A Cognitive Model of Figure Segregation

Akira SHBMAYA and Isamu YOROIZAWA
NTT Human Interface Laboratories
1-2356 Take Yokosuka-shi Kanagawa 238-03 Japan
E-mail: shimaya%nttcvg.NTT.jp©relay.cs.net yoro%nttcvg.NTTjp@relayxs.net

Abstract

When humans look at a complex and ambiguous figure, they divide it into several elemental figures. This human visual characteristic is called figure segiegation. There is a problem when constructing a cognitive model for figure segiegation. That is, one interpretation is selected by most people for some figures, and several interpretations are selected almost equally for other figures. This paper discusses the deciding selection frequency problem. First, a geometrical function is introduced for describing line figures. Next, several Gestalt features (such as symmetry, continuity, etc) are defined using the function. Then, by applying linear multiple regression analysis, the characteristic value of each interpretation is obtained, and the selection frequency is calculated. The results of a psychological experiment show that the model proposed here can simulate human visual perception in figure segregation fairly well.

1 Introduction

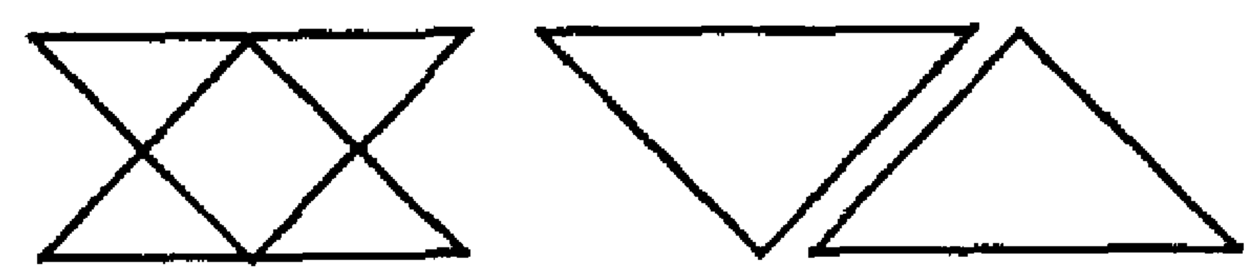
What is figure segregation ? When humans look at a complex and ambiguous figure, they divide it into several elemental figures (Figure 1). In psychology, this characteristic of human visual perception is called figure segregation. How can we construct a cognitive model of figure segregation ? Why is it difficult ? It is difficult because there are two processes to be considered.

(1) How to create a small number of natural interpretations from an infinite number of possible interpretations.

(2) How to decide selection frequencies among those interpretations, (in other words, how to estimate which interpretation will be selected by what percentage.)

Although research on figure segregation modeling is very limited, some research has been reported [Uesaka and Tajima, 1976] [Tuijl, 1980]. However, these models do not include process (1) at all and treat only process (2), and even in the process (2), there are the problems stated below.

(a) they decide segregation characteristic vale with a single measure; therefore, their algorithms are not robust and hard to improve.



(a) (b)

(a) An overlapped figure
(b) An interpretation of (a)
Figure. 1 An example of figure segregation

(b) their algorithms only estimate which interpretation is likely to be selected more; therefore, they cannot estimate selection frequency.

We have already reported the automatic creation of reasonable interpretations for complex and ambiguous figures to cover process (I) [Shiruaya and Yoroizawa, 1990a]- This paper discusses process (2). In order to solve problems (a) and (b), it is necessary to simulate human visual perception. We first introduce a generalised total curvature function for describing line figures in section 2. Five Gestalt factors(such as symmetry, continuity, etc) of each subfigure are defined with the total curvature function in section 3. Two more factors among the subfigures are defined in section 4- The effectiveness of each factor is examined from the results of a psychological experiment in section 5. Then by applying linear multiple regression analysis, the characteristic value of each interpretation is obtained, and the selection frequency is calculated in section 6. Section 7 is the conclusion that the model proposed here can simulate the human visual perception in figure segregation fairly well.

2 Description of a line figure

In order to check the geometrical characteristics of an interpretation, it is necessary to describe each subfigure in the interpretation. Initially, the subfigure is a closed line figure of two dimensions. Curvature can completely describe any two-dimensional smooth line figure. Let s be the length of a line from a starting point and $(x(s), y(s))$ represents a point using X-Y co-ordinates. Then curva-

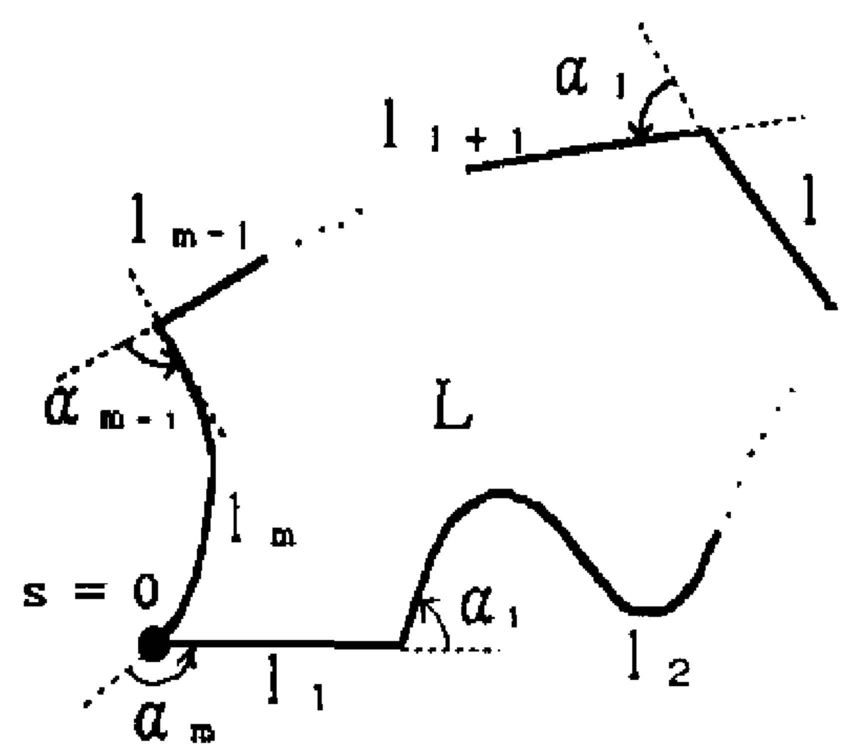


Figure. 2 Description of a line figure

ture k(s) of the point is given as

$$k(s) = \frac{dX(s)}{ds}\frac{d^2Y(s)}{ds^2} - \frac{dY(s)}{ds}\frac{d^2X(s)}{ds^2} \qquad (1)$$

