Geometric mean of auditory intervals assimilates visual apparent motion

Lihan Chen, Zhuanghua Shi, Hermann J. Müller, Xiaolin Zhou

Abstract

It has been shown our brain can extract statistical information from a group of objects or a sequence of events efficiently for decision-making. However, such statistical information could also introduce bias in judgment of individual samples or events, previously known as the central tendency effect. Here we examined how the mean and variance of the statistical information in a sequence of auditory temporal events affects the motion percept engendered by the imbedded 'Ternus-type' visual apparent motion. Even though the auditory intervals were entirely task-irrelevant and could so be ignored, we found the mean of the auditory interval to assimilate the subsequent interval between the Ternus visual display frames, causing a systematic bias of motion percept. Further we distinguished influences of the first-order mean and second-order variance on crossmodal assimilation, and found the main cause was the geometric mean, rather than the arithmetic mean.

Keywords: temporal averaging, central regression effect, audiovisual interaction

Introduction

Humans and some animals are endowed with the ability to process complex sensory events using rapid ensemble coding (perceptual averaging). Our visual system can accurately extract the mean visuo-spatial properties - in terms of number, spatial layout, and even emotion - for a group of simultaneously presented visual objects (e.g., faces) (Alvarez 2011; Chong and Treisman 2005). Similarly, the auditory system can process the mean pitch of sequentially presented auditory beeps, and employ their mean rhythm to facilitate the detection of a subsequently presented visual target by means of auditory entrainment (**???**).

-omitted-

In the study (as depicted in Fig. 1-C), the visual Ternus frames were preceded, (synchronously) accompanied, and followed by auditory beeps; the intervals between successive preceding beeps could be fixed or irregular, but with their arithmetic mean being either 70 ms shorter, equal to, or 70 ms longer than the transition threshold (for perceiving equal probabilities of 'element motion' and 'group motion') between the element and group motion percepts.

Methods

-omitted-

Results

1. Regular and irregular intervals assimilates visual apparent motion

In a pre-test session, for each individual, we measured PSE of Ternus apparent motion, constructed by two visual Ternus frames paired with two synchronous beeps (hereafter referred as to the 'audiovisual' Ternus frames).Following the pre-test session, the audiovisual Ternus frames were embedded in an auditory train containing 8-10 beeps (including those accompanying the (audio-) visual Ternus frames) with a fixed inter-stimulus interval, which could be 70 ms shorter, equal to or 70 ms longer than the previously estimated

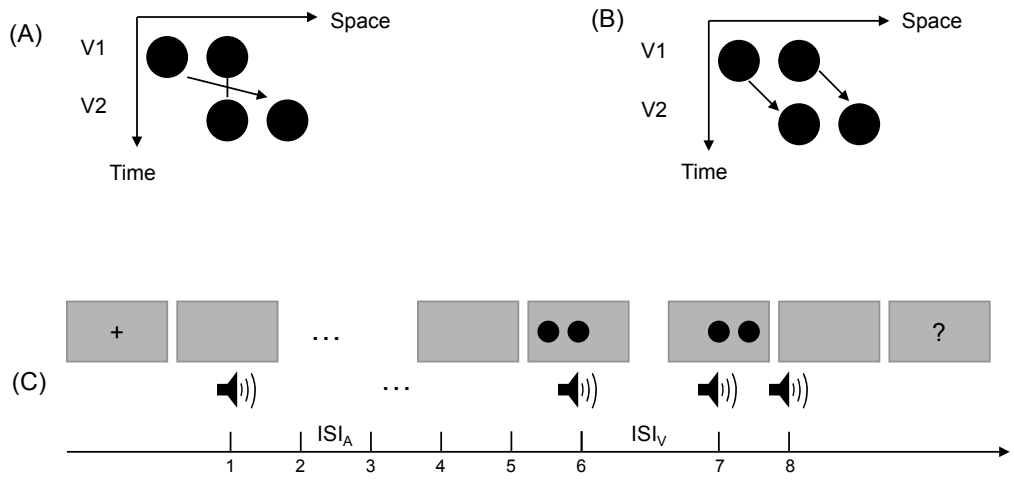

Figure 1: Fig.1. Ternus display and stimulus configurations. Two alternative motion percepts of the Ternus display: (A): 'element' motion for short ISIs. (B) 'group' motion for long ISIs. (C) Schematic illustration of the stimulus configurations used in the experiments. The auditory sequence contained 8-10 beeps. Two of the beeps were synchronously paired with the two visual Ternus frames, with an $ISI_{V(isual)}$ that varied from 50 to 230 ms. The other auditory ISIs ($ISI_{A(ditory)}$) were varied across different experiments.

transition threshold. The audiovisual Ternus frames were inserted at the sixth and seventh positions in 75% of all trials (the remaining trials were catch trials). Figure 2 depicts the average psychometric functions of Ternus apparent motion and the corresponding PSEs (151.2±6.2, 136.0±5.0, 127.7±5.2, 134.6±3.2 ms,) for short-, equal-, long-interval auditory sequences and for a baseline condition (implemented in the pre-test session). Auditory sequences with a long mean auditory interval, compared with a short mean auditory interval, elicited more reports of 'group motion' (indicated by the smaller PSE), $F(3,60)=10.23$, $p<0.001$, Bonferroni-corrected comparisons among all the conditions ('short', 'equal', 'long' and 'baseline') showed PSEs among the four experimental conditions were significant, all ps<0.05, all 0.330 (Fig. 2), but no difference between the 'equal'- and 'baseline', between 'long' and 'baseline', ps=1.

