

# 多感覚統合の時空間特性

佐藤好幸<sup>†1</sup>

**概要:**我々は常に複数の感覚器を通して外界を認識しながら日々を送っている。本稿では、このような多感覚の統合、特に多感覚統合の中でも研究が進んでいる視聴覚統合に関する解説を行う。視聴覚統合は空間的統合、時間的統合、刺激内容の統合などに大別することができるが、ここでは刺激位置や刺激タイミングに関する基礎的な部分での統合に焦点を当て、その性質などについて解説を行う。

## Spatio-temporal characteristics of multisensory integration

Yoshiyuki SATO<sup>†1</sup>

### 1. はじめに

我々が周囲の物体や出来事を知覚する時、基本的には複数の感覚器官を同時に使って情報を得ている。例えば、他人と会話している時には、聴覚で声を聞くと同時に、視覚でも相手の表情や身振りなどを観察している。さらには、握手などのスキンシップによって触覚を用いることもあるし、嗅覚によって相手の身につけている香水や先程の食事内容などもわかるかもしれない。このように、複数の異なる感覚の情報を用いることで、より正確に対象について知ることができる。

多感覚統合の中でも、視覚と聴覚の情報統合、すなわち視聴覚統合について特によく研究が行われている。視覚と聴覚は、距離の離れた物体に関する情報を速く得ることができるなど、比較的近い性質を持っている。例えば、視覚と聴覚どちらにおいても、位置と時間という物理的性質を知ることができる。そのため、刺激の位置やタイミングに関して、視聴覚で得られた情報を直接的に統合し、より信頼性の高い知覚を得ることができる。

また同時に、視覚と聴覚は、光と音という異なる物理的媒質を通して、質の異なる情報を得ることもできる。例えば、他人と会話しているときには、話している内容は聴覚だけではなく視覚からもある程度得ることができる。このような、異なる感覚器を通して得た「内容」を統合することで、より良く知覚・認知を行うことも可能となる。

本稿では以下において、主に視聴覚統合を対象として、その時空間特性について、基礎的に知られた性質と関連研究について紹介する。内容の統合についても簡単に言及する。

### 2. 多感覚刺激の時空間的統合

#### 2.1 空間的統合

空間的な視聴覚統合としては腹話術効果と呼ばれる現象がよく知られている[1]。視覚刺激（例えばLEDの短時間の点灯）と聴覚刺激（例えばスピーカーから発せられた短いホワイトノイズ）を異なる位置から同時に提示すると（図1）、それを受け取った人は、音が光のあたりから発せられたように知覚する。これは、人形の口を動かさながら（視覚刺激）、腹話術師ができるだけ口を動かさないようにしてしゃべる（聴覚刺激）と、まるで人形がしゃべっているように感じる、という腹話術に例えてつけられた名前である。

また、このような空間的統合は視聴覚間の統合だけで起こるわけではなく、同様の位置に関する統合が視覚-触覚間[2]や聴覚-触覚間[3]でも起こることも知られている。

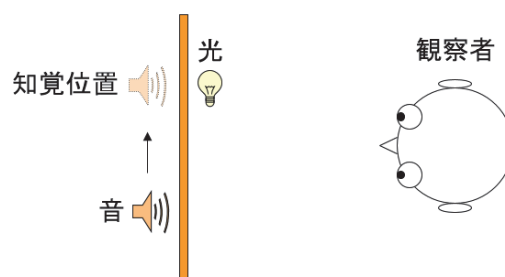


図1: 腹話術効果の概略図

#### 2.2 時間的統合

視聴覚刺激は時間的にも統合されることが知られている。例えばFendrichとCorballisの研究[4]では、瞬間的な視覚刺激（フラッシュ）と聴覚刺激（クリック音）を少し時間的にずらして提示すると、それぞれ時間的に引き寄せら

<sup>†1</sup> 東京大学大学院総合文化研究科  
Graduate School of Arts and Sciences, The University of Tokyo

れたタイミングに知覚されることが示された。これは、前述の腹話術効果と同様の引き寄せ現象が、空間だけではなく時間に対しても起こることを示している。このような類似性から、時間的な視聴覚刺激の統合は時間的腹話術効果とも呼ばれている[5], [6].

