ISSN 0202-338. PSICHOLOGIJA. 1994. 13

MONOCULAR LOCATION PRIMITIVES USED TO EVALUATE HORIZONTAL DISPARITY IN LOCAL STEREOPSIS

V. Petrauskas, H. Vaitkevičius, V. Mildažis

ABSTRACT

Depth perception of two point-like stimuli presented dixoptically was studied as a function of luminance distribution within the pair of stimuli in each half-field. The dots subtended approximately 1 arc min and were separated by 3, 4, 5 or 6 arc min. The depth changed gradually unless the separation between the stimuli exceeded 4 arc min. The depth change can not be derived solely on the basis of disparity change. The results may be accounted for if assumed that disparity is evaluated by the difference in the weighted mean (centroid) of luminance distribution over corresponding left and right retinal regions.

INTRODUCTION

In normal vision the two eyes receive slightly different views of the world. This difference is measured by horizontal disparity and i interpreted by the visual system as visual depth. This kind of depth perception, i. e. the perception originating from horizontal disparity is called stereoscopic depth perception, or STEREOPSIS. One may encounter a number of ways to represent disparity, which can be placed into two broad categories, local representation and distributed representation (Lehky and Sejnowski, 1991). According to local representation, the value of location is indicated by which neurone (neurone group) is fired differently from the others (one example, of the criteria might be maximum response). The most known from of local representation is „local sings" encoding in psychophysics and has been used widely in the models of stereopsis (Nelson, 1975; Marr and Pogio, 1976). To cover the entire range of disparities in the local representation there must be a large number of such narrowly-tuned units. According to the second type of encoding, location is represented by the pattern of activity within a population of neural units (or channels) each of which might be broadly tuned to location. The most familiar example of this form of encoding is used in colour vision. Also some attempts were made to apply distributed representation approach to encode binocular disparity as well (Vaitkevicius, 1984; Vaitkevicius et al., 1984; Petrauskas, 1986; Lehky and Seinowski, 1991). Most of the above approaches might be thought of as „pure" approaches except the first one. Vaitkevicius (1984) proproses location encoding theory in which distributed representation is taken as a first step, from which then local representation is derived.

Whereas „local signs" might be thought of as monocular primitives to calculate disparity in the local representation, however, there is no explicit pointers to what monocular location primitives are used to evaluate disparity in distributed representation or in other words, there is not made any kind of explicit transition from monocular to binocular stage.

METHODS

Stimuli generation, experimental procedure and data handling were controlled by the IBM PC XT computer.

The stimuli in our experiments consisted of bright (up to 3 cd/m) points generated on the oscilloscope screen and presented against a dark background (Figure 1 A). Orthogonally-oriented polarizers placed in front of the oscilloscope screen and the subject's eyes guaranteed haploscopic image presentation. However, the computer can readily generate stereograms that define objects at numerous different vergence distances, but no convenient way has yet been discovered to concomitantly covary the stimuli to accommodation. To decrease accommodation-vergence disparities, the crossed presentation was used, that is the left eye was presented with stimuli on the right

