

Baier, Daniel; Daniel, Ines; Frost, Sarah; Naundorf, Robert

Image data analysis and classification in marketing. (English) Zbl 1256.62003
Adv. Data Anal. Classif., ADAC 6, No. 4, 253-276 (2012).

Summary: Nowadays, the diffusion of smartphones, tablet computers, and other multipurpose equipment with high-speed Internet access makes new data types available for data analysis and classification in marketing. So, e.g., it is now possible to collect images/snaps, music, or videos instead of ratings. With appropriate algorithms and software at hand, a marketing researcher could simply group or classify respondents according to the content of uploaded images/snaps, music, or videos. However, appropriate algorithms and software are sparsely known in marketing research up to now. The paper tries to close this gap. Algorithms and software from computer science are presented, adapted and applied to data analysis and classification in marketing. The new SPSS-like software package IMADAC is introduced.

MSC:

- 62-07 Data analysis (statistics) (MSC2010)
- 90B60 Marketing, advertising
- 62-04 Software, source code, etc. for problems pertaining to statistics
- 62H30 Classification and discrimination; cluster analysis (statistical aspects)
- 62H35 Image analysis in multivariate analysis
- 65C60 Computational problems in statistics (MSC2010)

Cited in 1 Document

Keywords:

[image classification](#); [market segmentation](#); [speech recognition](#)

Software:

[IMADAC](#); [rapidminer](#); [SAS](#); [SURF](#); [Stata](#); [SIFT](#); [SPSS](#); [EMD](#)

Full Text: [DOI](#)

References:

- [1] Acock A (2005) SAS, Stata, SPSS: a comparison. *J Marriage Fam* 67(4):1093–1095 · doi:10.1111/j.1741-3737.2005.00196.x
- [2] ADM (2010) Jahresbericht 2010. Arbeitskreis deutscher Markt- und Sozialforschungsinstitute (ADM) e.V., Frankfurt/Main
- [3] Ahmad N, Omar A, Ramayah T (2010) Consumer lifestyles and online shopping continuance intention. *Bus Strateg Ser* 11(4):227–243 · doi:10.1108/17515631011063767
- [4] Albatineh A, Niewiadomska-Bugaj M (2011) Correcting jaccard and other similarity indices for chance agreement in cluster analysis. *Adv Data Anal Classif* 5(3):179–200 · Zbl 1274.62414 · doi:10.1007/s11634-011-0090-y
- [5] Anderson W, Golden L (1984) Lifestyle and psychographics: a critical review and recommendation. *Adv Consum Res* 11:405–411
- [6] Arabie P, Hubert L (1995) Advances in cluster analysis relevant to marketing research. *Stud Classif Data Anal Knowl Org* 6:3–19
- [7] ARD/ZDF (2011) ARD/ZDF-Onlinestudie 2011. ARD/ZDF. <http://www.ard-zdf-onlinestudie.de>
- [8] Baier D (2003) Classification and marketing research. *Taksonomia* 10:21–39
- [9] Baier D, Brusich M (2008) Marktsegmentierung. In: Herrmann A, Homburg C, Klarmann M (eds) *Handbuch Marktforschung. Methoden Anwendungen Praxisbeispiele*, Gabler, pp 769–790
- [10] Baier D, Daniel I (2012) Image clustering for marketing purposes. *Stud Classif Data Anal Knowl Org* 43:487–494 · Zbl 06014803 · doi:10.1007/978-3-642-24466-7_49
- [11] Baier D, Gaul W (1999) Optimal product positioning based on paired comparison data. *J Econ* 89(1):365–392 · Zbl 0936.62128 · doi:10.1016/S0304-4076(98)00067-0
- [12] Baier D, Stüber E (2010) Acceptance of recommendations to buy in online retailing. *J Retail Consum Serv* 17(3):173–180 · doi:10.1016/j.jretconser.2010.03.005
- [13] Bay H, Ess A, Tuytelaars T, Van Gool L (2008) Speeded-up robust features (SURF). *Comput Vis Image Understand* 110(3):346–359 · Zbl 05363027 · doi:10.1016/j.cviu.2007.09.014

