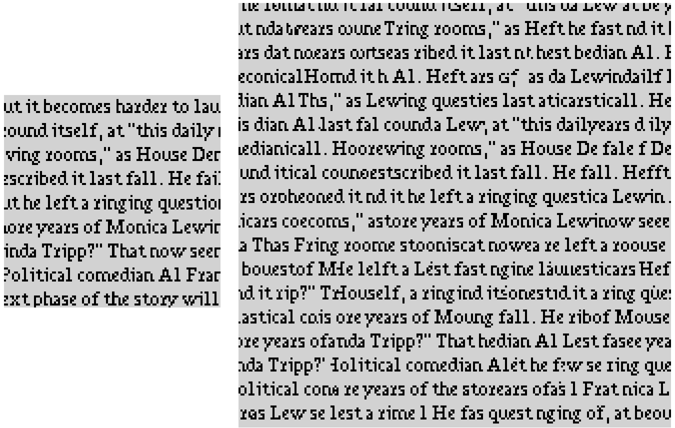
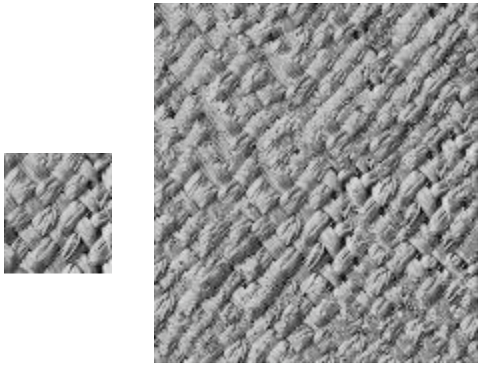


FIG. 5. A functional aspect of proximity-grouping is illustrated. The loci of clustered points may be described with choices from a smaller set of numbers if local origins are used.

Texture Synthesis by Non-parametric Sampling
A.A. Efros and T.K. Leung, ICCV 1999





From Gestalt Theory to Image Analysis

A Probabilistic Approach

Agnès Desolneux

Lionel Moisan

Jean-Michel Morel

The Helmholtz Principle

The Helmholtz principle can be formulated two ways. The first way is commonsensical. It simply states that we do not perceive any structure in a uniform random image. In this form, the principle was first stated by Attneave [Att54]. This gestaltist was to the best of our knowledge the first scientist to publish a random noise digital image. This image was actually drawn by hand by U.S. Army privates using a random number table. In its stronger form, of which we will make great use, the Helmholtz principle states that whenever some large deviation from randomness occurs, a structure is perceived.

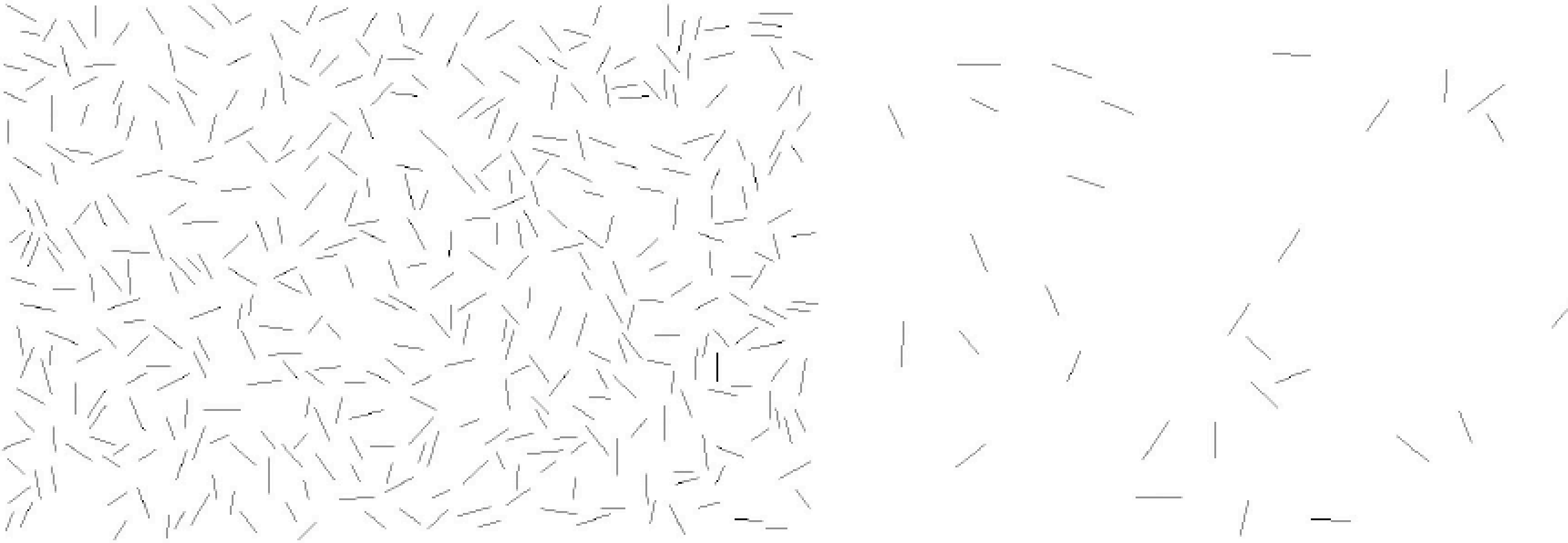


Fig. 3.3. The Helmholtz principle in human perception:

A group of four aligned segments exists in both images, but it can hardly be seen on the left-hand side image. Indeed, such a configuration is not exceptional in view of the total number of segments. In fact, the expectation of the number of aligned segments 4-tuples is about 10. In the right-hand image, we immediately perceive the alignment as a large deviation from randomness that could hardly happen by chance. In this image, the expectation of the number of groups of four aligned segments is about 1/4000.