However, the curvature itself is not suitable for describing line figures because it is not defined at the corners. Therefore, we introduce the total curvature function O*(s)* in order to describe general line figures. The total curvature function is originally the integral of curvature and defined with only smooth curves, but it can be applied to non-smooth curves by adding the angles at the discontinuous points in curvature[Uesaka and Tajima, 1976]. Now let's consider the line figure L (Figure 2). Let L be composed of m pieces of smooth arcs $L_1, l_2, \ldots, L_m$. *&i* is the discrete part between l_i and l_i+\. Let s_i be the length from the starting point to the end point of arc i. Then *8(s)* at a point at distance s $(s_n \leq s < s_{n+1})$ from the starting point is defined as

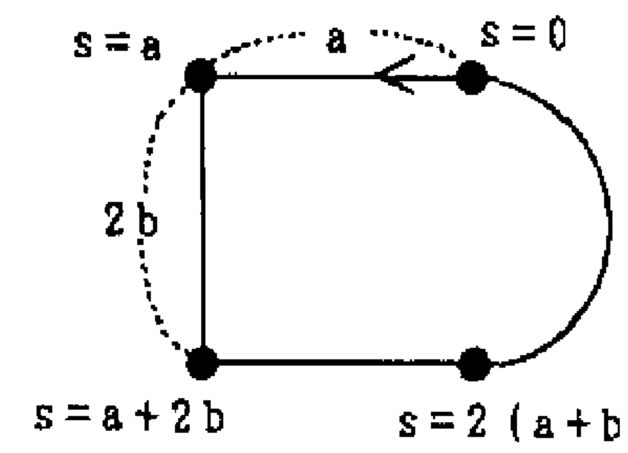
$$\theta(s) = \sum_{i=1}^{n}\{\int_{s_{i-1}}^{s_i} k_i(s)ds + \alpha_i\} + \int_{s_n}^{s} k_{n+1}(s)ds \qquad (2)$$

where

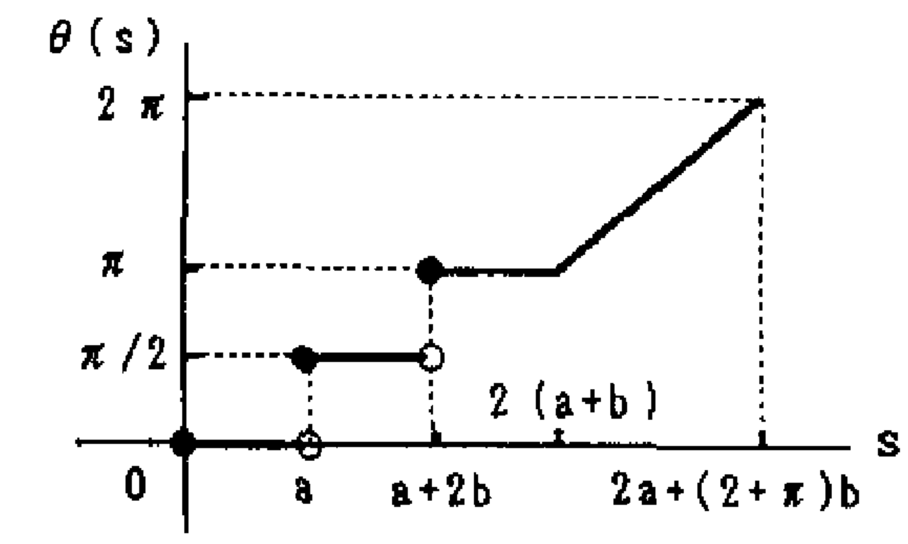
$k_i(a)$: curvature of a point on line segment i

a_i : the angle from the tangent at the end point of l_i, to the tangent at the starting point of $l_{\%+}$\. Let $-\pi < \alpha_i \leq \pi$.

An example of O*(s)* is shown in Figure 3. Note that if the starting point and tracing direction of a line figure is given, then O*(s)* of the figure is uniquely decided and that if 0(*s*) is given, the line figure is uniquely decided.

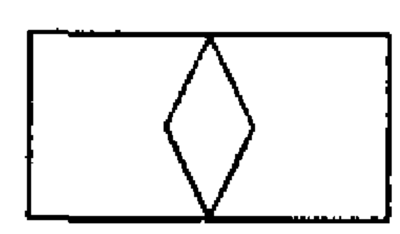


(a) A sample figure

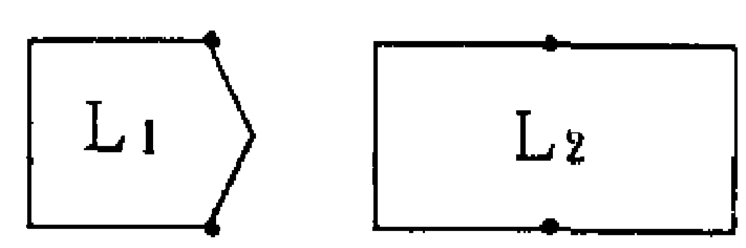


(b) Total curvature

Figure. 3 An example of total curvature function



(a) A sample figure



f1(L1)=2 f1(L2)=0

(b) Examples

Figure. 4 Simplicity

3 Description of Gestalt factors

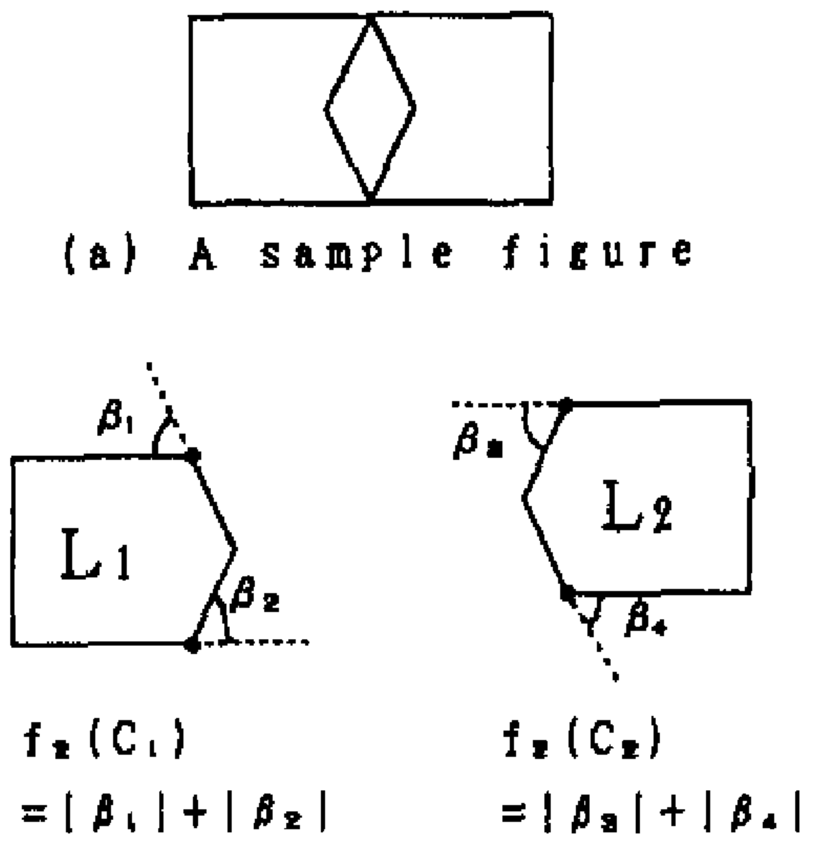
Figure segregation has been discussed in psychology, especially by Gestalt psychologists. They claimed that several factors(symmetry, continuity, etc) play an important role in figure segregation [Metzger, 1953] [Spoehr and Lehmkule, 1982], However, they only claimed the general effect of those Gestalt factors in figure segregation. In other words, their analyses were qualitative. In order to construct a cognitive figure segregation model, it is necessary to give quantitative measure to those factors [Shimaya and Yoroizawa, 1988], Therefore, we give quantitative definitions of the Gestalt factors in each subfigure with total curvature function. In this section, simplicity, continuity, symmetry, regularity, and convexity of subfigures are defined with O(s).