- [14] Bearden W, Netemeyer R, Teel J (1989) Measurement of consumer susceptibility influence. *J Consum Res* 15:473–481 · doi:10.1086/209186
- [15] Buckinx W, Van den Poel D (2005) Customer base analysis: partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting. *Eur J Oper Res* 164(1):252–268 · Zbl 1132.90349 · doi:10.1016/j.ejor.2003.12.010
- [16] Burget R, Karasek J, Smekal Z, Uher V, Dostal O (2010) RapidMiner image processing extension: a platform for collaborative research. In: *The 33rd international conference on telecommunication and signal processing, TSP*, pp 114–118
- [17] BVM (2011) *Marktforschung 2011/2012–BVM Handbuch der Institute und Dienstleister*. Berufsverband Deutscher Markt- und Sozialforscher (BVM) e.V., Berlin
- [18] Choras R (2007) Image feature extraction techniques and their applications for CBIR and biometrics systems. *Int J Biol Biomed Eng* 1(1):6–16
- [19] Daniel I, Baier D (2012) Lifestyle segmentation based on contents of preferred images versus ratings of items. *Stud Classif Data Anal Knowl Org* (to appear)
- [20] Datta R, Joshi D, Li J, Wang J (2008) Image retrieval: ideas, influences, and trends of the new age. *ACM Comput Surv* 40(2):5:1–60
- [21] Duda RO, Hart PE (1973) *Pattern classification and scene analysis*. Wiley, New York · Zbl 0277.68056
- [22] Fleuret F, Geman D (2008) Stationary features and cat detection. *J Mach Learn Res* 9:2549–2578 · Zbl 1225.68259
- [23] Foley JD, van Dam A, Feiner SK, Hughes JF (1999) *Computer graphics: principles and practice*. Addison-Wesley Publishing Company, Reading
- [24] Frost S, Baier D (2012) Using earth mover’s distance and its approximations for clustering images—a comparison. *Stud Classif Data Anal Knowl Org* (to appear)
- [25] Fu KS, Rosenfeld A (1976) *Pattern recognition and image processing*. IEEE Trans Comput 25(12):1336–1346 · Zbl 0348.68057
- [26] Gaul W, Baier D (1994) *Marktforschung und Marketing Management: Computerbasierte Entscheidungsunterstützung*. Oldenbourg, München
- [27] Gaul W, Schmidt-Thieme L (2002) Recommender systems based on user navigational behavior in the internet. *Behaviormetrika* 29(1):1–22 · Zbl 1033.68002 · doi:10.2333/bhmk.29.1
- [28] Gonzalez A, Bello L (2002) The construct lifestyle in market segmentation—the behavior of tourist consumers. *Eur J Mark* 36:51–85 · doi:10.1108/03090560210412700
- [29] Gonzalez R, Woods R (2002) *Digital image processing*. Prentice-Hall, Englewood Cliffs
- [30] Haralick R, Shanmugam K, Dinstein I (1973) Textural features for image classification. *IEEE Trans Syst Man Cybern* 3(6):610–621 · doi:10.1109/TSMC.1973.4309314
- [31] Hu M (1962) Visual pattern recognition by moment invariants. *IRE Trans Inf Theory* 8(2):179–187 · Zbl 0102.13304
- [32] Hubert L, Arabie P (1985) Comparing partitions. *J Classif* 2:193–218 · Zbl 0587.62128 · doi:10.1007/BF01908075
- [33] IBM (2009) IBM to acquire SPSS Inc. to provide clients predictive analytics capabilities. International Business Machines (IBM) Corp, News Release July 28, 2009
- [34] IBM (2011) IBM SPSS Statistics 20 core system user’s guide. <http://www.spss.com>
- [35] Kapferer JN (1995) Brand confusion: empirical study of a legal concept. *Psychol Mark* 12(6):551–569 · doi:10.1002/mar.4220120607
- [36] Kato T (1992) Database architecture for content-based image retrieval. In: Jambardino AA, Niblack WR (eds) *Image storage and retrieval systems*, Proc. SPIE 1662, San Jose, pp 112–123
- [37] KDNuggets (2011) KDNuggets polls. www.kdnuggets.com/polls
- [38] Kim Y, Street W (2004) An intelligent system for customer targeting: a data mining approach. *Decis Support Syst* 37(2):215–228 · doi:10.1016/S0167-9236(03)00008-3
- [39] Law M, Figueiredo M, Jain A (2004) Simultaneous feature selection and clustering using mixture model. *IEEE Trans Pattern Anal Mach Intell* 26(9):1154–1166 · Zbl 05112235 · doi:10.1109/TPAMI.2004.71
- [40] Lazer W (1963) Life style concepts and marketing. In: Greyser A (ed) *Toward Scientific Marketing*. American Marketing Association, Chicago, pp 130–139
- [41] Lee HJ, Lim H, Jolly L, Lee J (2009) Consumer lifestyles and adoption of high-technology products: a case of south korea. *J Int Consum Mark* 21(2):153–167 · doi:10.1080/08961530802153854
- [42] Ling H, Okada K (2007) An efficient earth mover’s distance algorithm for robust histogram comparison. *IEEE Trans Pattern Anal Mach Intell* 29(5):840–853 · Zbl 05340831 · doi:10.1109/TPAMI.2007.1058
- [43] Liu Y, Zhang D, Lu G, Ma WY (2007) A survey of content-based image retrieval with high-level semantics. *Pattern Recognit* 40:262–282 · Zbl 1103.68503 · doi:10.1016/j.patcog.2006.04.045
- [44] Lowe D (1999) Object recognition from local scale-invariant features. In: *The proceedings of the seventh IEEE international conference on computer vision*, vol 2, pp 1150–1157
- [45] Lowe D (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60(2):91–110 · Zbl 02244065 · doi:10.1023/B:VISI.0000029664.99615.94
- [46] MacCallum R (1983) A comparison of factor analysis programs in SPSS, BMDP, and SAS. *Psychometrika* 48(2):223–231 · Zbl 0534.62039 · doi:10.1007/BF02294017
- [47] Maheshwari M, Silakari S, Motwani M (2009) Image clustering using color and texture. In: *First international conference on*