### 2.3 多感覚統合の統計的最適性

では、このような視聴覚刺激という2つの刺激が時空間的に統合される時に、どちらがどちらに統合されるのだろうか？腹話術効果においては、通常は音の定位が光の方向に大きく引き寄せられる。これは、位置の知覚精度に関しては視覚のほうが聴覚に比べて高いために、位置を問うタスクでは視覚が優位になるからである、と旧来は思われていた (modality appropriateness 仮説[7])。しかし、逆に視覚的に刺激の位置がわかりにくいような場合には、むしろ視覚刺激が聴覚刺激の位置に引き寄せられて知覚されることがわかった[8]。さらに、この実験結果は、視覚・聴覚による位置に関する不確実性を考慮して、それらを統計的に最適な統合を行って刺激の位置が推定される、としたときの予想と非常によく一致する[8].

このような統計的最適性は、時空間的統合以外にも様々な側面で多感覚統合の性質を説明できるモデルとして、近年活発に研究が行われている[9].

### 2.4 統合における刺激の同一源性の影響

視覚と聴覚で受け取った情報を統合することで利益があるのは、それら刺激が同じ源に関する情報を共有するからである。もし無関係であれば統合するべきではないであろう。知覚の際に観測者がこのような刺激の同一源性というものを(無意識的にでも)考慮に入れているのではないかと、という考えは unity assumption[7]と呼ばれ、多感覚統合を考える上では重要な要素となっている。

例えば、前述の腹話術効果は、やかんの映像とホイッスル音のような刺激を用いることでより強くなるということが知られている[10]。これは、そのようなもともと関係がありそうな刺激を使うことで、視聴覚刺激が同一源から出ているという確信がより強くなるから、と解釈することができる。

視覚刺激と聴覚刺激を統合する際に、刺激の位置だけでなくそれらが源を共有するかどうかとも同時に推定し、共有するときは統合して共有しないときは統合しない、とするモデルにより視聴覚統合の時空間特性の様々な側面が説明できることも示されてきている[11], [12].

### 2.5 多感覚刺激の繰り返しによる残効

多感覚統合の性質は固定的なものではなく、受け取った刺激に応じて柔軟に変化することが知られている。

腹話術効果の例のように空間的に少し離れた刺激を繰り返し提示すると、観測者はそれが同じ位置であると感じるようになる[13], [14]。この現象は腹話術残効と呼ばれる。その際、聴覚刺激のみを提示した時の定位は、繰り返し提

示期間中の視覚刺激の方向にずれる。視覚定位も、聴覚よりもずれは小さいものの、繰り返し提示された聴覚刺激の方向にずれるという結果もある[15].

また、時間的統合に関しても残効が起こる。例えば、時間的にずれた視聴覚刺激を繰り返し提示することで、そのような時間差がついた刺激を同時だと感じやすくなるということが知られている[16].

## 3. 多感覚刺激の内容的統合

現実の視聴覚刺激は空間的・時間的情報だけではなく、それを発したものに関して多くの情報を含んでいる。我々は光や音から、その情報の「内容」を知ることで日常生活を送っている。内容の統合に関しての詳細は続く講演に譲り、ここでは一例だけ紹介する。

視聴覚による知覚内容の統合として、マガーク効果がよく知られている[17]。この現象は、「バ(ba)」という音声に「ガ(ga)」と発音している映像を合成した動画を視聴すると、「ダ(da)」という本来存在しない音が聞こえるという、音韻知覚における視聴覚統合現象である。マガーク効果に関しては、視聴覚統合、言語情報処理、言語発達など様々な観点から数多くの研究が行われているが、ここでは現象の紹介だけにとどめたい。