3*1 Simplicity

The number of discrete points in O*(s)* at the cross points of a subfigure represents the simplicity of the figure(Figure 4). Therefore, if *0(s)* of subfigure L has m discreation at cross points, then the simplicity of L is defined as

$$f_1 L) = m \qquad (3)$$

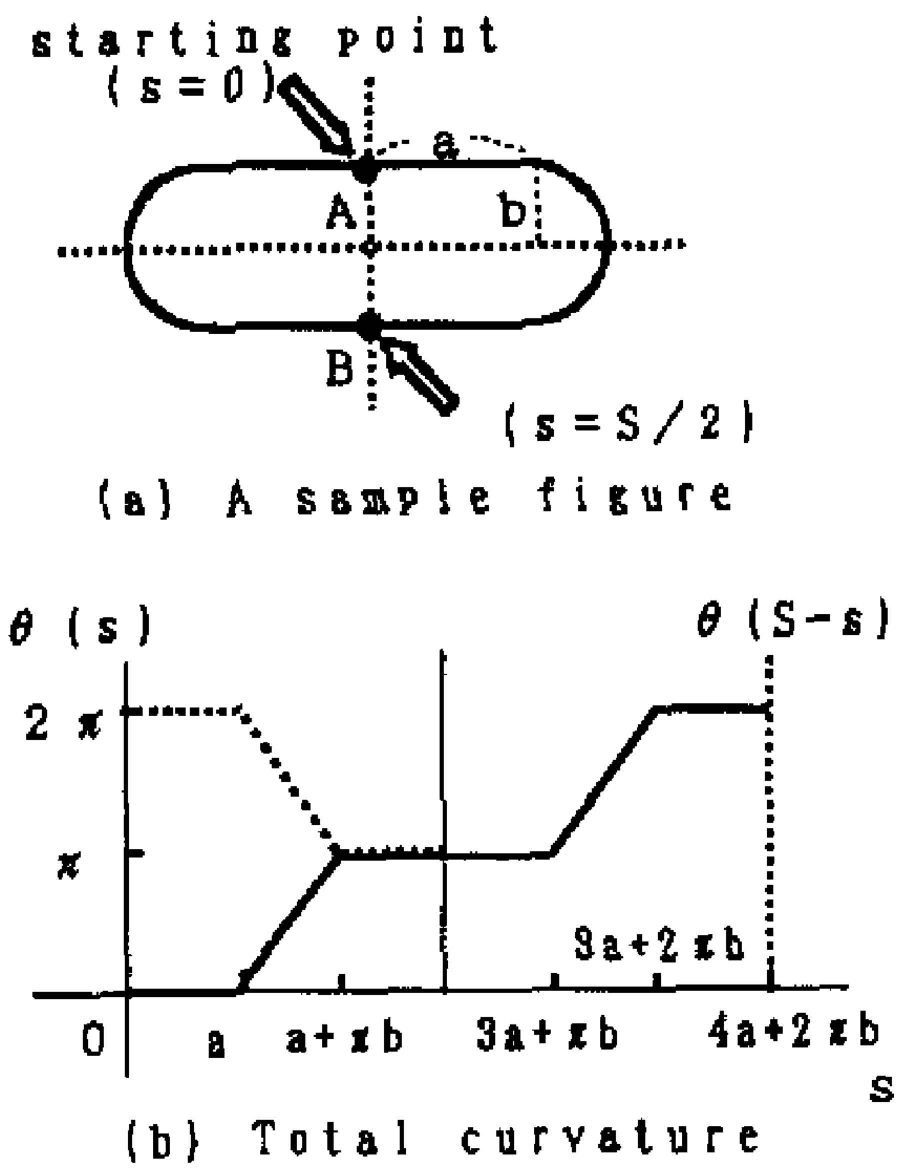
It can be said that the smaller f_1 of a subfigure is, the simpler it is.



(a) A sample figure

(b) Examples

Figure. 5 Continuity



(a) A sample figure

(b) Total curvature

Figure. 6 Symmetry with regard to axes

3.2 Continuity

The discrete parts of $\theta(s)$ at the cross points correspond to the continuity of subfigures in a certain segregation(Figure 5). Therefore, if the $\theta(s)$ of subfigure L has m discrete parts $(\beta_1, \beta_2, ..., \beta_m)$ at cross points, continuity of L is defined as

$$f_2(L) = \sum_{i=1}^{m} |\beta_i| \tag{4}$$

It can be said that the smaller f_2 of a subfigure is, the better it's continuity is.

3.3 Symmetry

If subfigure L (whose total curvature is $\theta(s)$) is symmetric with respect to an axis, there exists a starting point which satisfies the following equation(Figure 6).

$$\theta(s) + \theta(S - s) = \theta(S)$$

$$0 \le s \le \frac{S}{2} \tag{5}$$

where S is the total length of L. The line which connects the starting point and the point at a distance of S/2 becomes the symmetrical axis.

Thus, symmetry of L with regard to axes can be defined as

$$A_L = min\frac{1}{S}\int_0^S |\theta_x(s) + \theta_x(S - s) - \theta(S)|ds \tag{6}$$

where $\theta_x(s)$ is the generalized total curvature of L when the starting point is x. This definition means that the symmetry of a subfigure corresponds to the minimum distance of $\theta(s)$ between the subfigure and a symmetrical figure which has the same length.