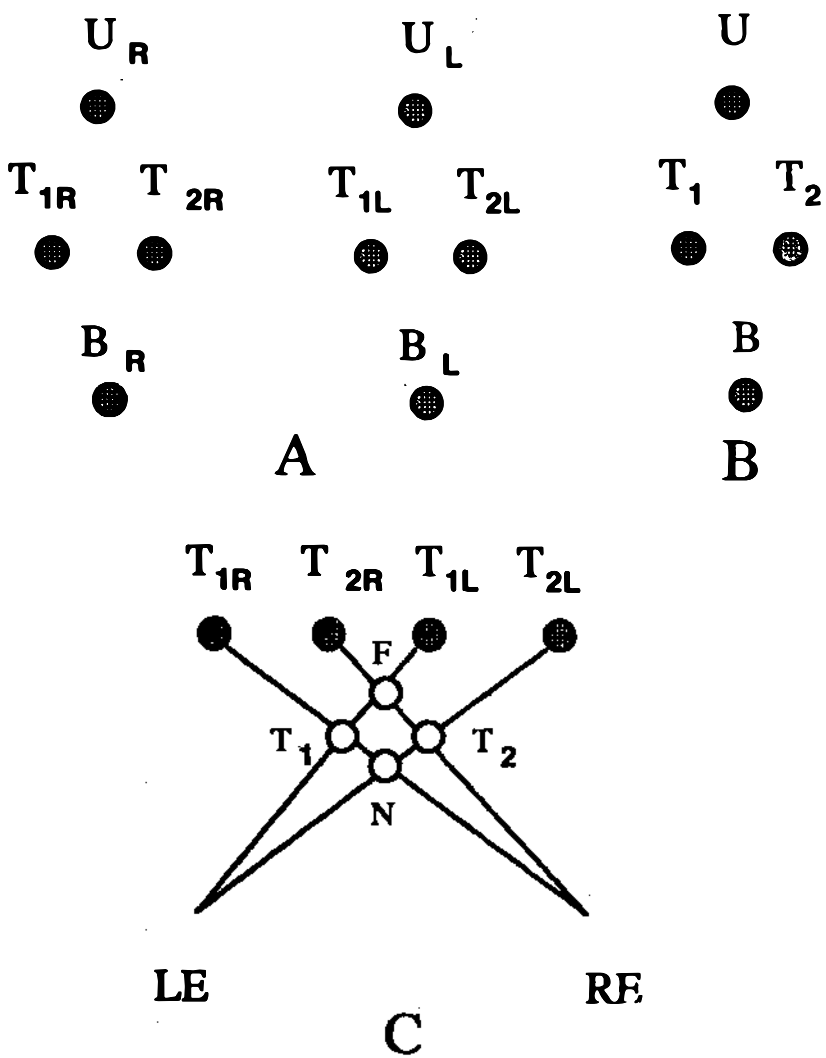

Fig. 1. Schematic representation of stimuli. A – front view of both half-fields. Index R stands for the right and L for the left half-fields; B – the percept of the A – part of this figure; C – the ilustration of fudion of the test stimuli; LE – stands for the left- and RE for the right eye optical centers respectiveiy.

side of the screen and vice versa. On the other hand, the separation between the left- and right eye targets did not exceed 8 arc min. Having in mind that visual system allows some tolerance for the accommodation and vergence alignment error, the negligible error in our experiments should not play a substantial role. Stimuli generation was performed using three 8-bit analogue-to-code converters to manipulate with X, Y, and Z channels of the oscilloscope. The viewing distance was 1.2 m. X and Y channels were adjusted to let dots be positioned to an accuracy of 5 arc sec. Target luminance was controlled by photometer. Each half-field consisted of one upper- (U_L, U_R) point, one bottom- (B_L, B_R) point, and, in the middle, of two horizontally displaced test-points (T_{1L}, T_{2L}, T_{1R}, T_{2R}). The vertical distance between test and upper (bottom) points was 12 arc/min. The horizontal distance between test-points was 3, 4, 5, 6 arc/min. For every value of horizontal distance the dependence of the test-points' depth on luminance contrast was examined. The luminance of either T_{1L} and T_{2R} or T_{2L} and T_{1R} was changed from 3 cd/m to zero while the luminance of other points was left constant. Six levels of luminance were used. Luminance contrast was evaluated by the following formula:

$$C = \frac{L(T_{2R}) - L(T_{1R})}{L(T_{2R}) + L(T_{1R})};$$

where letter L stands for the luminance of the stimulus denote in the parentheses. During an experiment the subject was asked to tell by pressing one of the two buttons whether the points U and B (Fig. 1 B) appear to be nearer or further than the testpoints, T_1 and T_2. For each luminance contrast condition the points, U and B were presented at 11 different disparities relative to the test points. Each disparity condition is presented 10 times in a random order. As a result, the number of responses „further" is plotted as the percentage of the total number of responses against each disparity condition, that is a psychometric function is generated. Ideally, this function will change from 0 to 100% as the disparity of points, U and B is changed from the most extreme crossed disparity to the most extreme uncrossed disparity of the tested set. To estimate the stimuli corresponding to the relevant response levels, acumulative normal curve was fitted to the data. The disparity associated with chance level of the response was interpreted as a measure of the perceived depth shift. The disparities associated with 20% and 80% „further" responses are interpreted as an indication of the range of uncertainty or spread.

