Audio and Visual Analytics in Marketing and Artificial Empathy

Other titles in Foundations and Trends® in Marketing

Language Research in Marketing
Ann Kronrod
ISBN: 978-1-68083-966-1

Reflections of Eminent Marketing Scholars
Dawn Iacobucci
ISBN: 978-1-68083-954-8

Social Media Analytics and Its Applications in Marketing
Sangkil Moon and Dawn Iacobucci
ISBN: 978-1-68083-952-4

Privacy and Consumer Empowerment in Online Advertising
W. Jason Choi and Kinshuk Jerath
ISBN: 978-1-68083-920-3

Studying Healthcare from a Marketing Perspective
Dawn Iacobucci and Deidre Popovich
ISBN: 978-1-68083-918-0

Aggregate Advertising Expenditure in the U.S. Economy: Measurement and Growth Issues in the Digital Era
Alvin J. Silk and Ernst R. Berndt
ISBN: 978-1-68083-872-5

Audio and Visual Analytics in Marketing and Artificial Empathy

Shasha Lu
University of Cambridge
UK
s.lu@jbs.cam.ac.uk

Hye-Jin Kim
Korea Advanced Institute of Science and Technology (KAIST)
South Korea
hyejinkim@kaist.ac.kr

Yinghui Zhou
Shenzhen University
China
yhzhou@szu.edu.cn

Li Xiao
Fudan University
China
lixiao@fudan.edu.cn

Min Ding
Pennsylvania State University
USA
minding@psu.edu

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Marketing

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

S. Lu *et al.*. *Audio and Visual Analytics in Marketing and Artificial Empathy*. Foundations and Trends® in Marketing, vol. 16, no. 4, pp. 422–493, 2022.

ISBN: 978-1-68083-971-5

Foundations and Trends® in Marketing
Volume 16, Issue 4, 2022

Editorial Board

Editorial Scope

Topics

Foundations and Trends® in Marketing publishes survey and tutorial articles in the following topics:

- B2B Marketing
- Bayesian Models
- Behavioral Decision Making
- Branding and Brand Equity
- Channel Management
- Choice Modeling
- Comparative Market Structure
- Competitive Marketing Strategy
- Conjoint Analysis
- Customer Equity
- Customer Relationship Management
- Game Theoretic Models
- Group Choice and Negotiation
- Discrete Choice Models
- Individual Decision Making
- Marketing Decisions Models
- Market Forecasting
- Marketing Information Systems
- Market Response Models
- Market Segmentation
- Market Share Analysis
- Multi-channel Marketing
- New Product Diffusion
- Pricing Models
- Product Development
- Product Innovation
- Sales Forecasting
- Sales Force Management
- Sales Promotion
- Services Marketing
- Stochastic Model

Information for Librarians

Foundations and Trends® in Marketing, 2022, Volume 16, 4 issues. ISSN paper version 1555-0753. ISSN online version 1555-0761. Also available as a combined paper and online subscription.

Contents

Audio and Visual Analytics in Marketing and Artificial Empathy

Shasha Lu[1], Hye-Jin Kim[2], Yinghui Zhou[3], Li Xiao[4] and Min Ding[5]

[1]*Associate Professor of Marketing, Cambridge Judge Business School, University of Cambridge, UK; s.lu@jbs.cam.ac.uk*
[2]*Assistant Professor of Marketing, School of Business and Technology Management, College of Business, Korea Advanced Institute of Science and Technology (KAIST), South Korea; hyejinkim@kaist.ac.kr*
[3]*Assistant Professor of Marketing, College of Management, Shenzhen University, China; yhzhou@szu.edu.cn*
[4]*Associate Professor of Marketing, School of Management, Fudan University, China; lixiao@fudan.edu.cn*
[5]*Bard Professor of Marketing at Smeal College of Business and the Affiliate Professor in the College of Information Sciences and Technology, Pennsylvania State University, USA; minding@psu.edu*

ABSTRACT

With the ever-cheaper digital equipment and the prevalent digital platforms such as Facebook and YouTube, more and more human behaviors and activities are digitalized in the form of images, videos, and audio. However, due to the information's unstructured nature, there has been a lack of useful framework and tools that can help businesses to effectively leverage this information to improve their practices. As a result, businesses are missing out on the opportunities to use this information to gain better customer insights, understand customer preference, improve customer experience, discover unmet needs and optimize marketing

Shasha Lu, Hye-Jin Kim, Yinghui Zhou, Li Xiao and Min Ding (2022), "Audio and Visual Analytics in Marketing and Artificial Empathy", Foundations and Trends® in Marketing: Vol. 16, No. 4, pp 422–493. DOI: 10.1561/1700000067.

effectiveness. In this monograph, the authors present an overview of audio and visual analytics and discuss how they can be used by marketers to improve business practices. This monograph first introduced a framework named *Artificial Empathy* (AE) to illustrate different contexts where the audio and/or visual information emitted by or presented to an individual are used to improve business decision making. Next, it presented a review of the cutting-edge techniques and methods used to mine valuable information and make useful inferences from the audio and visual data. Finally, it reviewed the use of A/V analytics in business practices and concluded with a discussion on the trends in applying audio and visual data analytics in business. This monograph aims to help readers understand how the new forms of rich data will affect the way we do business and gain insights into harnessing the power of audio, image, and video data to make useful inferences and improve business practices.