Similarly, if subfigure L is symmetric with regard to n th order of rotation (in other words, if L returns to it's original shape by rotating at least $360/n$ degree), then there exists a starting point which satisfies the following equation(Figure 7).

$$\theta(s) + \theta(\frac{S}{n}) = \theta(s + \frac{S}{n})$$

$$0 \le s \le S - \frac{S}{n} \tag{7}$$

Thus, symmetry of L with regard to rotation can be defined as

$$R_L = min\frac{1}{S}\int_0^S |\theta_x(s) + \theta_x(\frac{S}{n}) - \theta_x(s + \frac{S}{n})|ds \tag{8}$$

Then, symmetry of L is defined as

$$f_3(L) = min(A_L, R_L) \tag{9}$$

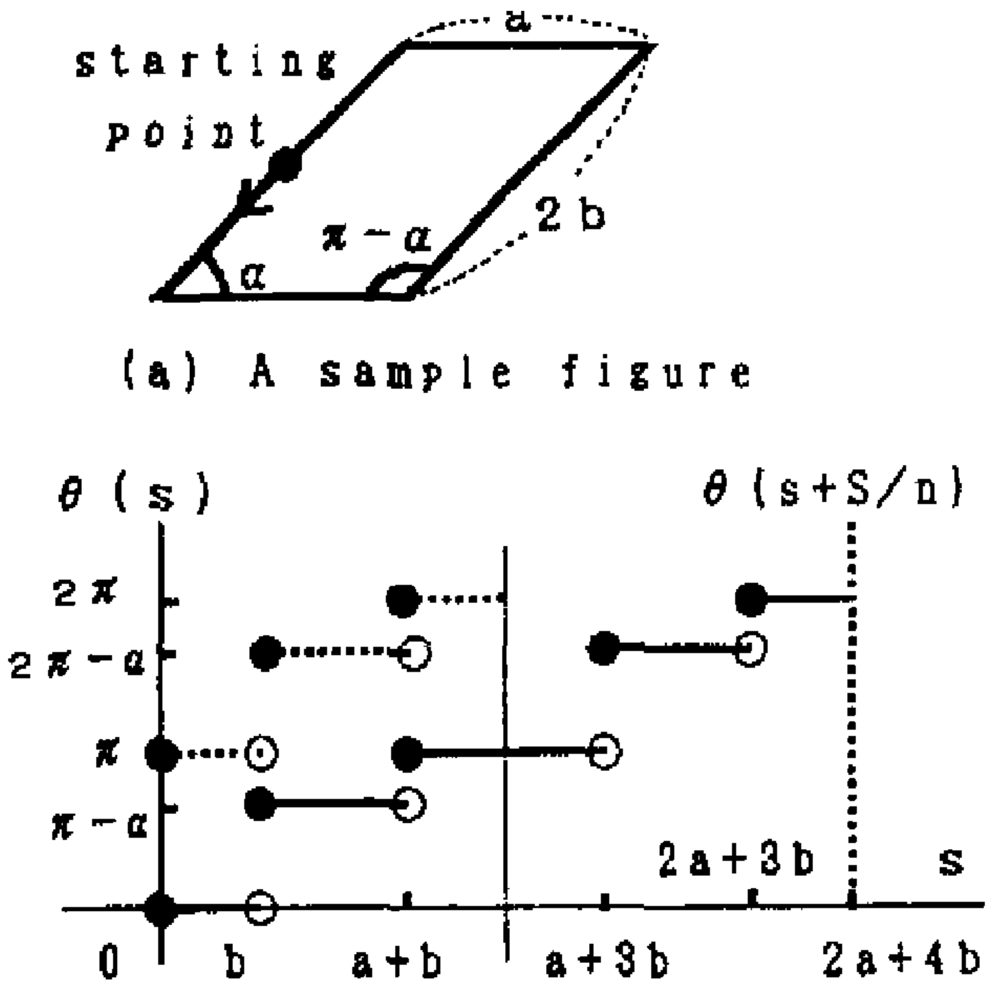
if L has n symmetrical axes and also is symmetric with m th order of rotation, then $f_3(L)$ is redefined as follows.

$$f_3(L) = min(\frac{1}{n}, \frac{1}{m-1}) \tag{10}$$

This means that parallelograms and trapezoids that have a symmetrical axis are the same level of symmetry. It can be said that the smaller f_3 of a subfigure is, the more symmetric it is.

3.4 Regularity

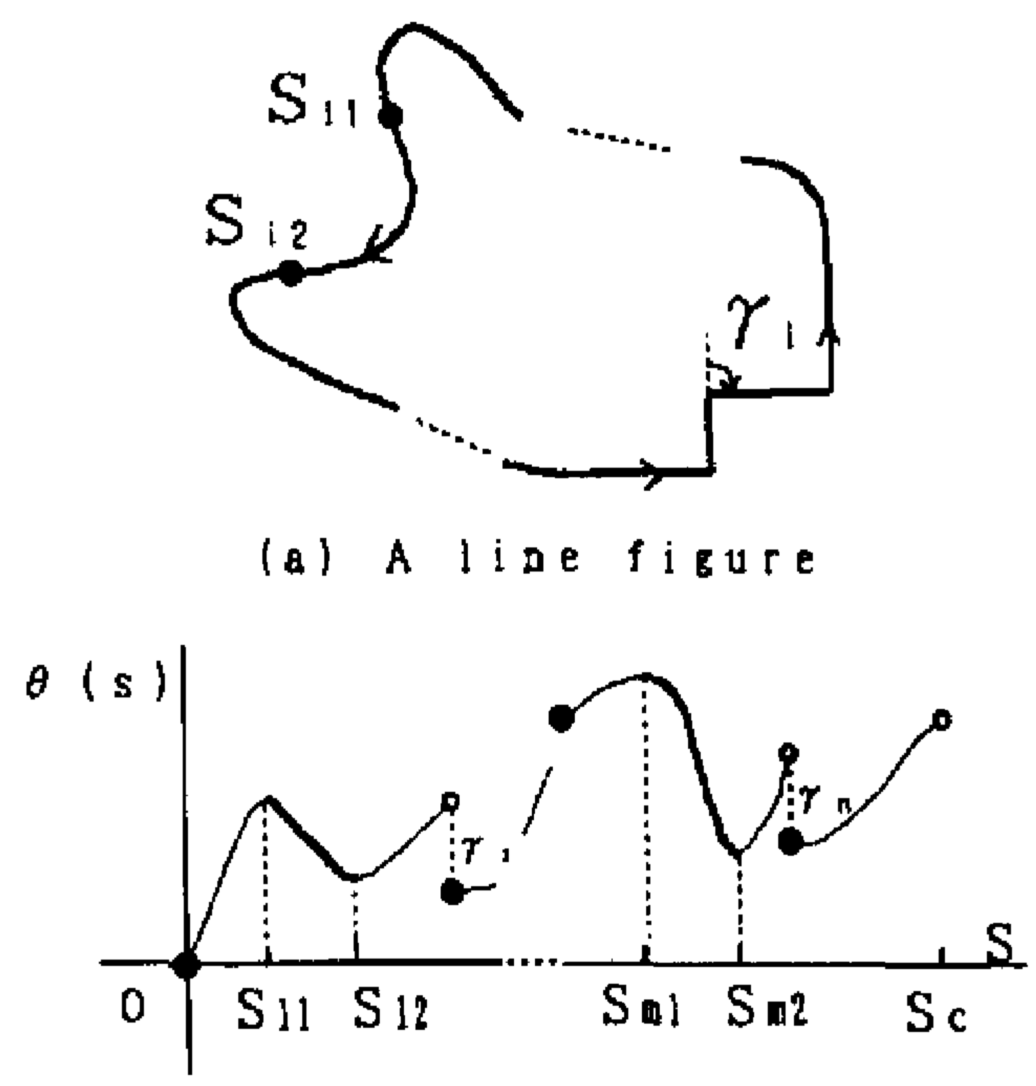
It is difficult to say how much regular a figure is. Humans feel circular arcs and straight lines are very regular. Therefore, in order to check the regularity of a subfigure, we first approximate $\theta(s)$ of component lines to that of straight lines or circular arcs. The curvature of circular arcs is constant and the curvature of straight lines is 0, therefore, both of their $\theta(s)$ can be described by a linear



