

Die Kunst als Ganzes. Heterogene Bilddatensätze als Herausforderung für die Kunstgeschichte und die Computer Vision.

Dieckmann, Lisa

lisa.dieckmann@uni-koeln.de

Universität zu Köln, Kunsthistorisches Institut

Bell, Peter

bell@uni-heidelberg.de

Ruprecht-Karls-Universität Heidelberg, Heidelberg

Collaboratory for Image Processing (HCI)

Den aufwändig nach bestimmten Schemata modellierten Datensätzen fachspezifischer Bilddatenbanken und der Verwendung von Ontologien, Thesauri und Normdaten steht die anschwellende Bilderflut des digitalen und vernetzten Zeitalters gegenüber. Auch im Bereich kulturhistorischer Artefakte liegen viele digitale Abbildungen in völlig unterschiedlichen Formen, mal mit Text und Metadaten, mal mehr oder weniger strukturiert und erschlossen vor. Dieser positiven Entwicklung einer Demokratisierung des Forschungsmaterials durch die digitale, vernetzte und a-hierarchische Bereitstellung, muss jedoch mit Verfahren begegnet werden, welche die heterogenen Bilder nach bestimmten Kriterien automatisiert erschließen und strukturieren und das Retrieval im Sinne von Precision und Recall optimieren.

Mit welchen Verfahren lässt sich die große Menge von Bildinhalten auffindbar und erschließbar machen? Wie kann der Heterogenität der Bild- und Metadaten begegnet werden? Und: Wie kann dennoch der Individualität von historisch-hermeneutischer Forschung Rechnung getragen werden? Neben automatisierten textanalytischen Verfahren, mit denen die Daten strukturiert, kategorisiert oder auch angereichert werden, können Bilder semantisch über rein bildanalytische Verfahren erschlossen werden. Darin besteht die größere und eigentliche Herausforderung. Denn das Bild ist im Vergleich zum Text schwerer in sinnhaltige und abgrenzbare Entitäten zerlegbar und seine Ikonographie nicht ohne weiteres über visuelle automatisierte Verfahren rekonstruierbar. Darin liegt der „semantic gap“.

Der Vortrag thematisiert die Erschließung von größeren und nur teilweise strukturierten Mengen von Bildern und Metadaten durch automatisierte Verfahren. Der Schwerpunkt liegt hierbei auf Verfahren zur visuellen Analyse. Anhand des heterogenen und verteilten Bilderpools des prometheus-Bildarchivs (prometheus

2001-2016) wird derzeit mit der Computer Vision Group der Universität Heidelberg (CompVis 2015) ein Projekt vorbereitet, in welchem die bereits sehr elaborierten Ansätze Anwendung finden.

Das prometheus-Bildarchiv (Universität zu Köln, Kunsthistorisches Institut) ist ein verteiltes digitales Bildarchiv, das über 80 strukturell und inhaltlich verschiedene Bilddatenbanken der Kunstgeschichte, Archäologie und weiterer bildbasierter Disziplinen und damit über 1,4 Mio. Bilder miteinander verknüpft.

Die Computer Vision Group des Heidelberg Collaboratory for Image Processing an der Ruprecht-Karls-Universität Heidelberg widmet sich der Grundlagenforschung zum automatischen Bildverstehen. Sie entwickelt Algorithmen zur Erschließung von Bildbestandteilen (Segmentierung), diskriminativer Objekterkennung und Szenenvergleich. Sie hat bereits in mehreren Projekten mit der Kunstgeschichte zusammengearbeitet.

Im vorzustellenden Projekt sollen Algorithmen zum automatischen Sehen und maschinellen Lernen speziell zur Erschließung und zur Recherche an kulturellem Erbe optimiert und entwickelt werden. Komplementär zur üblichen Textsuche werden die NutzerInnen des Bildarchivs nach Bildern und Bildpartien suchen können. Darüber hinaus sollen die Bilder aufgrund verschiedener Ähnlichkeiten sortiert und vorgeschlagen werden.

