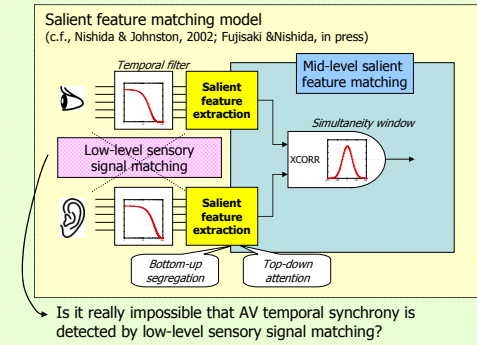


Audiovisual (AV) temporal synchrony perception: Which level?



Is AV temporal synchrony perception a low-level process or a higher level process?

□ Motion/stereo-depth perception with dense random-dot stimuli has been regarded as the evidence of low-level (pre-attentive) signal matching without explicit identification of matching features (e.g., Julesz, 1971; Braddick, 1974).

	Motion perception for random dot kinematogram (1 st order motion)	Motion perception based on feature tracking (3 rd order motion*)
Level	Low level	Mid level
Matching representation	Raw sensory signals (luminance contrast)	Salient features
Matching mechanism	Specialized sensor	General purpose mechanism

* c.f. Lu & Sperling (2001)

□ AV temporal synchrony can be correctly perceived for analogous dense random stimuli, if it is detected pre-attentively without identification of matching features.

→ Test AV temporal synchrony perception by means of synchrony- asynchrony discrimination with a temporally dense random pulse train made of luminance and sound-amplitude modulations

Expt. 1: Dense random pulse trains

Method

□ **Visual stimulus:**

- Luminance-modulated Gaussian blob (0-43 cd/m² increment from 21.5 cd/m² background)

□ **Auditory stimulus:**

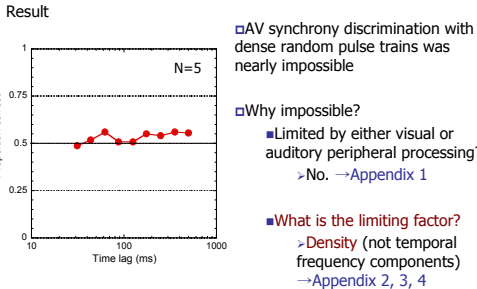
- 100% amplitude-modulated white noises (about 54 dB SPL at the peak of modulation)

□ **Modulation waveforms:**

- 2 sec random pulse trains (Refresh rate: 160 Hz, Pulse prob.: 1/2, Pulse density: ~80/s)

□ **Task:**

- AV temporal synchrony-asynchrony discrimination
- Feedback given after each response



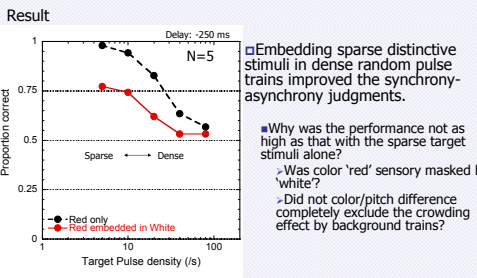
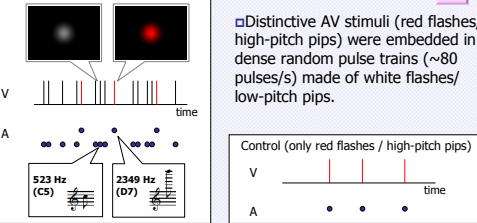
Suggests AV temporal synchrony perception is a mid-level feature-matching process

Expt. 2: Sparse in dense random pulse trains

Is AV temporal synchrony perception really a salient feature-based process?

■ Alternative hypothesis: direct comparisons of all AV signals

□ If feature based, embedding sparse temporal features in a dense random pulse trains may help performing the task



Density limit was not determined by the signal itself, but by the density of salient temporal features.

→ AV synchrony perception is a feature-matching process

Expt. 3a: Color-alternating random pulse trains

In Expt. 2, temporal features could be made salient either by stimulus-driven processes or top-down attention.

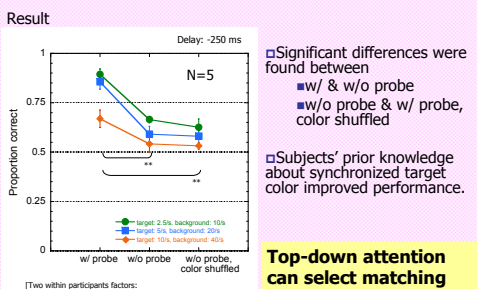
Can the temporal features be defined solely by top-down attention?

