

聴覚音源定位に刺激時間パターンが与える影響に関する考察

澤井 賢一^{1,2,a)} 佐藤 好幸³ 合原 一幸¹

概要: 本発表では、聴覚の空間知覚に対して刺激の時間的な要因が与える影響について考える。異なる場所から提示された刺激間の距離を知覚する際、それらの刺激の時間間隔が長くなると、知覚される空間的な距離も長くなることなどが知られている。この現象を説明するベイズモデルでは、ある1つの物体が信号を発生しながら移動する際、その速度は一定であることなどを脳が事前知識として持っていることと仮定している。一方で、別の聴覚の時間知覚モデルにおいて、聴覚系が時間を知覚する際に刺激の同一源性を無意識下で確率的に判断していることも示唆されている。そこで我々は、空間知覚の際にも同一源性が判断されると考え、その判断が空間知覚に与える影響について考察する。

1. 聴覚空間知覚に対する時間的要因の影響

音楽を聴くとき、人によって嗜好の違いはもちろんあるが、一般にステレオ録音はモノラル録音よりも好まれ、生演奏の方が録音したものよりも聴く価値が高いとされる傾向がある。これは、音高や音色、リズムなどの音楽的な要素以外に、空間知覚によって生み出される臨場感や緊張感などが大きく影響しているためと考えられる。このように聴覚における空間知覚は、人の音楽理解を考える上で重要な要素の一つである。

聴覚が音源の方向を推定する際、左右の耳に届く音の時間差や強度差、位相差などの、音源の空間的な位置に起因する情報が重要と考えられている [1]。一方で、そのような空間的要因以外の性質がどう影響するかについては、いくつかの研究はあるものの [2]、いまだ十分に解明されているとは言えない。本発表では、非空間的要因の一つである刺激の時間パターンが、空間知覚にどのような影響を及ぼし得るのかについて考える。

空間知覚と時間知覚は相互に影響を与えることが知られている。そして、時間的要因が空間知覚に影響を及ぼす現象の一つとしてタウ効果 (τ -effect) が、逆に空間的要因が時間知覚に影響を及ぼす現象の一つとしてカップ効果 (κ -effect) がある [3]。いま、1次元空間上に並んだ3つの

点を順に A, B, C とし、A-B 間、B-C 間の距離をそれぞれ Y_1, Y_2 とする。そして、A-B 間、B-C 間でそれぞれ時間間隔を T_1, T_2 だけ空けて刺激を与えることを考える。タウ効果とは、 $Y_1 = Y_2$ かつ $T_1 < T_2$ のときに、知覚される空間間隔 \hat{Y}_1, \hat{Y}_2 が $\hat{Y}_1 < \hat{Y}_2$ となる現象である。またカップ効果は、 $T_1 = T_2$ かつ $Y_1 < Y_2$ のときに、知覚される時間間隔 \hat{T}_1, \hat{T}_2 が $\hat{T}_1 < \hat{T}_2$ となる現象である。すなわち、時間的 (または空間的) に等間隔の刺激の知覚に、空間的 (または時間的) な刺激の大小関係が反映される。これらの現象は、触覚、視覚など、様々な感覚様相で観察されている [3]。また、聴覚においては時間と周波数の関係においてカップ効果が生じることが分かっている [4]。

触覚におけるこれらのタウ効果やカップ効果に対し、脳が外界を知覚する際にベイズ推定を行っているとして説明するモデルが提案されている [5]。また、視聴覚情報の統合に関し、時間知覚と空間知覚が影響する現象もベイズ推定によって説明されている [6]。ベイズ推定による知覚のモデル化は、このモデル以外にもこれまで多くの成功を収めている [7]。このベイズモデルの枠組みでは、脳が感覚器を通じて得た観測情報と、現象に対して事前に持っている知識とを組み合わせることで、観測誤差に強いロバストな知覚を実現していると考えられる。そしてこの様子をベイズ推定とみなし、確率モデルとして定式化する。

一方で、聴覚における短音で区切られた時間パターンの知覚に対し、我々はベイズモデルを過去に提案している [8]。そこでは、時間的に隣接する音が同じ音源から生じたかどうかを、脳が無意識下で判断していることが示唆されている。そこで我々は、この同一源性の判断が空間知覚においても行われていると考え、その判断の影響について考察する。

¹ 東京大学 生産技術研究所
Institute of Industrial Science, University of Tokyo
² JST FIRST 合原最先端数理モデルプロジェクト
FIRST Aihara Innovative mathematical Modelling Project, JST
³ 電気通信大学 大学院情報システム学研究所
Graduate School of Information Systems, University of Electro-Communications
a) ken1@sat.t.u-tokyo.ac.jp