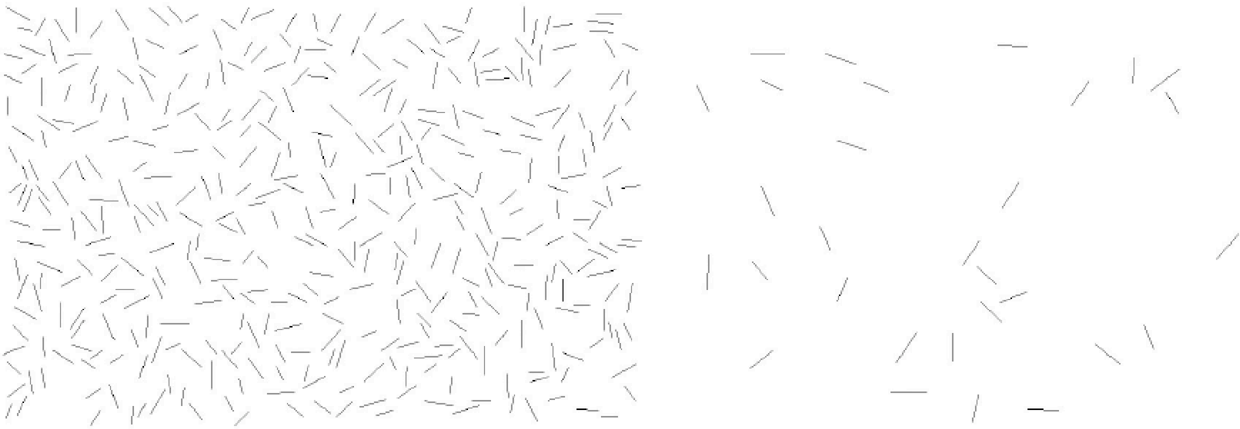


Fig. 3.3. The Helmholtz principle in human perception:

Assume that atomic objects O_1, O_2, ..., O_n are present in an image. Assume that k of them, say O_1,..., O_k, have a common feature (same color, same orientation, position, etc.). We then face a dilemma: Is this common feature happening by chance or is it significant and enough to group O_1, ..., O_k? To answer this question, let us make the following mental experiment: Assume *a priori* that the considered quality had been randomly and uniformly distributed on all objects O_1, ..., O_n. In the mental experiment, the observed position of objects in the image is a random realization of this uniform process. We finally ask the question: Is the observed repartition probable or not? If not, this proves *a contrario* that a grouping process (a gestalt) is at play. The Helmholtz principle states roughly that in such mental experiments, the numerical qualities of the objects are assumed to be uniformly distributed and independent.

A First Illustration: Playing Roulette with Dostoievski

That time, as if on purpose, a circumstance arose which, incidentally, recurs rather frequently in gambling. Luck sticks, for example, with red and does not leave it for ten or even fifteen turns. Only two days before, I had heard that red had come out twenty two times in a row in the previous week. One could never recall a similar case at roulette and it was spoken of with astonishment.

In the succession of fortuitous events, there is, if not a system, at least some kind of order. (...) It's very odd. On some afternoon or morning, black alternates with red, almost without any order and all the time. Each color only appears two or three times in a row. The next day or evening, red alone turns, for example, up to twenty times in a row.

Why 22? The probability that red appears 22 times in a row is $\left(\frac{18}{37}\right)^{22}$, namely about 10^{-7}. The computation of the probability that this happens in a series of n trials may be a bit intricate. We can, instead, directly compute the expected number of occurrences of the event as $\mathrm{NFA}(n) = (n-21) \times \left(\frac{18}{37}\right)^{22}$. The event is likely to happen if its NFA is larger than 1, which yields roughly $n \geq 10^7$. Thus, we are led to compute how many trials a passionate gambler may have done in his life. Considering that a professional gambler would play roulette at 100 evenings of 5 hours a year for 20 years, estimating in addition that a roulette trial may take about 30 seconds, we deduce that an experienced gambler would observe at the most, in his gambling life span, about $n = 20 \times 100 \times 5 \times 120 \simeq 10^6$ trials. We deduce that 1 out of 10 professional gamblers can have observed such a series of 22. Actually, Dostoievski's information about the possibility of 22 series is clearly based on conversations with specialists. The hero says:

I own a good part of these observations to Mr. Astley, who spends all of his mornings by the gambling tables but never gambles himself.