The different experimental procedures were used for the short and long presentations. In the case of the long presentation subjects were allowed to inspect the stimulus until they were aware of their response. In the case of short presentation stimuli were presented for 200 ms. In the latter case a subjective method of vergence control was used. The control was performed using a stereogram presented in Fig. 2. The right vergence is achieved, when the subject reports that the vertical line seen only by the left eye (remember the condition of crossed presentation) fell in the middle of the horizontal line seen only by the right eye. The experimental procedure which uses vergence control consists of the following steps. The computer starts each presentation cycle by setting the vergence control condition. An observer looks at the screen until he perceives the vertical line falls in the middle of the horizontal line, then he presses a button to start stimuli presentation lasting for 200 ms. After this presentation, the observer responds whether he perceives points, U and B further or nearer than the test ones by pressing one of the two buttons. In both cases (long and short presentation) in one experimental session the perceived depth is determined for one luminance ratio condition. Each session consisted of 110 presentations following after the 10 training presentations which were not stored in the computer's memory. At the end of each session the computer calculated and plotted a psychometric curve from which the perceived depth and spread were determined. The latter two values were stored in the computer memory. After all luminance ratio conditions were explored a graph of the perceived depth as a function of luminance contrast is plotted. The data were obtained for three adult subjects.

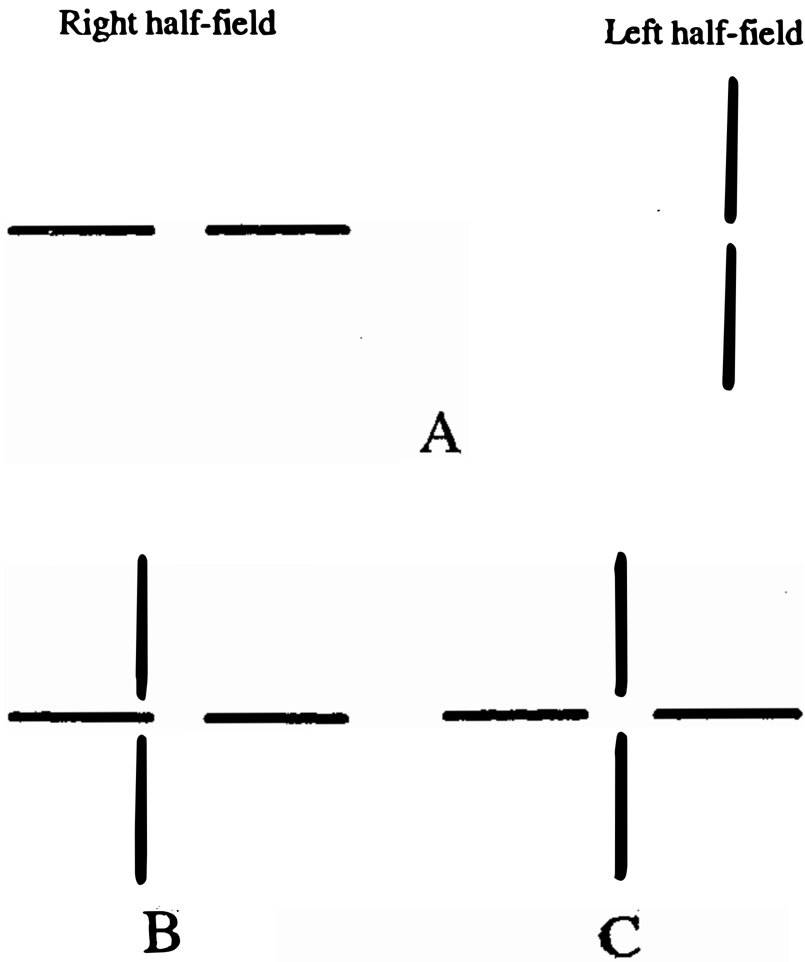

Fig. 2. Haploscopic method of adjusting vergebce. A – the test presented for the left (L) and for the ringt (R) eyes; B – stereoscopic percept when vergence is properly adjusted; C – stereoscopic percept of a maladjusted system.

RESULTS

Figures (1 B) and (1 C) show the subjects percept of the stimuli presented in figure (1 A). Figure (1 B) iliustrates the frontal view and figure (1 C) the view from above. In the basic luminance contrast (C=0), subjects perceived the two teststimuli, T_1 and T_2 (Fig. 1 C). At the extreme contrast conditions (C=–1 or C=1) subjects perceived respectively the stimuli, N or F. At the intermediate luminance contrast conditions the test-stimuli were perceived at some intermediate depth which depended on luminance contrast. The results for the observers MK, VR, AS obtained using long presentation are shown on Fig. 3, 4, 5, and using short presentation (for the observers AS, MK) in Fig. 6, 7. Different parts of the figures present results for the target (T_{1R}, T_{2R} and T_{1L}, T_{2L}) separation of 3, 4, 5, and 6 arc min (indicated in each graph). The comparison of Fig. 3, 4, 5 with Fig. 6, 7 indicates that the perceived depth is not affected substantially by the procedure (long or short presentation). Perhaps the only evident difference is the larger spread in the case of short presentation. On the other hand, however, an examination of the data obtained for different target separation indicates, that target-separation is an important factor for the perceived depth change. While the target-separation does not exceed 4 arc min, the perceived depth change can be thought of as continues. On the other hand, for target-separations of 5 and 6 arc min the transition of the perceived depth from continues to abrupt is obvious.