computational intelligence, communication systems and networks, Indore, pp 403–408

- [48] Niemann H (1983) Klassifikation von Mustern. Springer, Berlin · [Zbl 0537.68084](#)
- [49] Nunnally J (1978) Psychometric theory. McGraw-Hill, New York
- [50] Pele O, Werman M (2009) Fast and robust earth mover's distances. In: IEEE 12th international conference on computervision, vol 12, no. 1, pp 460–467
- [51] Peleg S, Werman M, Rom H (1989) A unified approach to the change of resolution: space and gray-level. *IEEE Trans Pattern Anal Mach Intell* 11(7):739–742 · [Zbl 05111841](#) · [doi:10.1109/34.192468](#)
- [52] Plummer J (1974) The concept and application of life style segmentation. *J Mark* 38(1):33–37 · [doi:10.2307/1250164](#)
- [53] Punj G, Stewart DW (1983) Cluster analysis in marketing research: review and suggestions for application. *J Mark Res* 20:134–148 · [doi:10.2307/3151680](#)
- [54] Puzicha J, Buhmann J, Rubner Y, Tomasi C (1999) Empirical evaluation of dissimilarity measures for color and texture. In: Proceedings of the seventh IEEE international conference on computer vision, vol 2, pp 1165–1172 · [Zbl 1021.68787](#)
- [55] Rapid-I (2010) Rapidminer 5.0 user manual. <http://www.rapid-i.com>
- [56] Resankova H, Husek D (2002) Comparison of SAS, SPSS and STATISTICA systems in the area of clustering variables. *Comput Stat Data Anal* 41(2):331–339
- [57] Rexer K (2010) 4th annual rexa analytics data miner survey. In: Predictive analytics world, Oct. 2010. Washington, D.C
- [58] Rosenfeld A (1969) Picture processing by computer. Computer science and applied mathematics. Academic Press, New York · [Zbl 0182.50603](#)
- [59] Rubner Y, Tomasi C (2001) Perceptual metrics for image database navigation. Kluwer Academic Publishers, Boston · [Zbl 0973.68661](#)
- [60] Rubner Y, Guibas L, Tomasi C (1997) The earth mover's distance, multi-dimensional scaling, and color-based image retrieval. In: Proceedings of the ARPA image understanding workshop, pp 661–668
- [61] Rubner Y, Tomasi C, Guibas L (2000) The earth mover's distance as a metric for image retrieval. *Int J Comput Vis* 40(2):99–121 · [Zbl 1012.68705](#) · [doi:10.1023/A:1026543900054](#)
- [62] Rui Y, Huang T, Chang SF (1999) Image retrieval: current techniques, promising directions, and open issues. *J Vis Commun Image Represent* 10(1):39–62 · [doi:10.1006/jvci.1999.0413](#)
- [63] Schmitt I (2005) Ähnlichkeitssuche in Multimedia-Datenbanken–retrieval. Suchalgorithmen und Anfragebehandlung, Oldenbourg · [Zbl 1100.68018](#)
- [64] Schwaiger M (2006) Wirkungskontrolle kommunikationspolitischer Maßnahmen. In: Tomczak T (ed) Reinecke S. Handbuch Marketingcontrolling, Gabler, pp 521–548
- [65] Schwarz M, Cowan W, Beatty J (1987) An experimental comparison of RGB, YIQ, LAB, HSV, and opponent color models. *ACM Trans Graph* 6(2):123–158 · [doi:10.1145/31336.31338](#)