(a) A sample figure

(b) Total curvature

Figure. 7 Symmetry with regard to rotation



(a) A line figure

(b) Total curvature

Figure. 8 Convexity

function of s. If a given figure L is composed of m pieces of smooth lines$(l_1, l_2, ..., l_m)$, it is approximated as

$$D(L) = min \sum_{i=1}^{m} \frac{1}{L_i} \int_{s_{i-1}}^{s_i} |\theta(s) - (a_i s + b_i)| ds \quad (11)$$

where

L_i : length of line l_i

s_i : length of the line from starting point to the end of line l_i

$\theta(s)$: total curvature of L

a_i, b_i : constants

Further, if a subfigure is composed of only circular arcs and straight lines, the figure is considered to be more regular. Therefore, regularity of subfigure L is defined as

$$f_4(L) = min(D_c(L), D_s(L)) \quad (12)$$

where $D_c(L)$ is the $D(L)$ when all the arcs of L is approximated by circular arcs, and $D_s(L)$ is the $D(L)$ approximated by straight lines. It can be said that the smaller f_4 of a subfigure is, the more regular it is.

3.5 Convexity

Convexity and concavity may have something to do with the selection of interpretations. Convex lines have a positive curvature, and concave lines have a negative curvature. Note one line may have both concave and convex segments in any order. Remember that the total curvature is originally the integral of curvature. Thus, decreasing O*(s)* values indicate that the subfigure is concave at these points. Note that there are two ways in which *&(s)* values decrease: continuous and discontinuous.

If O*(s)* of subfigure L is composed of m continuously decreasing parts and n discontinuously decreasing parts (Figure 8), then the convexity of L is defined as

$$f_5(L) = \sum_{i=1}^{m} |\theta(s_{i1}) - \theta(s_{i2})| + \sum_{i=1}^{n} |\gamma_i| \quad (13)$$

It can be said that the smaller f_5 of a subfigure is, the more convex it is.

3.6 Gestalt factors of an interpretation

Five Gestalt factors of a subfigure have been defined. In figure segregation, several subfigures are combined into an interpretation of the complex and ambiguous figure. Each subfigure has five Gestalt factors and by averaging the corresponding factors of the subfigures, an equivalent set of five Gestalt values can be created for each interpretation.

If interpretation I is composed of m subfigures, then the Gestalt values of each interpretation are defined as

$$F_i(I) = \frac{1}{m} \sum_{k=1}^{m} f_i(k)$$

$$i = 1, 2, 3, 4, 5 \quad (14)$$

where $f_i(k)$ is the Gestalt value of subfigure k.

4 Characteristics among the subfigures

The five Gestalt factors mentioned in section 3 are defined with each subfigure. However, it seems that there exist some characteristics that are defined among several subfigures [Shimaya and Yoroisawa, 1990b]. Therefore, we introduce two more characteristics.

4*1 Overlap

In figure segregation, it is felt that subfigures overlap each other or just connect with each other(Figure

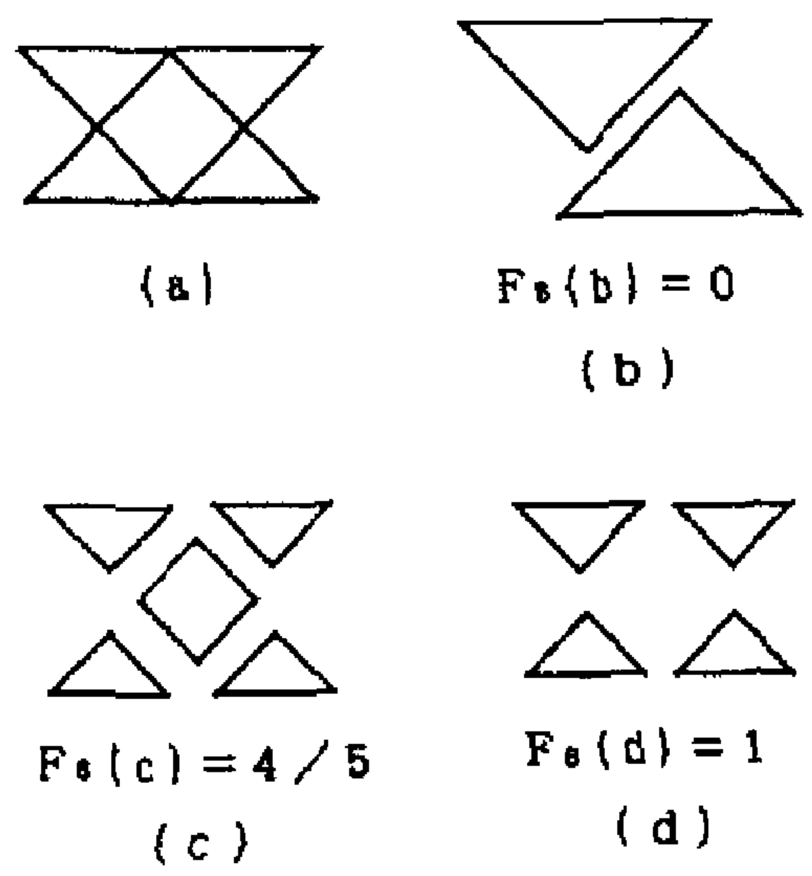


Figure. 9 Surrounding
(a) A sample figure
(b) - (d) Examples

9). This may have something to do with the selection frequency. Let interpretation I has n subfigures $C_1, C_2, ..., C_n$. then, the overlap of I is defined as

$$F_6(I) = \frac{1}{n}\sum_{i=1}^{n} f_6(C_i) \tag{15}$$

where
$f_6(C_i) = 0$: if C_i satisfies condition 1 or condition 2 but $f_6(C_i) = 1$: otherwise

condition 1: C_i does not share line elements in common with other subfigures and C_i overlaps other subfigures.

condition 2: all the line elements of C_i are inner lines (lines that are not the border lines between the original figure and the background), and all of them are shared with some other subfigures which do not include one another.