It needs stressing that though no qualitative difference obtained between the results by all three observers, quantitative variations are prominent (compare Fig. 3, 4, 5). As it is evident from these figures, the critical target-separation at which the transition from the continuous to the abrupt (step-wise) depth change occurs, varies from 4 arc min (observer VR) to 5 (observer AS). The difference of the inter subject results may be related to the subject visual acuity. For example subject MK and AS wore glasses, but only for subject MK vision was corrected to normal, whereas subject's AS vision was corrected to 80%. On the other hand, subject VR has perfect visual acuity of 120% when compared to normal. From this comparison an

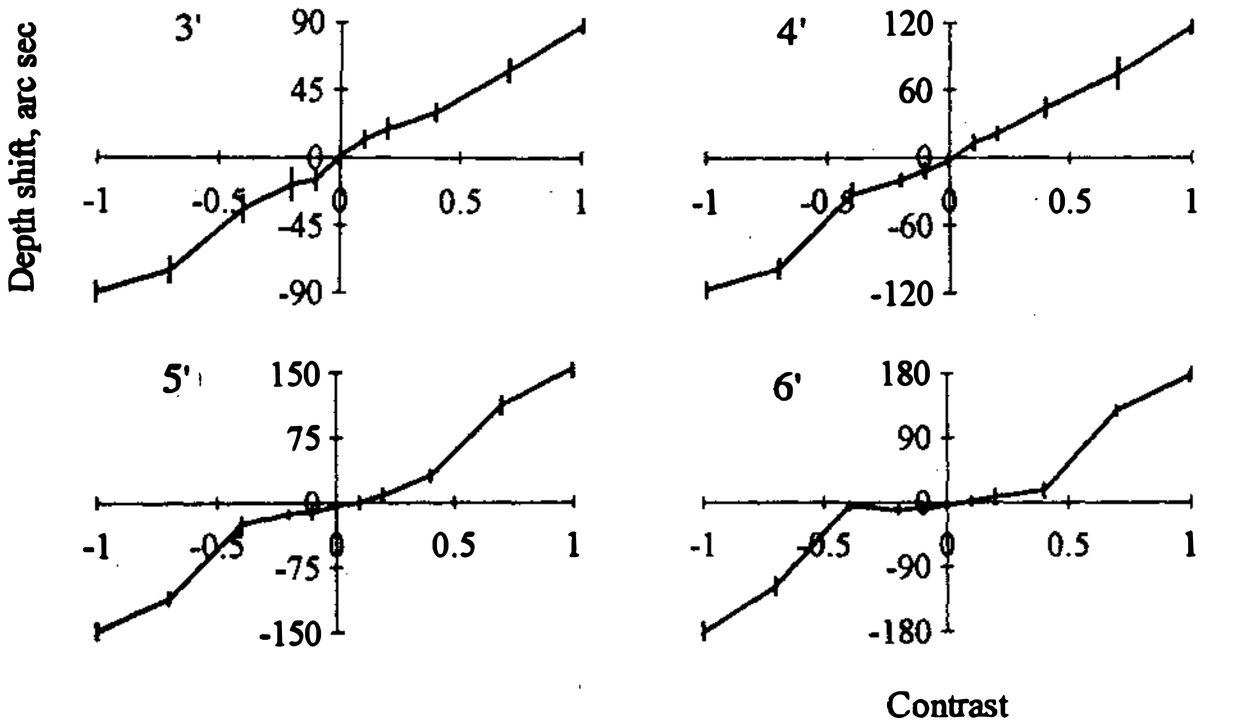

Fig. 3. Perceived depth shift (measured by disparity) of test stimuli T_1 and T_2 (shown in Fig. 1 B) as a function of the luminance contrast of test points presented in one half field. Data obtained with long presentation for observer MK. Different graphs are for different test target separations (see numbers inside each graph):