Keywords: artificial empathy; audio analytics; image analytics; video analytics

1

Introduction

With the ever-cheaper digital recording equipment and the prevalent use of online channels for interpersonal interactions such as Facebook, YouTube, TikTok and Zoom, more and more human behaviors and activities are digitalized in the form of images, videos and audio. In fact, one major trend in the digital economy is the exponentially growing number of images, videos and audios generated and consumed every day. For example, in one minute there are 700,000 hours of videos watched and 500 hours of videos uploaded on YouTube, 243,000 photos uploaded on Facebook and 1,000,000 swipes on Tinder, and 400,000 hours of music listened on Spotify.[1] In addition, the creation and consumption of the audio and visual content have been amplified by the global pandemic during the last two years. The ubiquity of audio and visual information has given birth to the development of technologies and applications to transform both online and traditional businesses and marketing practices. For example, face recognition technology has now been widely used in online and mobile payment systems.[2] Over 5 million face videos have been used by Affectiva, an emotion-analysis start-up,

[1] https://www.go-globe.com/blog/things-that-happen-every-60-seconds/.
[2] https://www.bbc.co.uk/news/business-55748964.

to understand how consumers react to video ads.[3] The audio, image, and video information contains much valuable information on many topics of interest to businesses, including advertising, branding, content marketing, retailing, e-commerce, customer relationship management, product recommendation, and data privacy etc. However, it's almost impossible for human beings to process this amount of information on a daily basis. In order to derive useful insights from this ever-growing ocean of information, organizations need more efficient and effective ways to process and analyze the rich audio/visual data and new models to assist business decision making related to these data.

Over the past decade, there has been growing interest among researchers across many disciplines including business management, computer science, neuroscience, and social science to use automated audio and visual data analytics for more efficient and effective decision making in various contexts. However, the applications of audio and visual data analytics are still very limited in practice and the majority of organizations have yet to unlock the potential of audio/visual data analytics for business decision making in the post-digital economy. To this end, many important questions have yet to be answered in this domain, such as, in what contexts should marketers consider using audio/visual data analytics? What types of business problems can be and are best addressed by audio/visual data analytics? What models are available for different types of problems and what skills should researchers and/or practitioners develop to better manage and leverage the potential of audio/visual data in various business contexts? This monograph aims to address these questions by (1) developing a framework for understanding the internal states of individuals based on audio/visual signals and incorporating works in the domain of audio/visual (A/V) data analytics into marketing research in order to identify future research opportunities; (2) providing an overview of methodologies that are commonly used in conducting research with A/V data; (3) providing a review of A/V analytics-based research in various business contexts; and

[3] https://blog.affectiva.com/the-worlds-largest-emotion-database-5.3-million-faces-and-counting.

(4) reviewing the business practices using A/V analytics and identifying the future trends in both research and business applications.

The rest of this monograph is organized as follows: we first propose a framework for A/V data-based research in the business domain and discuss how A/V data analytics can be used to support business decision making in various contexts. We then provide an overview of the key techniques and tools used in A/V data analytics and discuss the procedures and key methodological questions. Finally, we discuss how the A/V analytics has been used in business practices and its trend and future development.

References

Adaval, R., G. Saluja, and Y. Jiang (2019). "Seeing and thinking in pictures: A review of visual information processing". *Consumer Psychology Review*. 2(1): 50–69.

Bagozzi, R. P., M. Gopinath, and P. U. Nyer (1999). "The role of emotions in marketing". *Journal of the Academy of Marketing Science*. 27(2): 184–206.

Balducci, B. and D. Marinova (2018). "Unstructured data in marketing". *Journal of the Academy of Marketing Science*. 46(4): 557–590.

Belin, P., P. E. Bestelmeyer, M. Latinus, and R. Watson (2011). "Understanding voice perception". *British Journal of Psychology*. 102(4): 711–725.

Bengio, Y., A. Courville, and P. Vincent (2013). "Representation learning: A review and new perspectives". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 35(8): 1798–1828.

Berger, J., A. Humphreys, S. Ludwig, W. W. Moe, O. Netzer, and D. A. Schweidel (2020). "Uniting the tribes: Using text for marketing insight". *Journal of Marketing*. 84(1): 1–25.

Bernstein, A. V. and E. V. Burnaev (2018). "Reinforcement learning in computer vision". *Tenth International Conference on Machine Vision (ICMV 2017)*. 10696: 458–464.

Berry, D. S. and L. Z. McArthur (1985). "Some components and consequences of a babyface". *Journal of Personality and Social Psychology*. 48(2): 312.

Bharadwaj, N., M. Ballings, P. A. Naik, M. Moore, and M. M. Arat (2022). "A new livestream retail analytics framework to assess the sales impact of emotional displays". *Journal of Marketing*. 86(1): 27–47.

Bianco, S., D. Mazzini, D. Pau, and R. Schettini (2015). "Local detectors and compact descriptors for visual search: A quantitative comparison". *Digital Signal Process.* 44: 1–13.

Brei, V. A. (2020). "Machine learning in marketing: Overview, learning strategies, applications, and future developments". *Foundations and Trends® in Marketing*. 14(3): 173–236.

Brown, S. P., P. M. Homer, and J. J. Inman (1998). "A meta-analysis of relationships between ad-evoked feelings and advertising responses". *Journal of Marketing Research*. 35(1): 114–126.

Bruce, V. and A. Young (2013). *Face Perception*. Psychology Press.