Für die Computer Vision bietet das prometheus-Bildarchiv ein großes Potential und viele skalierbare Aufgaben, da viele der Bilder in einem stilistischen oder semantischen Zusammenhang stehen, der sich visuell oder über die Metadaten erschließen lässt. Es ist somit keine Anwendungsaufgabe, sondern dezidiert informatische Grundlagenforschung. Bisher existieren nur wenige Versuche und Prototypen, wobei es sich im Wesentlichen um die Adaption von bestehenden Computer Vision Ansätzen auf kunsthistorische Spezialprobleme handelt.

Die skizzierte automatische Bildanalyse geht verschiedene Problemfelder an. Zum einen kann die Abbildung als Ganzes analysiert werden. Dabei sollen nicht nur Duplikate, sondern auch Reproduktionen und Variationen eines Kunstwerks aufgefunden werden, wie z. B. Kunstwerke, die im Laufe der Geschichte immer wieder rezipiert wurden (Laokoon-Gruppe, Apoll von Belvedere). Die automatische Bildanalyse muss daher verschiedene Dimensionen von Ähnlichkeit im Blick behalten.

Es können aber auch einzelne Bildpartien und Motive eines Bildes analysiert werden. In vielen Fällen sind diese Motive nicht verschlagwortet (Kreuz, Schädel, Pferd) und erst die visuelle Analyse der Bildinhalte liefert die gewünschten Ergebnisse. Auch die spezifische Formbehandlung oder der Duktus des Künstlers – Phänomene, die quasi nie textlich beschrieben werden, können so in einen Zusammenhang gebracht werden. Schließlich kann die Bildanalyse auch syntaktische Unebenheiten und variierende Schreibweisen (Bildtitel, Künstlernamen) ausgleichen. Als alternative Suchstrategie

kann die Bildsuche diese Übersetzungsschwierigkeiten umgehen.

Ein Kernproblem ist die methodische Herausforderung, Ähnlichkeit in verschiedenen Dimensionen zu definieren. Hier bieten sich einerseits zahlreiche stilkundliche und kennerschaftliche Ansätze an, die jedoch in ihrer Reichweite umstritten sind und noch nicht auf diese Weise technisch evaluiert wurden. Andererseits lassen sich semantische Taxonomien (ICONCLASS) und kontextbezogene Metadaten hinzuziehen, deren Verhältnis zum visuellen Befund erst bestimmt werden muss. Daraus ergibt sich auch ein großer Bedarf an methodischer Reflektion und interdisziplinärer Diskussion zwischen den beiden Bildwissenschaften. Weiterhin sind technische Schwierigkeiten durch die sehr große Bildmenge und die wissenschaftlichen Standards entsprechende tiefe Erschließung gegeben. Dies kann nur gewährleistet werden, wenn die visuelle Suchanfrage prägnant genug ist, aber auch die Variationsweite definiert wird. Da dies nur in wenigen Fällen mit einem Suchbereich funktioniert, muss ein interaktives Modell geschaffen werden, in dem sich Mensch und Maschine dialogisch dem gewünschten Ergebnis nähern. Die verschiedenen Dimensionen von Ähnlichkeit und kontextueller Bezogenheit sind ein anspruchsvolles Grundlagenforschungsproblem der Computer Vision, das nicht mit vorgefertigten Algorithmen oder bestehenden Usecases adaptiv gelöst werden kann.

Die NutzerInnen sollen mehrere Suchfelder markieren und kombinieren können, sowie die Ergebnisse bewerten, um mit Positiv- und Negativbeispielen eine neue Suche auszulösen. Dazu bedarf es eines ergonomischen Userinterfaces, in dem die NutzerInnen die neuartige Suche leicht erlernen und nutzen können. Der Algorithmus lernt fortwährend an den durchgeführten Suchen und Feedbacks, welche Suchen für die NutzerInnen relevant sind und erschließt dadurch kontinuierlich den Bilddatensatz.