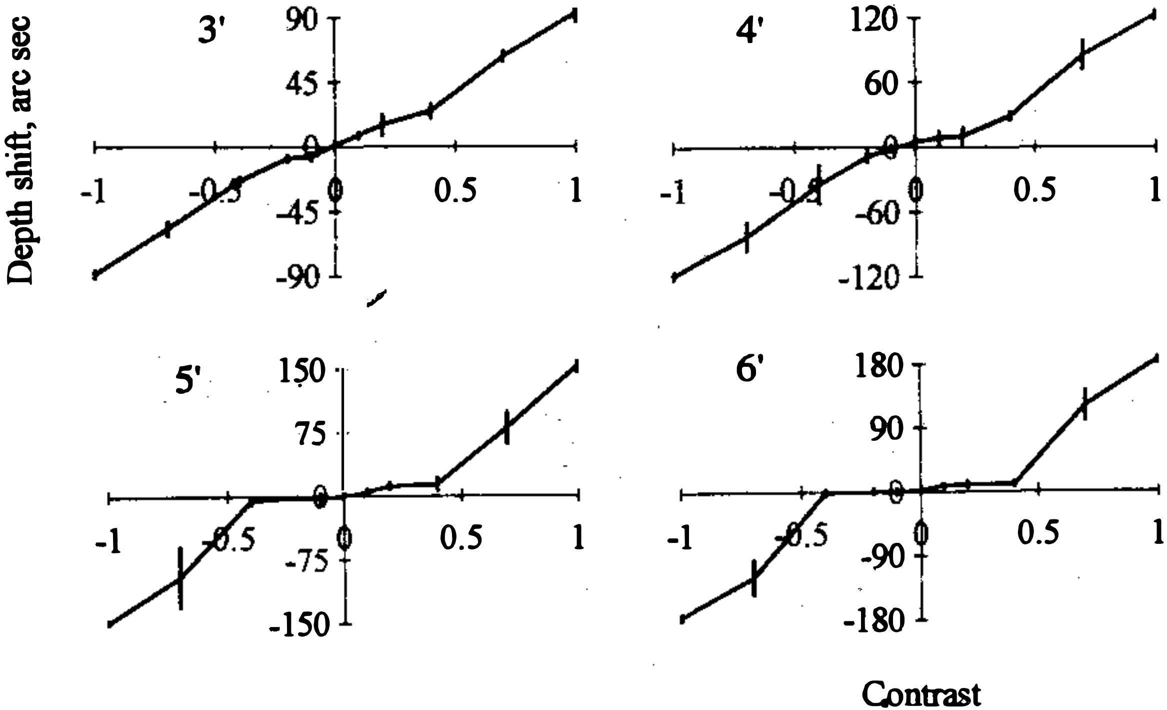

Fig. 4. The same as in Fig. 3 except that the data were obtained for observer VR.

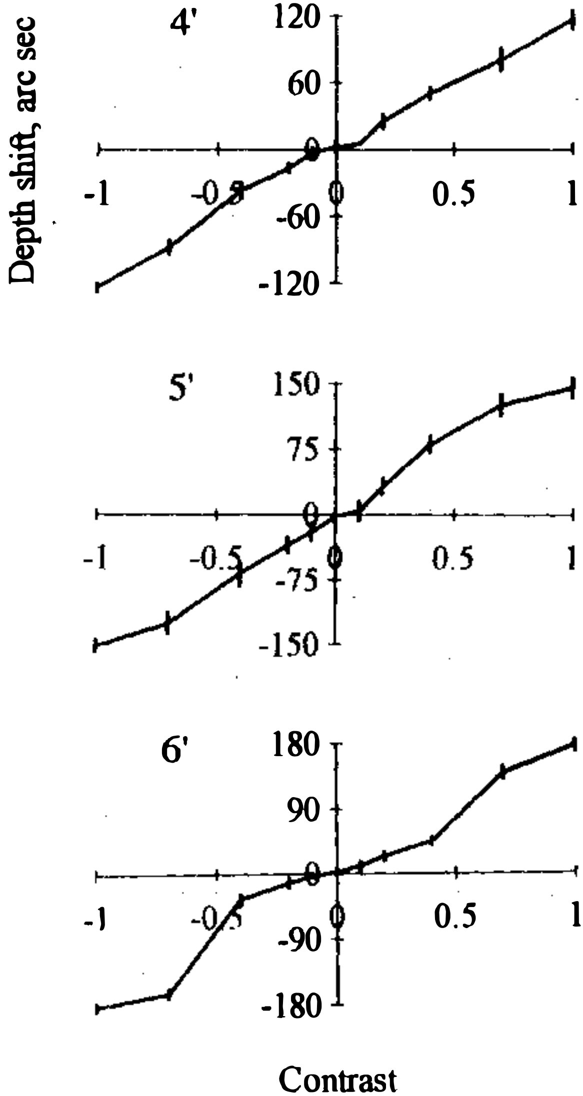

Fig. 5. The same as in Fig. 3 except that the data were obtained for observer AS.