Bruner, G. C. (1990). "Music, mood, and marketing". *Journal of Marketing*. 54(4): 94–104.

Buie, D. H. (1981). "Empathy: Its nature and limitations". *Journal of the American Psychoanalytic Association*. 29: 281–307.

Burnap, A., J. R. Hauser, and A. Timoshenko (2021). "Design and evaluation of product aesthetics: A human-machine hybrid approach". Available at SSRN 3421771.

Busso, C., M. Bulut, S. Lee, and S. Narayanan (2009). *Fundamental Frequency Analysis for Speech Emotion Processing. The Role of Prosody in Affective Speech*. Berlin, Germany: Peter Lang Publishing Group. 309–337.

Caldwell, C. and S. A. Hibbert (2002). "The influence of music tempo and musical preference on restaurant patrons' behavior". *Psychology and Marketing*. 19(11): 895–917.

Chaki, J. and N. Dey (2018). *A Beginner's Guide to Image Preprocessing Techniques*. CRC Press.

Chandon, P., J. W. Hutchinson, E. T. Bradlow, and S. H. Young (2009). "Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase". *Journal of Marketing*. 73(6): 1–17.

Chattopadhyay, A., D. W. Dahl, R. J. B. Ritchie, and K. N. Shahin (2003). "Hearing voices: The impact of announcer speech characteristics on consumer response to broadcast advertising". *Journal of Consumer Psychology*. 13(3): 198–204.

Cheang, H. S. and M. D. Pell (2008). "The sound of Sarcasm". *Speech Communication*. 50(5): 366–381.

Clarke, J. S., J. P. Cornelissen, and M. P. Healey (2019). "Actions speak louder than words: How figurative language and gesturing in entrepreneurial pitches influences investment judgments". *Academy of Management Journal*. 62(2): 335–360.

Connell, J., Q. Fan, P. Gabbur, N. Haas, S. Pankanti, and H. Trinh (2013). "Retail video analytics: An overview and survey". In: *Video Surveillance and Transportation Imaging Applications*. Vol. 8663.

Cunningham, M. R., A. P. Barbee, and C. L. Pike (1990). "What do women want? Facialmetric assessment of multiple motives in the perception of male facial physical attractiveness". *Journal of Personality and Social Psychology*. 59(1): 61–72.

Decety, J. and P. L. Jackson (2004). "The functional architecture of human empathy". *Behavioral and Cognitive Neuroscience Reviews*. 3(2): 71–100.

Decety, J. and C. Lamm (2006). "Human empathy through the lens of social neuroscience". *The Scientific World Journal*. 6: 1146–1163.

Decety, J. and W. Ickes (eds.) (2011). *The Social Neuroscience of Empathy*. MIT Press.

DeGroot, T. and S. J. Motowidlo (1999). "Why visual and vocal interview cues can affect interviewers' judgments and predict job performance". *Journal of Applied Psychology*. 84(6): 986–993.

DeShields, J., W. Oscar, K. Ali, and K. Erdener (1996). "Source effects in purchase decisions: The impact of physical attractiveness and accent of salesperson". *International Journal of Research in Marketing*. 13(1): 89–101.

Dew, R., A. Ansari, and O. Toubia (2021). "Letting logos speak: Leveraging multiview representation learning for data-driven branding and logo design". *Marketing Science*. Articles in Advance.

Ding, M. (2007). "A theory of intraperson games". *Journal of Marketing*. 71(2): 1–11.

Duan, C. and C. E. Hill (1996). "The current state of empathy research". *Journal of Counseling Psychology*. 43(3): 261.

Duarte, J., S. Siegel, and L. Young (2012). "Trust and credit: The role of appearance in peer-to-peer lending". *The Review of Financial Studies*. 25(8): 2455–2484.

Dzyabura, D., S. El Kihal, J. R. Hauser, and M. Ibragimov (2019). "Leveraging the power of images in managing product return rates". Available at SSRN 3209307.

Dzyabura, D. and R. Peres (2021). "Visual elicitation of brand perception". *Journal of Marketing*. 85(4): 44–66.

Ekman, P., M. O'Sullivan, W. V. Friesen, and K. R. Scherer (1991). "Invited article: Face, voice, and body in detecting deceit". *Journal of Nonverbal Behavior*. 15(2): 125–135.

Ert, E., A. Fleischer, and N. Magen (2016). "Trust and reputation in the sharing economy: The role of personal photos in Airbnb". *Tourism Management*. 55: 62–73.

Eslami, M., K. Vaccaro, K. Karahalios, and K. Hamilton (2017). "Be careful; things can be worse than they appear: Understanding biased algorithms and users' behavior around them in rating platforms". In: *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 11.

Eyben, F., K. R. Scherer, B. W. Schuller, J. Sundberg, E. André, and C. Busso, ... and K. P. Truong (2015). "The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing". *IEEE Transactions on Affective Computing*. 7(2): 190–202.

Fischer, A. H. and A. S. R. Manstead (2008). "Social functions of emotion". In: *Handbook of Emotions*. Ed. by M. Lewis, J. Haviland-Jones, and L. F. Barrett. 3rd edition. New York, NY: Guilford Press. 456–468.

Folkes, V. and S. Matta (2004). "The effect of package shape on consumers' judgments of product volume: Attention as a mental contaminant". *Journal of Consumer Research*. 31: 390–401.

Forsyth, D. and J. Ponce (2011). *Computer Vision: A Modern Approach*. Prentice Hall.