In Experiment 2, we varied the irregularity of the auditory sequence while keeping the mean intervals (comparable to the mean intervals in Experiment 1) at 70 ms shorter, equal to, or 70 ms longer than the pre-test transition threshold. The PSEs obtained were 155.1±7.7, 143.3±5.7, 137.9±5.3, and 135.3±4.5 ms for the auditory-sequence conditions with 'short', 'equal', and 'long' intervals and the pre-test baseline (paired audiovisual Ternus), $F(3,63)=8.190$, $p<0.001$, $=0.281$. Bonferroni- corrected comparisons showed that, relative to the 'equal' condition, the PSE was significantly larger in the 'short' condition, $p<0.01$, $=0.304$, and significantly smaller in the 'long' condition, $p<0.05$, $=0.264$. In addition, there were no differences between the 'equal' and 'baseline' ($p=0.611$), and the 'long' and 'baseline' ($p=1$) conditions. The fact that a crossmodal assimilation effect was still obtained with the irregular auditory sequence thus rules out a general effect of auditory entrainment.

...

Discussion

-omitted-

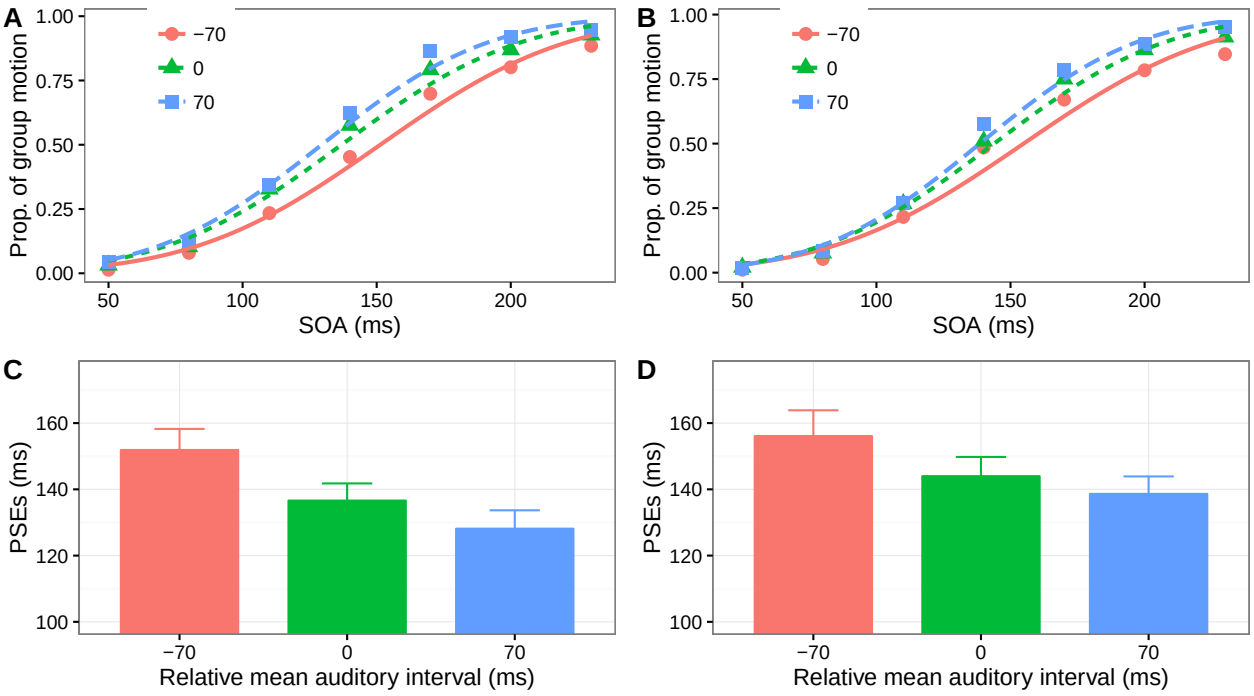

Figure 2: Psychometric curves and Mean PSEs for regular and irregular sequences

In summary, we revealed a new cross-modal perceptual averaging effect. The ensemble mean, in terms of geometric averaging, of the auditory sequence assimilates the visual interval, subsequently influencing perceptual decisions directed to the target event: visual apparent motion. The crossmodal assimilation has similar effects for regular and irregular auditory sequences, but becomes more marked when the target interval is embedded downstream of the auditory sequence, as the trailing interval is subjectively perceived to be dilated when more preceding auditory intervals can be sampled.

Appendix

	Effect	DFn	DFd	F	p	p<.05	ges
2	mIntv	2.000	40.000	19.092	0.000	*	0.134

Table 1: ANOVA for Regular sequence

	Effect	DFn	DFd	F	p	p<.05	ges
2	mIntv	2.000	42.000	13.094	0.000	*	0.061

Table 2: ANOVA for Irregular sequence

References

Alvarez, George A. 2011. "Representing multiple objects as an ensemble enhances visual cognition." *Trends in Cognitive Sciences* 15 (3): 122–31. doi:10.1016/j.tics.2011.01.003.

Chong, Sang Chul, and Anne Treisman. 2005. "Statistical processing: computing the average size in perceptual groups." *Vision Research* 45 (7): 891–900. doi:10.1016/j.visres.2004.10.004.