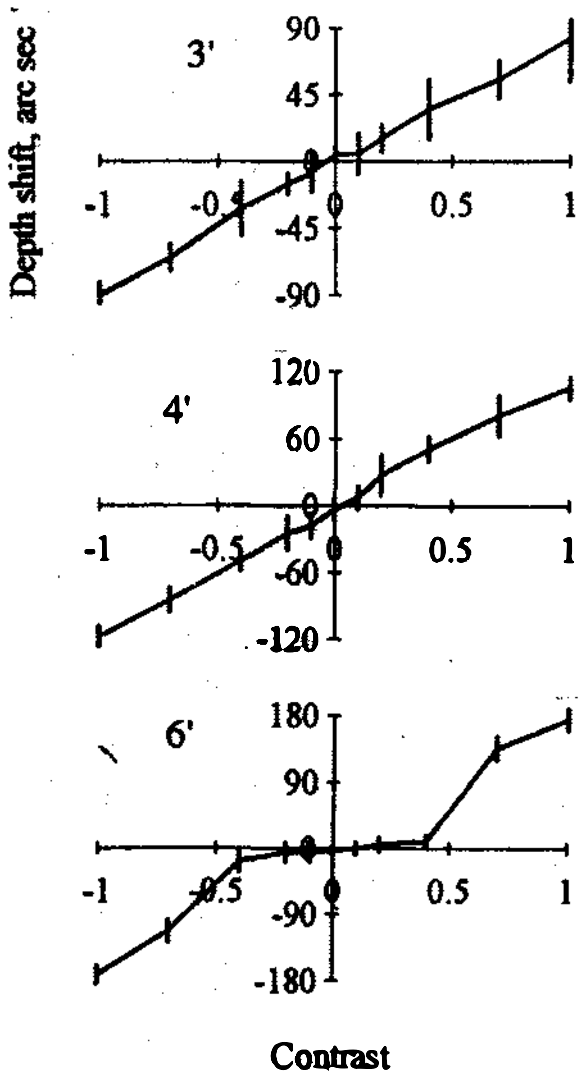

Fig. 6. Perceived depth shift (measured by disparity) of test stimuli T_1 and T_2 (shown in Fig. 1 B) as a function of the luminance contrast of the points presented in one half field. Data obtained with short presentation for observer MK. Different graphs are for different test target separations (see numbers inside each graph):

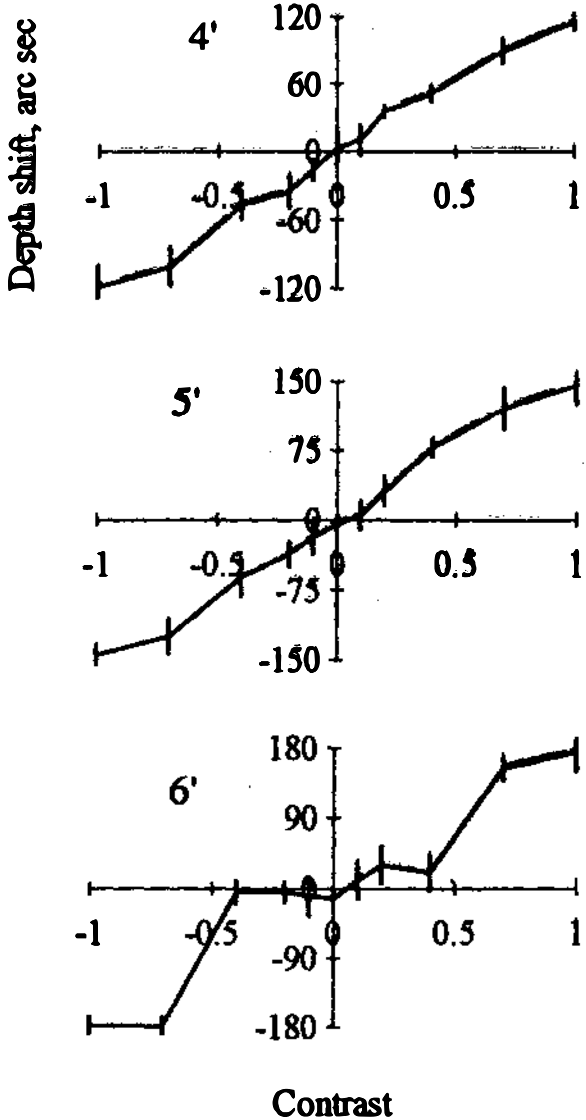

Fig. 7. The same as in Fig. 6 except that data are obtained for observer AS.

approximate conclusion follows: the better visual acuity, the lesser target-separation is needed for the transition from the continuous to abrupt changes in the perceived depth with variation in relative luminance ratio (contrast). This conclusion, however, should be tested for more observers to become a universal.