Fortun, D., P. Bouthemy, and C. Kervrann (2015). "Optical flow modeling and computation: A survey". *Computer Vision and Image Understanding*. 134: 1–21.

Fortune (2021). "A.I. in the beauty industry: How the pandemic finally made consumers care about it". URL: https://fortune.com/2021/01/11/ai-artificial-intelligence-personalized-beauty-cosmetics-brainstorm-reinvent/ Retrieved on 8 Oct. 2021.

Frühholz, S. and P. Belin (eds.) (2018). *The Oxford Handbook of Voice Perception*. Oxford University Press.

Gerhard, D. (2003). *Pitch Extraction and Fundamental Frequency: History and Current Techniques*. Regina, SK, Canada: Department of Computer Science, University of Regina. 0–22.

Giannakopoulos, T. and A. Pikrakis (2014). *Introduction to Audio Analysis: A MATLAB® Approach*. Academic Press.

Giertz, J. N., W. H. Weiger, M. Törhönen, and J. Hamari (2021). "Content versus community focus in live streaming services: How to drive engagement in synchronous social media". *Journal of Service Management*. 33(1): 33–58.

Goldman, A. (1993). "Ethics and cognitive science". *Ethics*. 103: 337–360.

Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, and S. Ozair, ... and Y. Bengio (2014). "Generative adversarial nets". *Advances in Neural Information Processing Systems*. 27.

Gorn, G. J., A. Chattopadhyay, J. Sengupta, and S. Tripathi (2004). "Waiting for the web: How screen color affects time perception". *Journal of Marketing Research*. 41(2): 215–225.

Gorn, G. J., A. Chattopadhyay, T. Yi, and D. W. Dahl (1997). "Effects of color as an executional cue in advertising: They're in the shade". *Management Science*. 43(10): 1387–1400.

Gorn, G. J., Y. Jiang, and G. V. Johar (2008). "Babyfaces, trait inferences, and company evaluations in a public relations crisis". *Journal of Consumer Research*. 35(1): 36–49.

Guan, Y., Y. Tan, Q. Wei, and G. Chen (2020). "Information or distortion? The effect of customer generated images on product rating dynamics". *Working Paper*.

Hagtvedt, H. and S. A. Brasel (2017). "Color saturation increases perceived product size". *Journal of Consumer Research*. 44(2): 396–413.

Hall, J. A., P. Verghis, W. Stockton, and J. X. Goh (2014). "It takes just 120 seconds: Predicting satisfaction in technical support calls". *Psychology and Marketing*. 31(7): 500–508.

Hartmann, J., M. Heitmann, C. Schamp, and O. Netzer (2021). "The power of brand selfies". *Journal of Marketing Research*. 58(6): 1159–1177.

Hauser, J. R., G. L. Urban, G. Liberali, and M. Braun (2009). "Website morphing". *Marketing Science*. 28(2): 202–223.

He, K., X. Zhang, S. Ren, and J. Sun (2016). "Deep residual learning for image recognition". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 770–778.

Henderson, P. W., J. A. Cote, S. M. Leong, and B. Schmitt (2003). "Building strong brands in Asia: Selecting the visual components of image to maximize brand strength". *International Journal of Research in Marketing*. 20: 297–313.

Hennig-Thurau, T., M. Groth, M. Paul, and D. D. Gremler (2006). "Are all smiles created equal? How emotional contagion and emotional labor affect service relationships". *Journal of Marketing*. 70(3): 58–73.

Hickman, L., N. Bosch, V. Ng, R. Saef, L. Tay, and S. E. Woo (2021). "Automated video interview personality assessments: Reliability, validity, and generalizability investigations". *Journal of Applied Psychology*. Articles in Advance.

Hodges-Simeon, C. R., S. J. Gaulin, and D. A. Puts (2010). "Different vocal parameters predict perceptions of dominance and attractiveness". *Human Nature*. 21(4): 406–427.

Hoegg, J., J. W. Alba, and D. W. Dahl (2010). "The good, the bad, and the ugly: Influence of aesthetics on product feature judgments". *Journal of Consumer Psychology*. 20: 419–430.

Hossein-Nejad, Z., H. Agahi, and A. Mahmoodzadeh (2021). "Image matching based on the adaptive redundant keypoint elimination method in the SIFT algorithm". *Pattern Analysis and Applications*. 24(2): 669–683.

Hughes, C., V. Swaminathan, and G. Brooks (2019). "Driving brand engagement through online social influencers: An empirical investigation of sponsored blogging campaigns". *Journal of Marketing*. 83(5): 78–96.

Hui, S. K., P. S. Fader, and E. T. Bradlow (2009). "Path data in marketing: An integrative framework and prospectus for model building". *Marketing Science*. 28(2): 320–335.

Hui, S. K., Y. Huang, J. Suher, and J. Inman (2013). "Deconstructing the 'first moment of truth': Understanding unplanned consideration and purchase conversion using in-store video tracking". *Journal of Marketing Research*. 50(4): 445–462.

Hul, M. K., L. Dube, and J. C. Chebat (1997). "The impact of music on consumers' reactions to waiting for services". *Journal of Retailing*. 73(1): 87–104.

Ickes, W. (1997). *Empathic Accuracy*. New York: The Guilford Press.

Jalali, N. Y. and P. Papatla (2016). "The palette that stands out: Color compositions of online curated visual UGC that attracts higher consumer interaction". *Quantitative Marketing and Economics*. 14(4): 353–384.