If this professional observer spent his time by several tables, maybe 10 simultaneously, he is, according to our computations, likely to have observed a series of 22. As we computed, 22 is somewhat a limit for an observable series. On the other hand, the hero mentions this occurrence as having happened just a few days before he was playing. There is no contradiction here, since, according to Aristotle, it is a rule of poetry, epics, and tragedy to put their heroes in exceptional situations. As he notices in his *Poetics*, exceptional situations do happen. Dostoievski twice puts his hero in an unlikely, but not impossible, situation. First, as we mentioned, is when a series of 22 occurs just a few days before the hero gets interested in roulette, second, a few days later, is when the hero observes a series of 14 reds and takes advantage of it to win a fortune. A series of 14 is unlikely to be observed by a beginner. The NFA of this happening to the hero during the three evenings he plays at the Roulettenbourg casino is, by the same kind of calculations as above, about $\text{NFA} = 3 \times 5 \times 120 \times \left(\frac{18}{37}\right)^{14} \simeq 4.10^{-3}$. Thus, this event is unlikely, yet, again, not impossible and therefore fits Aristotle's criterion.

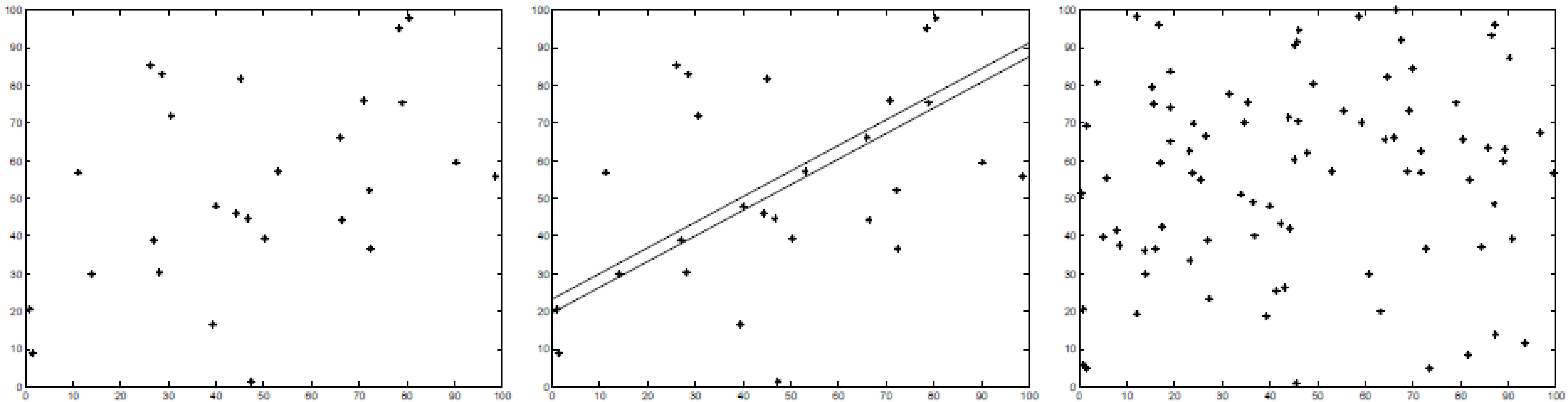


Fig. 3.4. The Helmholtz principle:

Noncasual alignments are automatically detected by the Helmholtz principle as a large deviation from randomness. Left: 20 uniformly randomly distributed dots and 7 aligned added. Middle: This meaningful and visible alignment is detected as a large deviation from randomness. Right: same alignment added to 80 random dots. The alignment is no more meaningful (and no longer visible). In order to be meaningful, it would need to contain at least 12 points.

LSD: a Line Segment Detector - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils ?

Google

Page précédente Page suivante Actualiser Arrêter Accueil http://mw.cmla.ens-cachan.fr/megawave/algo/lsd/ Aperçu Imprimer

Débuter avec Firefox À la une Google News France - ...

> mw > megawave > algo

LSD: A LINE SEGMENT DETECTOR

Contacts

- Rafael Grompone von Gioi grompone[AT]cmla.ens-cachan.fr
- Jérémie Jakubowicz jakubowi[AT]cmla.ens-cachan.fr
- Jean-Michel Morel morel[AT]cmla.ens-cachan.fr
- Gregory Randall randall[AT]fing.edu.uy

References

1. R. Grompone von Gioi, J. Jakubowicz, J.-M. Morel, G. Randall, *LSD: A Fast Line Segment Detector with a False Detection Control*. IEEE Trans. on PAMI, 19 Dec. 2008. preprint

Overview

A linear-time line segment detector that gives accurate results, a controlled number of false detections, and requires no parameter tuning. The method is based in Burns et al. method, and uses an *a contrario* validation approach.

On-line Demo

An on-line demo that allows you to try LSD with your own images is available here.

Software

Implementations in C programming language and Megawave2 framework are available here.

Video

The video here (mp4 file, 56 MB) shows the result of applying LSD, frame per frame, to the original video here.

Examples

Terminé