DISCUSSION

Kaufman et al. (1973) were first to show continuous changes in the perceived depth caused by variation in relative luminance of superimposed random-dot stereograms. The latter studies (Foley, 1976; Foley and Richards 1978) were intended to demonstrate the existence of depth mixture over much larger disparity (up to 2 arc deg) differences using vertical slit stimuli. Contrary to Foley (1976), Foley and Richards (1978), Krol (1982) has found only a slight dependence of perceived depth on luminance ratio with the target-separation of 15 arc min. However, while Kaufman's (1973) data show continuous perceived depth change with luminance ratio, the data presented in Foley (1976), Foley and Richards (1978) could be hardly treated as an indication of the continuous change. The results reported by Foley (1976), Foley and Richards (1978), Krol (1982) indicate that the perceived depth changes rapidly only at the extreme luminance ratio conditions. These results are comparable with our findings obtained for target separation of 5 and 6 arc min. Such results were interpreted by Krol (1982) as a brake-down of normal stereopsis. And a slight perceived depth change at relative luminance ratio conditions, where target luminance did not differ very much, according to Krol, may be interpreted as an artefact of eye vergence. However, our results obtained for the separation between test-stimuli of 3–4 arc min might be considered as an indication of continues depth change as opposed to the step-like change when the separation was larger. In addition, the comparison of the results obtained for the condition when eye vergence was not controlled (fig. 3, 4, 5) and for the condition when it was controlled (Fig. 6, 7), seem not to support the idea of the perceived depth change being an artefact of eye vergence thus allowing us to conclude the gradual depth change obtained in our experiments being a consequence of some kind of spatial processing taking place at some still not know stages of the visual system.

As our data show there exists a remarkable difference in the performance when target-separation is 3–4 arc min and when it is 5–6 arc min. This difference may be interpreted by different mechanisms involved. Most likely that step-like change in the perceived depth can be explained in the foolowing way (Fig. 1 C). In the basic luminance contrast condition the subjects perceive two test-stimuli, T_1 and T_2. When contrast is of +1 (only T_{1R} and T_{2L} are present) the subjects perceive target F (Far), and when contrast is of –1 (only T_{2R} and T_{1L} are present) the subjects perceive target N (Near). Hence, the perceived depth change is possibly due to the change in the stereo pair matched. Such an explanation could be applied in principal to all target separations. However, it is not clear how the change in fusion alone could account for the continuous change in the perceived depth at intermediate values of relative luminance contrast when target-separation does not exceed 3–4 arc min.

At least two hypothesis can be entertained for the gradual disparity change obtained here. The change could be a consequence of processing taking place at purely monocular stage or during the actual establishment of stereo correspondence.

Badcock and Westheimer (1985) found that flanks influence the localization of test-targets in a Vernier acuity task. For the flanks within a zone extending approximately 3–4 arc min attraction takes place. As could be seen from Fig. 1 A variation of luminance was performed in both left and right half-fields of the stereogram. In this view, the change of the perceived depth could be a consequence of the change in perceived location due to the attraction thus resulting in processing taking place at a purely monocular stage. However, this kind of change reported in the above study is too small (about 5 arc sec) to account alone for the continues depth change within the range from –4 to +4 arc min observed in our experiments.

On the other hand, our results might be interpreted as disparity averaging of targets F and N (Fig. 1 C). However such an interpretation is hardly possible because of differences in the disparity range explored in our experiments (minimal disparity of points F and N (Fig. 1 C) was ±3 arc min) and the disparity range where disparity averaging takes place (Shumer, 1979; Parker and Yang, 1989). They have shown that random-dot stereogram with two disparity values may be perceived as acontinuous surface with an intermediate disparity value if the disparities range from about 80 arc sec at the fixation to as much as 240 arc sec at the average disparity offset from the fixation plane.