Jia, H., B. K. Kim, and L. Ge (2020). "Speed up, size down: How animated movement speed in product videos influences size assessment and product evaluation". *Journal of Marketing*. 84(5): 100–116.

Jiang, L., D. Yin, and D. Liu (2019). "Can joy buy you money? The impact of the strength, duration, and phases of an entrepreneur's peak displayed joy on funding performance". *Academy of Management Journal*. 62(6): 1848–1871.

Jiang, Y., G. J. Gorn, M. Galli, and A. Chattopadhyay (2016). "Does your company have the right logo? How and why circular- and angular-logo shapes influence brand attribute judgments". *Journal of Consumer Research*. 42: 709–726.

Joshi, K. A. and D. G. Thakore (2012). "A survey on moving object detection and tracking in video surveillance system". *International Journal of Soft Computing and Engineering*. 2(3): 44–48.

Kachur, A., E. Osin, D. Davydov, K. Shutilov, and A. Novokshonov (2020). "Assessing the big five personality traits using real-life static facial images". *Scientific Reports*. 10(1): 1–11.

Kahle, L. R. and P. M. Homer (1985). "Physical attractiveness of the celebrity endorser: A social adaptation perspective". *Journal of Consumer Research*. 11(4): 954–961.

Kang, D., P. Bailis, and M. Zaharia (2019). "Challenges and opportunities in DNN-based video analytics: A demonstration of the Blazeit video query engine". In: *CIDR*.

Kazama, M., S. Gotoh, M. Tohyama, and T. Houtgast (2010). "On the significance of phase in the short term Fourier spectrum for speech intelligibility". *The Journal of the Acoustical Society of America*. 127(3): 1432–1439.

Keh, H. T., R. Ren, S. R. Hill, and X. Li (2013). "The beautiful, the cheerful, and the helpful: The effects of service employee attributes on customer satisfaction". *Psychology and Marketing*. 30(3): 211–226.

Kent, R. D. and C. Read (1992). *The Acoustic Analysis of Speech*. London: Whurr.

Kim, H. J., Y. Wang, and M. Ding (2021). "Brand voiceprint". *Customer Needs and Solutions*: 1–14.

Kingma, D. P. and M. Welling (2013). "Auto-encoding variational bayes". arXiv preprint arXiv: 1312.6114.

Klofstad, C. A., R. C. Anderson, and S. Peters (2012). "Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women". *Proceedings of the Royal Society B: Biological Sciences*. 279(1738): 2698–2704.

Klostermann, J., A. Plumeyer, D. Böger, and R. Decker (2018). "Extracting brand information from social networks: Integrating image, text, and social tagging data". *International Journal of Research in Marketing*. 35(4): 538–556.

Kreiman, J. and D. Sidtis (2011). "Voices and listeners: Toward a model of voice perception". *Acoustics Today*. 7(4): 7–15.

Krider, R. E., P. Raghubir, and A. Krishna (2001). "Pizzas: π or square? Psychophysical biases in area comparisons". *Marketing Science*. 20(4): 405–425.

Kumar, A., R. Bezawada, R. Rishika, R. Janakiraman, and P. K. Kannan (2016). "From social to sale: The effects of firm-generated content in social media on customer behavior". *Journal of Marketing*. 80(1): 7–25.

Landwehr, J. R., D. Wentzel, and A. Herrmann (2013). "Product design for the long run: Consumer responses to typical and atypical designs at different stages of exposure". *Journal of Marketing*. 77(5): 92–107.

Le, N., V. S. Rathour, K. Yamazaki, K. Luu, and M. Savvides (2021). "Deep reinforcement learning in computer vision: A comprehensive survey". *Artificial Intelligence Review*: 1–87.

Levenson, R. W. and A. M. Ruef (1992). "Empathy: A physiological substrate". *Journal of Personality and Social Psychology*. 63: 234–246.

Li, H., D. Simchi-Levi, M. X. Wu, and W. Zhu (2019a). "Estimating and exploiting the impact of photo layout in the sharing economy". Available at SSRN 3470877.

Li, S. and W. Deng (2020). "Deep facial expression recognition: A survey". *IEEE Transactions on Affective Computing*: 1–20.

Li, X., M. Shi, and X. S. Wang (2019b). "Video mining: Measuring visual information using automatic methods". *International Journal of Research in Marketing*. 36(2): 216–231.

Li, Y. and Y. Xie (2020). "Is a picture worth a thousand words? An empirical study of image content and social media engagement". *Journal of Marketing Research*. 57(1): 1–19.

Liaukonyte, J., T. Teixeira, and K. C. Wilbur (2015). "Television advertising and online shopping". *Marketing Science*. 34(3): 311–330.

Lin, Y., D. Yao, and X. Chen (2021). "Happiness begets money: Emotion and engagement in live streaming". *Journal of Marketing Research*. 58(3): 417–438.

Liu, H., S. Chen, and N. Kubota (2013). "Intelligent video systems and analytics: A survey". *IEEE Transactions on Industrial Informatics*. 9(3): 1222–1233.

Liu, L., D. Dzyabura, and N. Mizik (2020). "Visual listening in: Extracting brand image portrayed on social media". *Marketing Science*. 39(4): 669–686.

Liu, X., S. W. Shi, T. Teixeira, and M. Wedel (2018). "Video content marketing: The making of clips". *Journal of Marketing*. 82(4): 86–101.