## 4. おわりに

本稿では、多感覚統合の中でも特に広く研究されている視聴覚統合について、時空間特性という観点から簡単に紹介を行った。本稿で触れたのは視聴覚統合に関する研究でもごく一部の内容であり、その他の話題については例えば[18]等を参照されたい。

多感覚統合は、それぞれの単感覚の研究に比べると比較的研究の歴史も浅く、まだまだわかっていないことも多い。これからのますますの発展が期待できる豊かな研究分野である。

## 参考文献

- [1] I. Howard and W. Templeton, Human spatial orientation. John Wiley & Sons, 1966.
- [2] F. Pavani, C. Spence, and J. Driver, "Visual Capture of Touch: Out-of-the-Body Experiences With Rubber Gloves," Psychol. Sci., vol. 11, no. 5, pp. 353–359, 2000.
- [3] A. Caclin, S. Soto-Faraco, A. Kingstone, and C. Spence, "Tactile 'capture' of audition.," Percept. Psychophys., vol. 64, no. 4, pp. 616–30, May 2002.
- [4] R. Fendrich and P. M. Corballis, "The temporal cross-capture of audition and vision.," Percept. Psychophys., vol. 63, no. 4, pp. 719–725, 2001.
- [5] G. Aschersleben and P. Bertelson, "Temporal ventriloquism: crossmodal interaction on the time dimension 2. Evidence from sensorimotor synchronization," Int. J. Psychophysiol., vol. 50, pp. 157–163, 2003.
- [6] S. Morein-Zamir, S. Soto-Faraco, and A. Kingstone, "Auditory capture of vision: Examining temporal ventriloquism," Cogn.

- Brain Res., vol. 17, no. 1, pp. 154–163, 2003.
- [7] R. B. Welch and D. H. Warren, “Immediate perceptual response to intersensory discrepancy,” *Psychol. Bull.*, vol. 88, no. 3, pp. 638–667, 1980.
- [8] D. Alais and D. Burr, “The ventriloquist effect results from near-optimal bimodal integration,” *Curr. Biol.*, vol. 14, no. 3, pp. 257–62, Feb. 2004.
- [9] J. Trommershauser, K. Kording, and M. S. Landy, Eds., *Sensory Cue Integration*. New York, Oxford: Oxford University Press, 2011.
- [10] C. V. Jackson, “Visual factors in auditory localization,” *Q. J. Exp. Psychol.*, vol. 5, no. 2, pp. 52–65, 1953.
- [11] Y. Sato, T. Toyoizumi, and K. Aihara, “Bayesian inference explains perception of unity and ventriloquism aftereffect: identification of common sources of audiovisual stimuli,” *Neural Comput.*, vol. 19, no. 12, pp. 3335–55, Dec. 2007.
- [12] K. P. Körding, U. Beierholm, W. J. Ma, S. Quartz, J. B. Tenenbaum, and L. Shams, “Causal inference in multisensory perception,” *PLoS One*, vol. 2, no. 9, p. e943, Jan. 2007.
- [13] R. Held, “Shifts in binaural localization after prolonged exposures to atypical combinations of stimuli,” *Am. J. Psychol.*, vol. 68, no. 4, pp. 526–548, 1955.
- [14] L. K. Canon, “Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation,” *J. Exp. Psychol.*, vol. 84, no. 1, pp. 141–147, 1970.
- [15] J. Lewald, “Rapid Adaptation to Auditory-Visual Spatial Disparity,” *Learn. Mem.*, vol. 9, no. 5, pp. 268–278, Sep. 2002.
- [16] W. Fujisaki, S. Shimojo, M. Kashino, and S. Nishida, “Recalibration of audiovisual simultaneity,” *Nat. Neurosci.*, vol. 7, no. 7, pp. 773–8, Jul. 2004.
- [17] H. McGurk and J. MacDonald, “Hearing lips and seeing voices,” *Nature*, vol. 264, no. 5588, pp. 746–748, 1976.
- [18] B. E. Stein, *The new handbook of multisensory processing*. Mit Press Cambridge, MA, 2012.