Although it is hard to attribute our findings to one of the above phenomena one point is stil common to the above phenomena and to our findings. That is the spatial zone of 3–4 arc min over which both monocular (Badcock and Westheimer, 1985) and stereoscopic targets (Westheimer, 1986; Westheimer and Levi, 1987) pool their location signals in a facilitatory way. In addition, the spatial zone over which depth averaging (Tyler, 1974; Parker and Yang, 1989) and transition from continuous to abrupt changes in depth perception found in our experiments is also about 3–4 arc min. These facts allow, to or opinion, some speculations about the disparity processing mechanism.

As it is known, binocular disparity between any two point-like stimuli is the difference in visual angles between them when seen by left and right eye. Such a definition of binocular disparity might be useful to describe depth of point-like stimuli and/or points on the surfaces. To use the same definition to assess disparity to the whole surface we should be able to calculate disparity of separate points on that surface. The problem we meet is the neural implementation of a mechanism which would be able to do such a job. Given the understanding of neurones and their receptive fields, it is more likely thar disparity is assessed to finite regions of the visual field instead of separate points. If that is true, the next question to be answered is what dimensions of such a region, i. e. of the region over which disparity information is integrated. The most obvious candidate findings to determine the dimensions of such a region might be stereo acuity of 22 c/deg found by Westheimer and Mckee (1980) along with the above mentioned facts about spatial zone of 3–4 arc min. Given that, we posulate, that the visual system assigns binocular disparity to overlapping spatial zones of approximately 3 arc min.

The next question we would like to proceed with is related to possible mechanisms performing disparity processing. We start from the question what are location primitives in corresponding left and right-eye spatial zones feeding binocular spatial zone to which disparity is assigned. One of the candidate location primitives might be that proposed by Westheimer and McKee (1977) according to which location is assigned to the centroid in luminance over a spatial zone of a few arc min. Extending this idea to the disparity processing we get local mechanism for the processing of disparity between any two small regions in visual field: the disparity of two small regions of any surface is a difference in centroid (or any other simmilar integral parameter) over corresponding regions projected to the left and right eye retinas.

With respect to the two discussed above groups of models it is not clear how local representation models might be fitted to perform centroid processing. Quite the opposite is with distributed representation approach. Not going in to details of what kind of centroid (linear or weighted) is to be used, in all cases it's processing involves responses of serveral units. This means that centroid-like processing is one of the concrete forms of distributed representation. As one of the first attempts to use distributed representation approach to model stereo vision is the model proposed by Vaitkevicius et al. (1984). The basic idea behind this mechanism is that the retinal image could be represented by the outputs of four filters defining weighted centroids and by the weighted energy over the overlapping areas in retinal image. The other attempt to use distributed representation approach in finding monocular location primitives under conditions of normal noise were derived using the maximum likelihood principle (Petrauskas and Stankevicius, 1990). The derived units may successfully be used to process disparity.

Finally, it seems possible to integrate psychophysical mechanisms under study with physiological ones. For example, the adjacent simple striate cells with odd- and even- symmetric weighting functions (Kulikowski and Marcelja, 1982) could be in principle fitted to perform a weighted centroid calculation in the above discussed models.

CONCLUSIONS

1. Findings of Foley (1976), Foley and Richards (1978), Krol (1982) could not be interpreted as depth mixture (or averaging) of stimuli F and N (Fig. 1 C).

2. The change in the perceived depth with variation in relative luminance is gradual when target separation is less than 4 arc min (Fig. 1 C), and steep-wise when it exceeds 4 arc min. To account for the differences at least two mechanisms could be postulated.

3. The gradual depth change obtained for target separation less than 4 arc sec could be understood if weighted centroids evaluated over the overlapping retinal areas with linear dimensions of about 3–4 arc min used as spatial primitives to which disparity is assigned, whereas the step-wise depth change is perhaps due to the change in fusion.

MONOKULINIAI POŽYMIAI, NAUDOJAMI REGOS SISTEMOJE, ĮVERTINANT BINOKULINĮ DISPARATIŠKUMĄ

V. Petrauskas, H. Vaitkevičius, V. Mildažis,

Reziumė

Buvo tiriamas dviejų taškų gylio suvokimas, keičiant jų ryškumo santykį monokulinėse dichoptinio vaizdo dalyse. Gylis keitėsi tolydžiai, jei atstumas tarp taškų neviršijo 4 regimųjų minučių. Rezultatus galima paaiškinti remiantis prielaida, jog regos sistemoje skaičiuojamas ryškumų pasiskirstymo atitinkančiose kairės ir dešinės tinklainių srityse disparatiškumas.