Computer Vision kann also, indem direkt auf die Bildinformationen zugegriffen wird, Beschreibungen vornehmen und Verbindungen zwischen Kunstwerken aufzeigen, die vom menschlichen Auge nicht oder nur unter größtem Zeitaufwand gesehen werden können. Als Vorschlagssystem findet es nicht nur genau die Partien nach denen der Anwender sucht, sondern weist auch auf Bilder mit ähnlichen Formen, Texturen oder Farbwerten hin und visualisiert die ähnlichen Bilder in Form übersichtlicher Synopsen. Die Arbeit mit digitalen Bildrepositorien wird nicht nur effektiver, sondern auch assoziativer, durch das Vorschlagen neuer Verbindungen zwischen Kunstwerken. Dieser visuelle Zugang zu wissenschaftlichen Bilddatenbanken bereichert nicht nur die historischen Bildwissenschaften um ein Forschungsinstrument, sondern erleichtert auch den benachbarten Disziplinen (Archäologie, Kunstpädagogik etc.) einen intuitiven Zugriff auf das Material jenseits fachlicher Terminologie.

Bibliographie

Bell, Peter / Ommer, Björn (2015): "Training Argus, Ansätze zum automatischen Sehen in der Kunstgeschichte", in: *Kunstchronik* 68, 8: 414-420.

Bell, Peter / Ommer, Björn / Schlecht, Joseph (2013): "Nonverbal Communication in Medieval Illustrations Revisited by Computer Vision and Art History", in: *Visual Resources. An International Journal of Documentation* 29, 1-2: 26-37.

Bell, Peter / Dieckmann, Lisa / Ommer, Björn / Takami, Masato (2015): "Passion Search. Prototype of an unrestricted image search of the crucifixion", in: <http://hci.iwr.uni-heidelberg.de/COMPVIS/projects/suchpassion/> [letzter Zugriff 05. Januar 2016]

Chung, Joon Son / Arandjelovi#, Relja / Bergel, Giles/ Franklin, Alexandra/ Zisserman, Andrew (2014): "Representations of Art Collections", in: Agapito, Lourdes / Bronstein, Michael M. / Rother, Carsten (eds.): *Computer Vision - ECCV 2014 Workshops*. Heidelberg / New York / Dordrecht / London: Springer 85-100.

CompVis(2015): *Computer Vision Group*. Ruprecht-Karls-Universität Heidelberg <http://hci.iwr.uni-heidelberg.de/COMPVIS/> [letzter Zugriff 08. Januar 2016].

Crowley, Elliot J. / Zisserman, Andrew (2014): "In Search of Art", in: Agapito, Lourdes / Bronstein, Michael M. / Rother, Carsten (eds.): *Computer Vision - ECCV 2014 Workshops*. Heidelberg / New York / Dordrecht / London: Springer 54-70.

Dieckmann, Lisa (2015): "prometheus – das verteilte digitale Bildarchiv für Forschung & Lehre e. V.", in: Euler, Ellen / Hagedorn-Saupe, Monika/ Maier, Gerald/ Schweibenz, Werner/ Sieglerschmidt, Jörn (eds.): *Handbuch Kulturportale*. Online-Angebote aus Kultur und Wissenschaft. Berlin / Boston: DeGruyter 223-229.

Johnson, C. Richard Jr. / Wang, James Z. et al. (2008): "Image Processing for Artist Identification - Computerized Analysis of Vincent van Gogh's Painting Brushstrokes", in: *IEEE Signal Processing Magazine* 25, 4: 37-48.

Kohl, Jeanette / Srinivasan, Ramya / Roy-Chowdhury, Amit / Rudolph, Conrad (2013): "Quantitative Modeling of Artists Styles in Renaissance Face Portraiture", in: Second International Workshop on Historical Document Imaging and Processing <http://www.ee.ucr.edu/~amitrc/publications/icdar2013.pdf> [letzter Zugriff 05. Januar 2016]

prometheus (2001-2016)*prometheus – Das verteilte digitale Bildarchiv für Forschung & Lehre*. Kunsthistorisches Institut der Universität zu Köln <http://prometheus-bildarchiv.de/index> [letzter 08. Januar 2016].

Vaughan, William (1997) "Computergestützte Bildrecherche und Bildanalyse", in: Hubertus Kohle (ed.): *Kunstgeschichte digital*. Eine Einführung für Praktiker und Studierende. Berlin: Reimer 97-105.