Liu, Y., K. J. Li, H. Chen, and S. Balachander (2017). "The effects of products' aesthetic design on demand and marketing-mix effectiveness: The role of segment prototypicality and brand consistency". *Journal of Marketing*. 81(1): 83–102.

Lu, S., L. Xiao, and M. Ding (2016). "A video-based automated recommender (VAR) system for garments". *Marketing Science*. 35(3): 484–510.

Lu, S., D. Yao, X. Chen, and R. Grewal (2021). "Do larger audiences generate greater revenues under pay what you want? Evidence from a live streaming platform". *Marketing Science*. 40(5): 964–984.

Luffarelli, J., M. Mukesh, and A. Mahmood (2019a). "Let the logo do the talking: The influence of logo descriptiveness on brand equity". *Journal of Marketing Research*. 56(5): 862–878.

Luffarelli, J., A. Stamatogiannakis, and H. Yang (2019b). "The visual asymmetry effect: An interplay of logo design and brand personality on brand equity". *Journal of Marketing Research*. 56(1): 89–103.

Ma, X., Z. Wu, J. Jia, M. Xu, H. Meng, and L. Cai (2018). "Emotion recognition from variable-length speech segments using deep learning on spectrograms". In: *Interspeech*. 3683–3687.

Ma, Z. and L. Dubé (2011). "Process and outcome interdependency in frontline service encounters". *Journal of Marketing*. 75(3): 83–98.

Malik, N., P. V. Singh, D. D. Lee, and K. Srinivasan (2019). "A dynamic analysis of beauty premium". Available at SSRN 3208162.

Marinova, D., S. K. Singh, and J. Singh (2018). "Frontline problem-solving effectiveness: A dynamic analysis of verbal and nonverbal cues". *Journal of Marketing Research*. 55(2): 178–192.

Mattila, A. S. and C. A. Enz (2002). "The role of emotions in service encounters". *Journal of Service Research*. 4(4): 268–277.

Mayew, W. J. and M. Venkatachalam (2012). "The power of voice: Managerial affective states and future firm performance". *Journal of Finance*. 67(1): 1–43.

McAleer, P., A. Todorov, and P. Belin (2014). "How do you say 'Hello'? Personality impressions from brief novel voices". *PloS One*. 9(3): e90779.

Mei, T., X. S. Hua, L. Yang, and S. Li (2007). "VideoSense: Towards effective online video advertising". In: *Proceedings of the 15th ACM International Conference on Multimedia*. September. 1075–1084.

Messer, U. and S. Fausser (2019). "Predicting social perception from faces: A deep learning approach". *Working Paper*.

Miller, N., G. Maruyama, R. J. Beaber, and K. Valone (1976). "Speed of speech and persuasion". *Journal of Personality and Social Psychology*. 34(4): 615.

Moeslund, T. B., A. Hilton, and V. Krüger (2006). "A survey of advances in vision-based human motion capture and analysis". *Computer Vision and Image Understanding*. 104(2–3): 90–126.

Nanne, A. J., M. L. Antheunis, C. G. van der Lee, E. O. Postma, S. Wubben, and G. van Noort (2020). "The use of computer vision to analyze brand-related user generated image content". *Journal of Interactive Marketing*. 50: 156–167.

Nelson, R. G. and D. Schwartz (1979). "Voice-pitch analysis". *Journal of Advertising Research*. 19(5): 55–59.

Nighswonger, N. J. and C. R. Martin, Jr. (1981). "On using voice analysis in marketing research". *Journal of Marketing Research*. 18(3): 350–355.

Oleszkiewicz, A., K. Pisanski, K. Lachowicz-Tabaczek, and A. Sorokowska (2017). "Voice-based assessments of trustworthiness, competence, and warmth in blind and sighted adults". *Psychonomic Bulletin and Review*. 24(3): 856–862.

Oosterhof, N. N. and A. Todorov (2009). "Shared perceptual basis of emotional expressions and trustworthiness impressions from faces". *Emotion*. 9(1): 128.

Orth, U. R. and R. C. Crouch (2014). "Is beauty in the aisles of the retailer? Package processing in visually complex contexts". *Journal of Retailing*. 90(4): 524–537.

Orth, U. R. and K. Malkewitz (2008). "Holistic package design and consumer brand impressions". *Journal of Marketing*. 72(3): 64–81.

Pennington, A. L. (1968). "Customer-salesman bargaining behavior in retail transactions". *Journal of Marketing Research*. 5(3): 255–262.

Peschel, A. O. and J. L. Orquin (2013). "A review of the findings and theories on surface size effects on visual attention". *Frontiers in Psychology*. 4: 902.

Pieters, R. and M. Wedel (2004). "Attention capture and transfer in advertising: Brand, pictorial, and text-size effects". *Journal of Marketing*. 68(2): 36–50.

Pieters, R., M. Wedel, and R. Batra (2010). "The stopping power of advertising: Measures and effects of visual complexity". *Journal of Marketing*. 74(5): 48–60.

Podder, P., T. Z. Khan, M. H. Khan, and M. M. Rahman (2014). "Comparative performance analysis of hamming, hanning and blackman window". *International Journal of Computer Applications*. 96(18): 1–7.

Pon-Barry, H. and S. M. Shieber (2011). "Recognizing uncertainty in speech". *EURASIP Journal on Advances in Signal Processing*. Special Issue on Emotion and Mental State Recognition from Speech.