2. 刺激の同一源性判断を踏まえた時空間知覚モデル

ここでは、脳が3つの短音の発生時刻 t_1, t_2, t_3 と発生位置 y_1, y_2, y_3 を感覚器を通して観測し、その観測情報をもとに真の時間間隔 $T_1 = t_2 - t_1, T_2 = t_3 - t_2$ と真の空間間隔 $Y_1 = y_2 - y_1, Y_2 = y_3 - y_2$ を推定する状況を考える。この推定は、誤差を含んだそれぞれの音の観測時刻 s_1, s_2, s_3 と観測位置 x_1, x_2, x_3 を得た上で行われるため、事後分布 $P(T_1, T_2, Y_1, Y_2 | s_1, s_2, s_3, x_1, x_2, x_3)$ にもとづくと考えられる。そして、たとえば Y_2 の知覚を考える場合は、事後分布を Y_2 以外の確率変数について周辺化したものに対し、事後確率最大化 (MAP) 推定などによって知覚結果を計算する。この事後分布は、ベイズの定理より

$$P(T_1, T_2, Y_1, Y_2 | s_1, s_2, s_3, x_1, x_2, x_3) \propto P(s_1, s_2, s_3, x_1, x_2, x_3 | T_1, T_2, Y_1, Y_2) \times P(T_1, T_2, Y_1, Y_2) \quad (1)$$

と表せる。この右辺第1項は感覚器による観測、第2項は事前知識に対応し、それぞれ尤度関数、事前分布と呼ばれる。ここでは尤度関数に対し、真の値を中心とし、変数間に相関のない正規分布を用いる。

次に、先行研究の時空間知覚モデル [8] と同様に、隣り合う音が同じ音源から生じたかどうかを表す変数 C を導入する。いまは3つの短音の列を考えているため、この変数は次の4通りの値をとる：

- $C = 1$: 3つの音すべてが異なる音源から生じた；
- $C = 2$: はじめの2つの音は同じ音源から、3つ目は別の音源から生じた；
- $C = 3$: うしろの2つの音は同じ音源から、1つ目は別の音源から生じた；
- $C = 4$: 3つの音すべてが同一の音源から生じた。

この変数 C を用いると、事前分布は

$$P(T_1, T_2, Y_1, Y_2) = \sum_C P(T_1, T_2, Y_1, Y_2, C) = \sum_C P(C) P(T_1, T_2, Y_1, Y_2 | C) \quad (2)$$

と書ける。ここでは先行研究 [5], [8] をもとに、同一音源から生じる音に関する脳内の事前知識として次を仮定する：

- (1) 音の時間間隔は短い；
- (2) 音源の移動速度は遅い；
- (3) 音の時間間隔および空間間隔は等間隔である；
- (4) 音源の移動速度は一定である。

これらの仮定は、 $C = 1, 2, 3, 4$ のそれぞれに対して $P(T_1, T_2, Y_1, Y_2 | C)$ を定めることで定式化できる。たとえば $C = 2$ の場合、仮定 (1) より、1つ目と2つ目の音がなす時間間隔 T_1 は短く、かつ、仮定 (2) により音源の移動

速度 $V_1 := Y_1/T_1$ は小さいという事前確率が高い。また、3つ目の音のはじめの2つの音とは音源が異なるため、時間間隔 T_2 と空間間隔 Y_2 に関しては事前知識によるバイアスがかからず、事前分布は T_2 と Y_2 に関しては一様である。そのため、 $C = 2$ に対する事前分布は、

$$P(T_1, T_2, Y_1, Y_2 | C = 2) = P(T_1, Y_1 | C = 2) P(T_2, Y_2 | T_1, Y_1, C = 2) \quad (3)$$

$$= \frac{1}{2\pi L \sigma_T \sigma_V T_1} \exp \left[- \left(\frac{T_1^2}{2\sigma_T^2} + \frac{Y_1^2}{2\sigma_V^2 T_1^2} \right) \right] \quad (4)$$

となる。ただし、 L は積分範囲を表す定数、 σ_T と σ_V はそれぞれ仮定 (1) と (2) の知覚的バイアスの強さを表すパラメータである。この確率分布の概形を図1に示す。他の C の値に対する事前分布も同様に定めることができ、式 (2) により最終的な1つの事前分布を得て、さらに式 (1) によって事後分布を得る。