- [66] Sebe N, Lew M (2001) Texture features for content-based retrieval. In: Lew M (ed) Principles of visual information retrieval. Springer, London, pp 51–85
- [67] Serratos F, Sanroma G (2008) A fast approximation of the earth-movers distance between multidimensional histograms. *Int J Pattern Recognit Artif Intell* 22(8):1539–1558 · [doi:10.1142/S0218001408006880](#)
- [68] Sinus (2009) Informationen zu den Sinus-Milieus 2009. Sinus Sociovision GmbH, Heidelberg
- [69] Smith W (1956) Product differentiation and market segmentation as alternative marketing strategies. *J Mark* 21:3–8 · [doi:10.2307/1247695](#)
- [70] Swain M, Ballard D (1991) Color indexing. *Int J Comput Vis* 7(1):11–32 · [doi:10.1007/BF00130487](#)
- [71] Tamura H, Mori S, Yamawaki T (1978) Texture features corresponding to visual perception. *IEEE Trans Syst Man Cybern* 8(6):460–473 · [doi:10.1109/TSMC.1978.4309999](#)
- [72] Tuma M, Decker R, Scholz S (2011) A survey of the challenges and pitfalls of cluster analysis application in market segmentation. *Int J Mark Res* 53(3):391–415 · [doi:10.2501/IJMR-53-3-391-414](#)
- [73] van Horen F, Pieters R (2012) When high-similarity copycats lose and moderate-similarity copycats gain: the impact of comparative evaluation. *J Mark Res* 49(2):83–91 · [doi:10.1509/jmr.08.0405](#)
- [74] van House (2007) Flickr and public image-sharing: distant closeness and photo exhibition. In: Conference on human factors in computing systems, San Jose, pp 2717–2722
- [75] Venables W, Smith D, Team RDC (2011) An introduction to R: notes on R. Department of Statistics and Mathematics, Vienna University of Economics and Business, Vienna, A Programming Environment for Data Analysis and Graphics
- [76] Wahbeh A, Al-Radaideh Q, Al-Kabi M, Al-Shawakfa E (2011) A comparison study between data mining tools over some classification methods. *Int J Adv Comput Sci Appl Spec Issue* 3:18–26
- [77] Wedel M, Kamakura W (2000) Market segmentation: conceptual and methodological foundations. Kluwer, Dordrecht
- [78] Weihs C, Ligges D, Mörchen F, Müllensiefen D (2007) Classification in music research. *Adv Data Anal Classif* 1(3):255–291 · [Zbl 1183.62109](#) · [doi:10.1007/s11634-007-0016-x](#)
- [79] Wells W, Tigert D (1971) Activities, interests and opinions. *J Advert Res* 11:27–35
- [80] Wyszecki G, Stiles W (1982) Color science. Concepts and methods, quantitative data and formulae, 2nd edn. Wiley, New York

- [81] Yankelovich D (1964) New criteria for market segmentation. *Harvard Bus Rev* 42(March–April):83–90
- [82] Yankelovich D, Meer D (2006) Rediscovering market segmentation. *Harvard Bus Rev* 84(2):122–131
- [83] Zellhöfer D, Schmitt I (2009) A preference-based approach for interactive weight learning—learning weights within a logic-based query language. *Distrib Parallel Databases* 27(1):31–51 · [Zbl 05661026](#) · [doi:10.1007/s10619-009-7049-4](#)

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.