Poor, M., A. Duhachek, and H. S. Krishnan (2013). "How images of other consumers influence subsequent taste perceptions". *Journal of Marketing*. 77(6): 124–139.

Pouyanfar, S., Y. Yang, S. C. Chen, M. L. Shyu, and S. S. Iyengar (2018). "Multimedia big data analytics: A survey". *ACM Computing Surveys (CSUR)*. 51(1): 1–34.

Pugh, S. D. (2001). "Service with a smile: Emotional contagion in a service encounter". *Academy of Management Journal*. 44(5): 1018–1027.

Raghubir, P. and A. Krishna (1999). "Vital dimensions in volume perception: Can the eye fool the stomach?" *Journal of Marketing Research*. 36: 313–326.

Rajaram, P. and P. Manchanda (2020). "Video influencers: Unboxing the mystique". arXiv preprint arXiv: 2012.12311.

Reik, T. (1949). *Character Analysis*. New-York: Farrar, Strauss and Giroux.

Rezlescu, C., B. Duchaine, C. Y. Olivola, and N. Chater (2012). "Unfakeable facial configurations affect strategic choices in trust games with or without information about past behavior". *PloS One*. 7(3): e34293.

Rooderkerk, R. P. and D. R. Lehmann (2021). "Incorporating consumer product categorizations into shelf layout design". *Journal of Marketing Research*. 58(1): 50–73.

Rule, N. O. and N. Ambady (2008). "The face of success: Inferences from chief executive officers' appearance predict company profits". *Psychological Science*. 19(2): 109–111.

Scherer, K. R. (1995). "Expression of emotion in voice and music". *Journal of Voice*. 9(3): 235–248.

Schroeder, J. and N. Epley (2015). "The sound of intellect speech reveals a thoughtful mind, increasing a job candidate's appeal". *Psychological Science*. 26(6): 877–891.

Schuller, B., B. Vlasenko, F. Eyben, M. Wöllmer, A. Stuhlsatz, A. Wendemuth, and G. Rigoll (2010). "Cross-corpus acoustic emotion recognition: Variances and strategies". *IEEE Transactions on Affective Computing*. 1(2): 119–131.

Schweinberger, S. R., H. Kawahara, A. P. Simpson, V. G. Skuk, and R. Zäske (2014). "Speaker perception". *Wiley Interdisciplinary Reviews: Cognitive Science*. 5(1): 15–25.

Shen, T., H. Kim, and M. Ding (2021). "Voice-based empathetic method for enhancing voice technology applications in marketing with certainty/uncertainty". *Working Paper*.

Shin, D., S. He, G. M. Lee, A. B. Whinston, S. Centintas, and K. C. Lee (2020). "Enhancing social media analysis with visual data analytics: A deep learning approach". *Management Information Systems Quarterly*. 44(4): 1459–1492.

Simonyan, K. and A. Zisserman (2014). "Very deep convolutional networks for large-scale image recognition". *Computer Science*.

Singh, S., D. Marinova, J. Singh, and K. R. Evans (2018). "Customer query handling in sales interactions". *Journal of the Academy of Marketing Science*. 46(5): 837–856.

Small, D. A. and N. M. Verrochi (2009). "The face of need: Facial emotion expression on charity advertisements". *Journal of Marketing Research*. 46(6): 777–787.

Smith, A. N., E. Fischer, and C. Yongjian (2012). "How does brand-related user-generated content differ across YouTube, Facebook, and Twitter?" *Journal of Interactive Marketing*. 26(2): 102–113.

Smith, B. L., B. L. Brown, W. J. Strong, and A. C. Rencher (1975). "Effects of speech rate on personality perception". *Lang Speech*. 18(2): 145–152.

Sundar, A. and T. J. Noseworthy (2014). "Place the logo high or low? Using conceptual metaphors of power in packaging design". *Journal of Marketing*. 78(5): 138–151.

Sutton, R. and A. G. Barto (2018). *Reinforcement Learning: An Introduction*. MIT Press.

Székely, E., Z. Ahmed, S. Hennig, J. P. Cabral, and J. Carson-Berndsen (2014). "Predicting synthetic voice style from facial expressions. An application for augmented conversations". *Speech Communication*. 57: 63–75.

Tahon, M. and L. Devillers (2015). "Towards a small set of robust acoustic features for emotion recognition: Challenges". *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 24(1): 16–28.

Tegtmeier, S., O. Niebuhr, and T. Schweisfurth (2021). "How money talks: Speech melody and venture evaluation". Available at SSRN 3795093.

Teixeira, T., R. Picard, and R. El Kaliouby (2014). "Why, when, and how much to entertain consumers in advertisements? A web-based facial tracking field study". *Marketing Science*. 33(6): 809–827.

Teixeira, T., M. Wedel, and R. Pieters (2012). "Emotion-induced engagement in internet video ads". *Journal of Marketing Research*. 49(2): 144–159.

Tellis, G. J., D. J. MacInnis, S. Tirunillai, and Y. Zhang (2019). "What drives virality (sharing) of online digital content? The critical role of information, emotion, and brand prominence". *Journal of Marketing*. 83(4): 1–20.

Texeira, T., M. Wedel, and R. Pieters (2012). "Emotion-induced engagement in internet video ads". *Journal of Marketing Research*. 49(2): 144–159.

Todorov, A. (2017). *Face Value: The Irresistible Influence of First Impressions*. Princeton University Press.