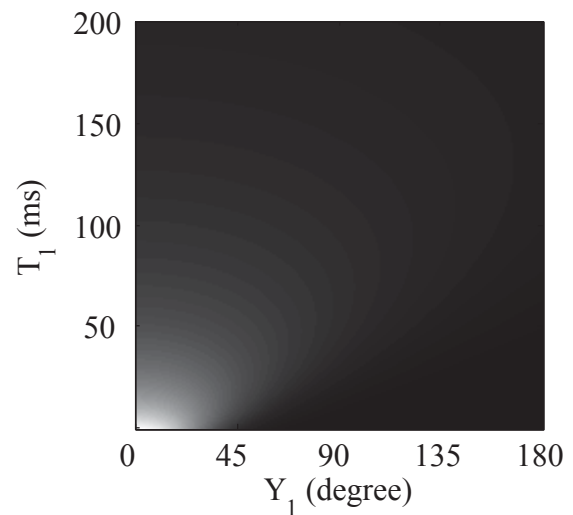


図1 3つの音のうち、はじめの2つだけが音源を共有していた場合の、1つ目の時間間隔 T_1 (ms) と空間間隔 Y_1 (degree) に対する事前分布。明るさが確率の大きさを表す。

3. 現象の予測例

ここでは、実験刺激のパラメータを (T_1, T_2, Y_1, Y_2) とし、 Y_2 の知覚に注目する。そのとき、前節の提案モデルにもとづく、以下の現象が起きると予測される。

- $T_1 < T_2, Y_1 > Y_2$ の場合。例: $(T_1, T_2, Y_1, Y_2) = (100\text{ms}, 400\text{ms}, 60^\circ, 30^\circ)$ 。このとき、刺激を発生する物体に対する速度一定の事前知識 [5] によると、 Y_2 は伸びて知覚されるはずである。しかし、この場合は1つ目と2つ目の間隔における音源の速度の差が大きいため、3つ目の音のはじめの2つの音とは異なる音源から発生した確率が高いと判断されやすい。そのため、2つ目の間隔に対しては知覚的バイアスが発生せず、錯覚も生じないと予測される。

- $V_1 = V_2$ ($V_i := Y_i/T_i, i = 1, 2$) かつ $T_1 < T_2$ の場合.
例: $(T_1, T_2, Y_1, Y_2) = (100\text{ms}, 200\text{ms}, 30^\circ, 60^\circ)$. このとき, 速度一定の事前知識 [5] によると錯覚は起きない. しかし, 時間知覚モデル [8] で示唆された「時間間隔は等間隔になりやすい」という事前知識を踏まえると, T_2 および Y_2 が短く知覚されることが予測される.

4. おわりに

本発表では, 同一源性を踏まえた時空間知覚モデルを提案し, 予測される現象について論じた. 現在, 予測される現象を確認する実験の方法について検討中である. さらに, 実験結果を触覚などにおける先行研究と比較することで, 各感覚様相の時間分解能や空間分解能の違いについて議論が可能と考えている.

謝辞 本研究は, 総合科学技術会議により制度設計された最先端研究開発支援プログラム (FIRST 合原最先端数理モデルプロジェクト) により, 日本学術振興会を通して助成を受けた. また, JSPS 科研費 23700309 の助成を受けた.

参考文献

- [1] Carr, C. E.: Processing of temporal information in the brain, *Annual Review of Neuroscience*, Vol. 16, No. 1, pp. 223–243 (1993).
- [2] Deutsch, D.: Two-channel listening to musical scales, *The Journal of the Acoustical Society of America*, Vol. 57, No. 5, pp. 1156–1160 (1975).
- [3] Jones, B. and Huang, Y. L.: Space-time dependencies in psychophysical judgment of extent and duration: Algebraic models of the tau and kappa effects, *Psychological Bulletin*, Vol. 91, No. 1, pp. 128–142 (1982).
- [4] Cohen, J., Hansel, C. E. M. and Sylvester, J. D.: Interdependence of Temporal and Auditory Judgments, *Nature*, Vol. 174, No. 4431, pp. 642–644 (1954).
- [5] Goldreich, D.: A Bayesian perceptual model replicates the cutaneous rabbit and other tactile spatiotemporal illusions, *PLoS ONE*, Vol. 2, No. 3, e333 (2007).
- [6] Sato, Y., Toyoizumi, T. and Aihara, K.: Bayesian inference explains perception of unity and ventriloquism aftereffect: identification of common sources of audiovisual stimuli, *Neural Computation*, Vol. 19, No. 12, pp. 3335–3355 (2007).
- [7] Körding, K.: Decision theory: what “should” the nervous system do?, *Science*, Vol. 318, No. 5850, pp. 606–610 (2007).
- [8] Sawai, K., Sato, Y. and Aihara, K.: Auditory time-interval perception as causal inference on sound sources, *Frontiers in Psychology*, Vol. 3, 524 (2012).