The value of F_6 is small when the original figure is interpreted as several subfigures that overlap(Figure 9(b)). If the original figure is interpreted otherwise, then F_6 becomes smaller if there exist subfigures that are surrounded by other subfigures(Figure 9(c),(d)).

4.2 Similarity

It seems that subfigures resembling each other may have something to do with figure segregation. Humans feel two figures look like each other when they are similar or when they are mirror images of each other.

Suppose $\theta_i(s)$ is the total curvature of subfigure C_i, and $\theta_j(s)$ is that of C_j, S_i is total length of C_i, and S_j is that of C_j. Let $S_i \leq S_j$.

If subfigures C_i and C_j are similar, then there exists a start point which satisfies the following equation.

$$\theta_i(s) = \theta_j^*(s)$$
$$0 \leq s \leq S_1 \tag{16}$$

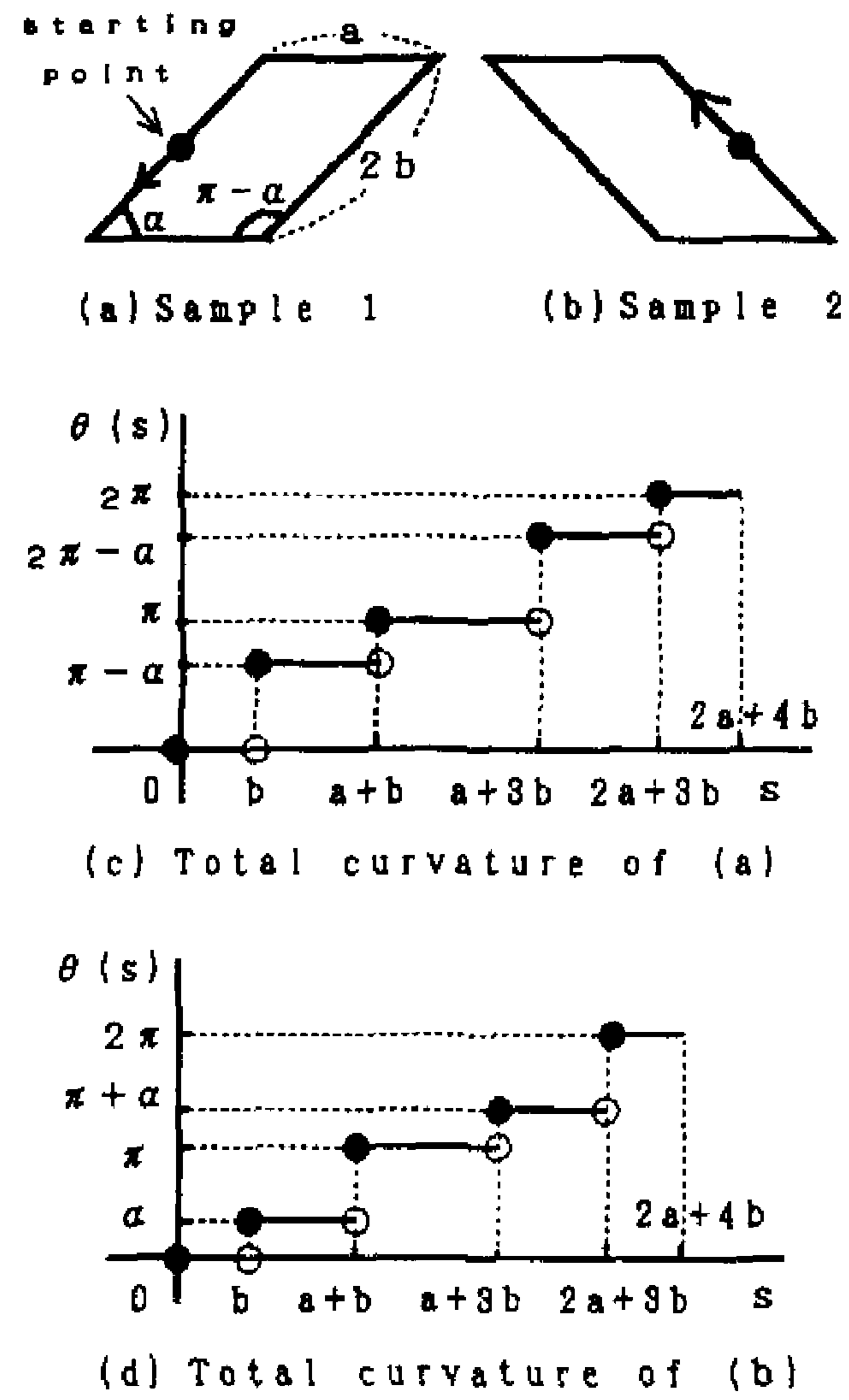


Figure. 10 Mirror Images

where

$$\theta_j^*(s) = \theta_j(s * \frac{S_i}{S_j}) \tag{17}$$

Thus, similarity of C_i and C_j can be defined as

$$T_s = min\frac{1}{S_i}\int_0^{S_i} |\theta_{ix}(s) - \theta_{jx}^*(s)|ds \tag{18}$$

where $\theta_{ix}(s), \theta_{jx}^*(s)$ is $\theta_i(s), \theta_j^*(s)$ when the starting point is x.

Mirror images are not similar, but humans think they look like each other. If C_i and C_j are mirror images of each other, then there exists a start point which satisfies the following equation(Figure 10).

$$\theta_i(s) + \theta_j^*(S_i - s) = \theta_i(S_i)$$
$$0 \leq s \leq S_1 \tag{19}$$

Thus, similarity of C_i and C_j concerning mirror image can be defined as

$$T_m = min\frac{1}{S_i}\int_0^{S_i} |\theta_{ix}(s) + \theta_{jx}^*(S_i - s) - \theta_i(S_i)|ds \tag{20}$$