Test with the stimuli in which salient features cannot be segregated from a background by bottom-up segmentation

a) Pre-cueing what color to attend by a probe

b) No probe

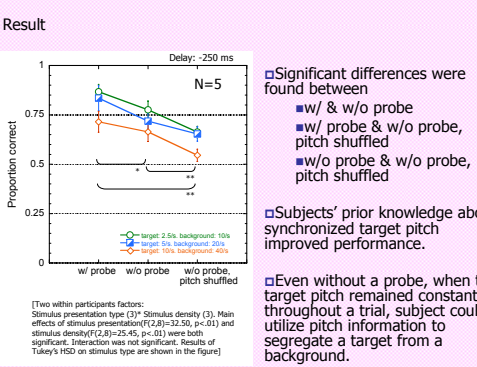
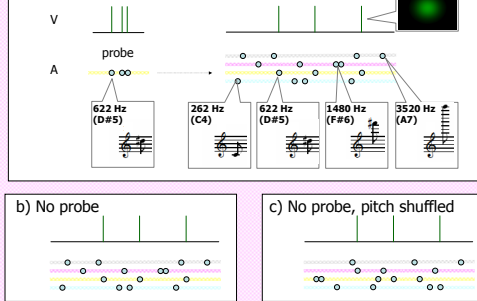
c) No probe & color shuffled



Top-down attention can select matching features

Expt. 3b: Pitch-alternating random pulse trains

a) Pre-cueing which pitch to attend by a probe



Summary

□ AV temporal synchrony perception, tested by means of synchrony-asynchrony discrimination, was nearly impossible with temporally dense random pulse trains.

□ The synchrony judgment was considerably improved when the stimulus included temporally sparse distinctive stimuli embedded in a dense pulse train.

□ Prior knowledge about a matching attribute improved the synchrony judgment.

→ These results suggest AV temporal synchrony perception is based on temporally sparse features that are made salient by stimulus-driven processes or top-down attention.

Discussion

□ Overall results support the salient feature extraction model

- AV temporal synchrony perception is a mid-level feature matching process (not a low-level, raw sensory signal matching process).
- Comparisons based on salient temporal features
 - Beneficial to avoid false matching, combinational explosion
 - Compare V and A signals with a common language (=salience)
- Temporal limit = salient feature extraction limit?

References

- Braddick O. (1974). A short-range process in apparent motion. *Vision Res.* 1974;14(7):519-27.
- Fujisaki, W. & Nishida, S. (in press). Temporal frequency characteristics of synchrony-asynchrony discrimination of audiovisual signals. *Experimental Brain Research*.
- Julesz, B. (1971). Foundations of cyclopean perception. Chicago: University of Chicago Press.
- Lu, ZL, Sperling G. (2001). Three-systems theory of human visual motion perception: review and update. *J Opt Soc Am A Opt Image Sci Vis.* Sep;19(9):2333-70.
- Nishida, S. & Johnston A. (2002). Marker correspondence, not processing latency, determines temporal binding of visual attributes. *Curr Biol* 12, 359-68.

Appendix 1: Dense random pulse trains (within modality)

W: Left half flicker vs. right half flicker (Gap 0.15 deg)

AA: Left ear flutter vs. right ear flutter (carriers were uncorrelated)

Density: Pulse density: ~80/s

Time lag: -250 ms

■ Almost perfect performance in both W and AA conditions

■ Indicates result of Expt. 1 was not due to the limit of within-modal peripheral processing

Appendix 2: High-cut random pulse trains

Cutoff: 4Hz, 16Hz

■ Removing high TF modulations from the dense random pulse train made AV synchrony discrimination easier.

■ Indicates that the dense random pulse trains contained a certain amount of detectable low-frequency components; however, those components were somehow masked by high-frequency components and therefore could not be extracted from dense random pulse trains.

Appendix 3: Sparse random pulse trains

Density: ~5/s, ~20/s

■ Reducing pulse densities also made the task easier.

■ Suggests the individuation or isolation of temporal features is a key factor.

■ Not 'frequency' but 'density' (sparseness) is important.

■ Is the density limit determined by

- V density limit?
- A density limit?

→ Appendix 4

Appendix 4: Matching sparse and dense random pulse trains

A fixed, V fixed, Both changed

■ Fixing the stimulus density of one modality sparse while changing the other from sparse to dense did not improve performance.

■ Little difference between A fix and V fix

■ Density limit was not determined by either V or A density limit alone.

■ Why is sparse-dense performance even worse than dense-dense performance?

- difficult to compare different temporal structures?