

Distanzmaße in der Bildähnlichkeitsanalyse – Neue Verfahren und deren Anwendung im Marketing

Bilder haben für das Marketing seit jeher eine große Bedeutung, die zudem aufgrund der Verbreitung digitaler Bilder im Internet ständig wächst. In sozialen Netzwerken, wie Facebook oder Twitter, werden täglich viele Millionen Bilder hochgeladen. Allerdings wird diese Datenflut von der Marktforschung noch kaum verwendet, da Methoden zur systematischen, feature-basierten Analyse der Bildbestände im Marketingkontext fehlen.

Die Voraussetzung für eine Klassifikation von großen Bildbeständen, wie sie für Marketinganwendungen eingesetzt werden kann, ist die Durchführung einer Ähnlichkeitsanalyse. Die Ergebnisse einer solchen Analyse hängen im Wesentlichen von den verwendeten Distanzmaßen ab. Dabei spielen neben der Qualität der Ergebnisse auch die Laufzeit der Algorithmen sowie die Robustheit bezüglich unterschiedlicher Bildqualitäten eine maßgebliche Rolle.

In der vorliegenden Arbeit wurde geprüft, welche Distanzmaße hinsichtlich Robustheit, Rechenaufwand und Klassifikationsergebnis geeignet sind, um große Bildbestände für die Marktforschung nutzbar zu machen. Dazu wurden eine Reihe empirischer Untersuchungen, wie eine Monte-Carlo-Analyse, ein Verwechslungstest sowie zahlreiche feature-basierte Bildähnlichkeitsanalysen, durchgeführt und beschrieben. Ein eigens entwickelter Algorithmus, der „Increasing Radius Algorithmus“ zur Approximation der, in der Bildanalyse weit verbreiteten, „Earth Mover’s Distance“, tat sich dabei besonders hervor. Dieser Algorithmus stellte sich als besonders robust bezüglich gängiger Qualitätsmängel in der digitalen Fotografie heraus. Diese können z. B. Überbelichtung, Rauschen oder eine zu geringe Auflösung sein.

Die vorliegende Arbeit zeigt eine Vielzahl von Forschungsbereichen im Marketing auf, für die die Bildähnlichkeitsanalyse in Zukunft relevant sein kann. Vor allem die Markenpositionierung anhand von Werbeanzeigen sowie die Verwechslung von Marken aufgrund visuell ähnlicher Anzeigen wurden hierzu näher beleuchtet. Die gewonnen Erkenntnisse sind aber auch für Gebiete wie die Lebensstilsegmentierung, Kundensegmentierung, die Suche von geeigneten Kampagnenbildern, die Neuproduktentwicklung oder die Unterscheidung von Käufern und Nichtkäufern vielversprechend. Ebenso ist zu erwarten, dass in Zukunft häufiger Bilder in Umfragen genutzt werden, welche mit den vorgestellten Methoden automatisiert ausgewertet werden können.

Distance measures in the analysis of image similarity – New approaches and their application in marketing

The importance of images in marketing has always been high and is still increasing because of the growing number of digital images uploaded to the Internet. Millions of photos are uploaded to social networks like Facebook or Twitter every day. But this flood of data has not been used for marketing approaches yet. Because there are no systematic, feature-based methods for analysing image databases in context of marketing.

The requirement for classifying huge image data bases, in a way that is useful for marketing, is the implementation of a similarity analysis. The results of such an analysis depend mostly on the used distance measures. Besides of the quality of the clustering results also the computation time and the robustness with regard to the image quality are important.

This work verifies which distance measures are appropriate to classify huge image data bases with regard to robustness, computation time and clustering results with the aim to make them usable for marketing approaches. Therefore a lot of empirical studies had to be implemented and are described in this book, like a Monte Carlo analysis, a confusion test, and many feature-based image similarity analyses. There was one algorithm that was especially robust against typical failures in the digital photography, like noise or overexposure: the 'increasing radius algorithm', which is an approximation of the well-known 'earth mover's distance' for image similarity analysis.

There are a lot of applications in marketing for which the image similarity analysis could be relevant in the future. This book especially shows how the method of using visual image information could be used for analysing brand positioning and brand confusion. But the obtained findings can also be used in areas like lifestyle segmentation, customer segmentation, searching for campaign images, new product development, or to distinguish buyers from non-buyers. It is also very likely that there will be more images used for questionnaires in the future.