Thus, similarity of C_i and C_j can be defined as

$$T_{ij} = min(T_s, T_m) \tag{21}$$

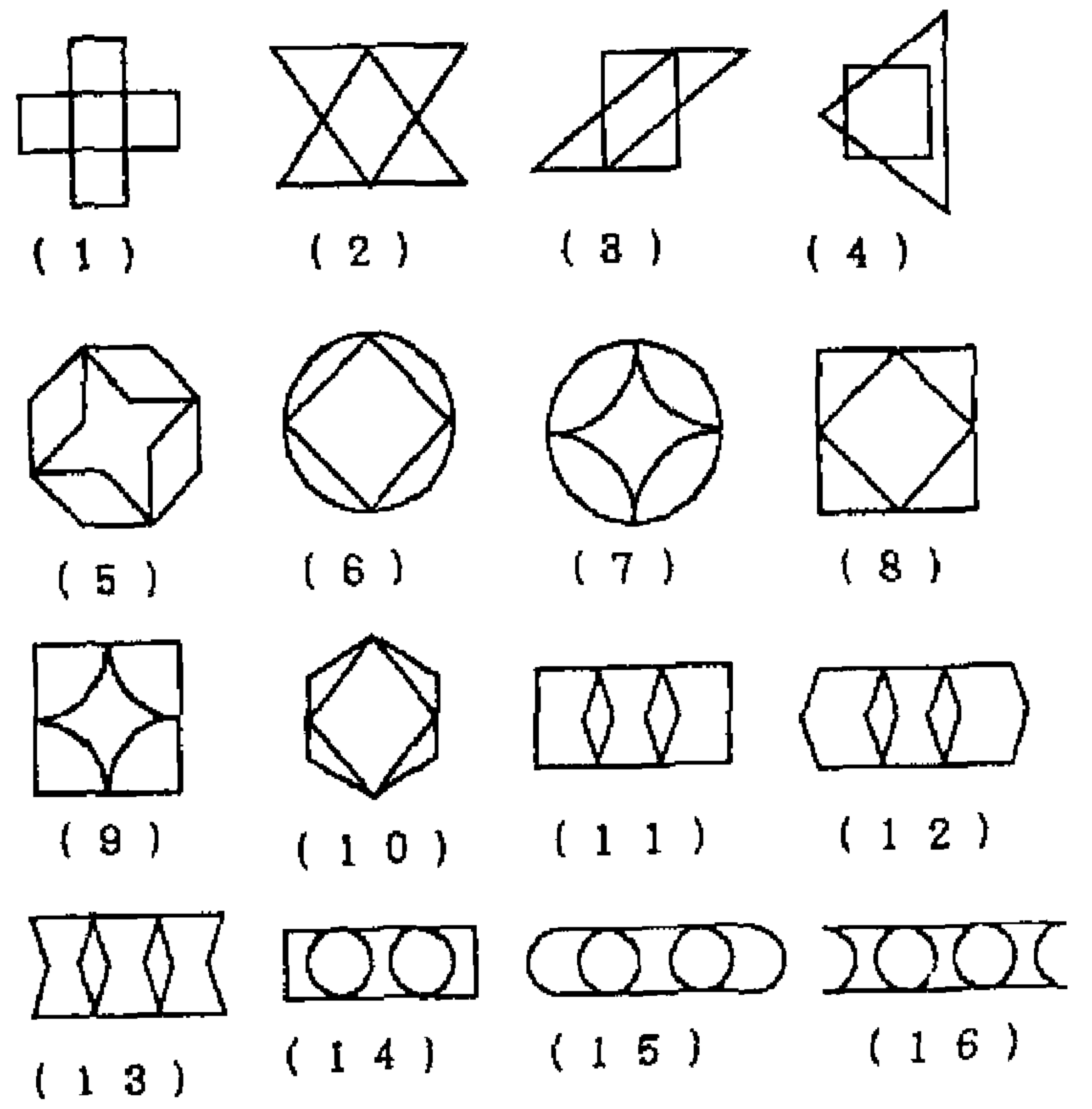


Figure 11 Sample figures

	F_1	F_2	F_3	F_4	F_5	F_6	F_7	G
F_1	1.00	0.50	0.28	0.16	0.21	0.71	-0.10	-0.68**
F_2		1.00	0.05	0.03	0.67	0.62	-0.03	-0.41**
F_3			1.00	0.29	-0.04	0.37	-0.42	-0.60**
F_4				1.00	0.02	0.16	-0.19	-0.29*
F_5					1.00	0.27	0.09	-0.14
F_6						1.00	-0.26	-0.66**
F_7							1.00	0.40**
G								1.00

* ···significant with 5% risk
**···significant with 0.5% risk

Table 1 Correlation coefficient

If interpretation I has k subfigures, then similarity of I can be defined as the average of T_{ij}.

$$F_7(I) = \frac{1}{nC2}\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} T_{ij} \tag{22}$$

It can be said that the smaller F_7 of an interpretation is, the more similar the subfigures are in it.

5 Significance of each Gestalt factor

In order *to* check the significance of the *seven Gestalt* factors, a psychological experiment was conducted. Sixteen ambiguous figures were shown to twenty subjects. Figure 11 shows the sample figures. Each subject was asked to draw the most natural interpretation for each sample figure. Fifty-six kinds of different interpretations were obtained. Gestalt values of each interpretation can be calculated as shown in sections 3 and 4.

The significance of each Gestalt factor is calculated from the correlation coefficient between the Gestalt values and the selection result as shown as Table 1. In Table 1, column G corresponds with the number of subjects who drew each interpretation.

Statistics give the limit of significance *JM*, by the next equation.

$$\gamma_M = \frac{e^{2T}-1}{e^{2T}+1} \tag{23}$$

where

$$T = \frac{t_{n-1}(\epsilon)}{\sqrt{n-3}} \tag{24}$$

n: number of samples
ϵ: risk
$t_{n-1}(\epsilon)$: t-distribution with $n-1$ degrees of freedom and with the risk of ϵ

Because the number of samples is 56, the limit of significance γ_M is as follows.

$$\gamma_M = 0.23(with 5\% risk)$$
$$\gamma_M = 0.32(with 1\% risk) \tag{25}$$
$$\gamma_M = 0.35(with 0.5\% risk)$$

The following results can be seen in Table 1.

(1) Simplicity F_1, symmetryF_3, and overlapF_6 are much more significant with 0.5% risk. Therefore, they are very important factors in figure segregation.

(2) Continuity F_2 and similarityF_7 are significant with 0.5% risk. Therefore, these two are also important factors.

(3) RegularityF_4 is significant with 5% risk-

(4) Is is generally believed that humans prefer convex figures rather than concave figures [Kanizsa, 1979]. However, according to this psychological experiment, convexity F_5 is not significant. It can be said that in figure segregation, even if the convexity seems significant, there exist other more significant factors.

6 Figure segregation estimation

Linear multiple regression analysis was conducted using Gestalt values as predictor variables and the number of the people who selected the interpretation as the criterion variable. By doing this, an equation which estimates the interpretation selection frequency is obtained as follows.