Todorov, A., A. N. Mandisodza, A. Goren, and C. C. Hall (2005). "Inferences of competence from faces predict election outcomes". *Science*. 308(5728): 1623–1626.

Townsend, C. and S. B. Shu (2010). "When and how aesthetics influences financial decisions". *Journal of Consumer Psychology*. 20: 452–458.

Urban, G. L., G. Liberali, E. MacDonald, R. Bordley, and J. R. Hauser (2014). "Morphing banner advertising". *Marketing Science*. 33(1): 27–46.

Vernon Richard, J. W., C. A. M. Sutherland, A. W. Young, and T. Hartley (2014). "Modeling first impressions from highly variable facial images". *Proceedings of the National Academy of Sciences*. 111(32): 3353–3361.

Voulodimos, A., N. Doulamis, A. Doulamis, and E. Protopapadakis (2018). "Deep learning for computer vision: A brief review". *Computational Intelligence and Neuroscience*.

Wang, X. S., J. He, and R. Grewal (2020). "Image features and demand in the sharing economy: A study of airbnb". *Working Paper*.

Wang, X. S., S. Lu, X. I. Li, M. Khamitov, and N. Bendle (2021). "Audio mining: The role of vocal tone in persuasion". *Journal of Consumer Research*. 48(2): 189–211.

Wang, Z., S. N. Singh, Y. J. Li, S. Mishra, M. Ambrose, and M. Biernat (2017). "Effects of employees' positive affective displays on customer loyalty intentions: An emotions-as-social-information perspective". *Academy of Management Journal*. 60(1): 109–129.

Wedel, M. and P. K. Kannan (2016). "Marketing analytics for data-rich environments". *Journal of Marketing*. 80(6): 97–121.

Wedel, M. and R. Pieters (2008). "A review of eye-tracking research in marketing". In: *Review of Marketing Research*. Ed. by N. K. Malhotra. Vol. 4. 123–147.

Weninger, F., J. Krajewski, A. Batliner, and B. Schuller (2012). "The voice of leadership: Models and performances of automatic analysis in online speeches". *IEEE Transactions on Affective Computing*. 3(4): 496–508.

Whittle, M. W. (2014). *Gait Analysis: An Introduction*. Butterworth-Heinemann.

Xiao, L. and M. Ding (2014). "Just the faces: Exploring the effects of facial features in print advertising". *Marketing Science*. 33(3): 338–352.

Xiao, L., H.-J. Kim, and M. Ding (2013). "An introduction to audio and visual research and applications in marketing". In: *Review of Marketing Research*. Ed. by N. Malhotra. Bingley: Emerald Group Publishing Limited. 213–253.

Xu, J. and M. Ding (2021). "Transparent model of unabridged data (TMUD)". Available at SSRN 3849871.

Yan, X., H. Gong, Y. Jiang, S.-T. Xia, F. Zheng, X. You, and L. Shao (2020). "Video scene parsing: An overview of deep learning methods and datasets". *Computer Vision and Image Understanding*. 201: 103077.

Yaoyuneyong, G., J. Foster, E. Johnson, and D. Johnson (2016). "Augmented reality marketing: Consumer preferences and attitudes toward hypermedia print ads". *Journal of Interactive Advertising*. 16(1): 16–30.

Yilmaz, A., O. Javed, and M. Shah (2006). "Object tracking: A survey". *ACM Computing Surveys (CSUR)*. 38(4). 13-es.

Young, A. W., S. Frühholz, and S. R. Schweinberger (2020). "Face and voice perception: Understanding commonalities and differences". *Trends in Cognitive Sciences*. 24(5): 398–410.

Zebrowitz, L. A., J. A. Hall, N. A. Murphy, and G. Rhodes (2002). "Looking smart and looking good: Facial cues to intelligence and their origins". *Personality and Social Psychology Bulletin*. 28(2): 238–249.

Zhang, M. and L. Luo (2022). "Can consumer-posted photos serve as a leading indicator of restaurant survival? Evidence from yelp". Forthcoming at *Management Science*.

Zhang, Q., W. Wang, and Y. Chen (2020). "Frontiers: In-consumption social listening with moment-to-moment unstructured data: The case of movie appreciation and live comments". *Marketing Science*. 39(2): 285–295.

Zhang, S., D. Lee, P. Singh, and K. Srinivasan (2022). "What makes a good image? Airbnb demand analytics leveraging interpretable image features". *Management Science*. Forthcoming.

Zhang, W., P. K. Chintagunta, and M. U. Kalwani (2021). "Social media, influencers, and adoption of an eco-friendly product: Field experiment evidence from rural China". *Journal of Marketing*. 85(3): 10–27.

Zhang, X., S. Li, R. R. Burke, and A. Leykin (2014). "An examination of social influence on shopper behavior using video tracking data". *Journal of Marketing*. 78(5): 24–41.

Zhang, Z. (2016). "Mechanics of human voice production and control". *The Journal of the Acoustical Society of America*. 140(4): 2614–2635.

Zhou, Y., S. Lu, and M. Ding (2020). "Contour-as-face framework: A method to preserve privacy and perception". *Journal of Marketing Research*. 57(4): 617–639.

Zoghaib, A. (2019). "Persuasion of voices: The effects of a speaker's voice characteristics and gender on consumers' responses". *Recherche et Applications en Marketing (English Edition)*. 34(3): 83–110.