	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7
coefficient	11.36	-5.53	-1.98	-4.36	-1.24	0.75	-1.09	2.20
t-value		3.70	1.09	3.62	0.88	0.47	0.84	1.87

Table 2 Partial regression coefficient and t-value

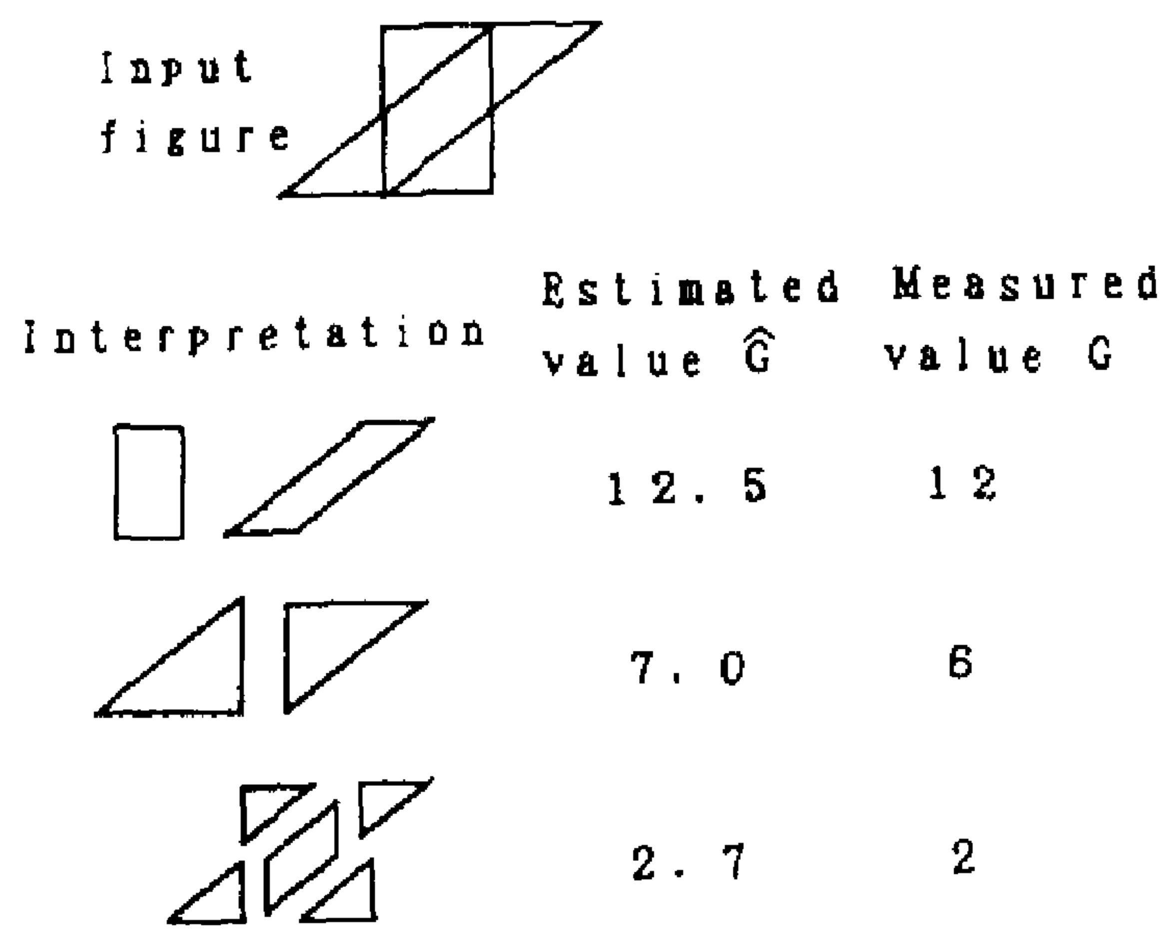


Figure 12 An estimation example

$$\hat{G}(I) = \beta_0 + \sum_{j=1}^{7} \beta_j * F_j(I)$$
$$j = 1, 2, ..., 7 \qquad (26)$$

where *G(I):* the criterion variable of interpretation *L*
BO: a constant,
BJ: the partial regression coefficient.
FJ: Gestalt values of interpretation I-

Table 2 shows the partial regression coefficients and their t-values. Figure 12 shows an example. *G* is the frequency selection value obtained by equation (26) and *G* is the number of people who actually selected the interpretation.

Multiple correlation *R* between *G* and G is,

$$R = 0.84 \qquad (27)$$

This indicates that equation (26) agrees fairly well with the experimental results. The characteristics of this model are as follows.

1) Ability of estimation

Figure segregation models have been reported, but they only indicate which interpretation is most likely to be selected. This model can completely estimate which interpretation is selected most and in addition to that, can estimate the selection frequency of each interpretation fairly well.

2) Robustness of the algorithm

The Gestalt factors shown in this paper are quite general and independent of sample figures. Also, if a new Gestalt factor is found, it is very easy to put the factor into this model. This easy modification character is very important when constructing a model of psychological effects because it is very difficult to make an initially complete model. This model is well constructed from this stand point.

7 Conclusion

It is very difficult to simulate human visual perception using computers. Figure segregation is one of the hardest problems to conquer. In order to achieve this, it is necessary to describe human visual characteristics in figure recognition. Therefore, we defined Gestalt factors such as symmetry and continuity with total curvature function. The significance of each factor was examined by a psychological experiment. Then, by applying linear multiple regression analysis, selection frequency of each interpretation was estimated. This model can simulate human visual perception in figure segregation very well. Also, it should be noted that the model proposed here and in [Shimaya and Yoroizawa, 1990a] is quite general and easy to apply to practical applications such as map recognition or technical parts recognition.

References

[Kanizsa, 1979] G, Kanizsa. *Organization in Vision: Essays on Gestalt Perception.* Praeger Publishers, 1979.

[Metzger, 1953] W. Metigex. *GESETZE DES SEHENS.* Verlag von Walderaar Kramer, Frankfurt am Main, 1953.

[Shimaya and Yoroizawa, 1988] A. Shimaya and 1. Yoroizawa. A figure segregation criterion based on human visual characteristics. *IEICE Technical Report,* PRU88-72, 1988. (in Japanese).

[Shimaya and Yoroizawa, 1990a] A. Shimaya and I. Yoroisawa. Automatic creation of reasonable interpretations for complex line figures. *Proc. of 10th International Conference on Pattern Recognition,* Vol. I:pages 480-484, 1990.

[Shimaya and Yoroizawa, 1990b] A, Shimaya and I. Yoroizawa. Quantitative analysis on figure segregation. *ITEJ Technical Report,* VAr90-18:pages 55-60, 1990. (in Japanese).

[Spoehr and Lehmkule, 1982] K. T. Spoehr and S. W. Lehmkule. Visual *Information Processing. W.*h. Freeman and Company, New York, 1982.

[Tuijl, 1980] H.F.J.M.van Tuijl. Perceptual interpretation of complex line patterns, *journal of Experimental Psychology: Human Perception and Performance,* vol.6, No.2, 1980.

[Uesaka and Tajima, 1976) Y. Uesaka and K- Tajima. An interpretative model of figure segregation. *Trans. of IEICE,* Vol.59, D-I, 1976, (in Japanese).