

---

# Computer Vision in Digital Humanities

## Martijn Kleppe

[martijn.kleppe@kb.nl](mailto:martijn.kleppe@kb.nl)

National Library of the Netherlands, the Netherlands

## Matthew Lincoln

[mllincoln@getty.edu](mailto:mllincoln@getty.edu)

Getty Research Institute, United States of America

## Melvin Wevers

[m.j.h.f.wevers@uu.nl](mailto:m.j.h.f.wevers@uu.nl)

Utrecht University, the Netherlands

## Mark Williams

[mark.j.williams@dartmouth.edu](mailto:mark.j.williams@dartmouth.edu)

Dartmouth College, United States of America

## Benoit Seguin

[benoit.seguin@epfl.ch](mailto:benoit.seguin@epfl.ch)

École Polytechnique Fédérale de Lausanne, Switzerland

## Thomas Smits

[t.smits@let.ru.nl](mailto:t.smits@let.ru.nl)

Raboud University, the Netherlands

---

## Introduction

Although the majority of Digital Humanities scholars still focus on textual analysis, we see an increasing number of studies using digitised visual sources and taking the first steps in the field of ‘Visual Big Data’ (Ordelman et al, 2014). Scholars have increasingly used both large-scale digitised visual datasets as well as computational methods to analyse these datasets in new ways, see for example the work of Lev Manovich on Manga comics (2012), Time covers, and selfies (Manovich & Tifentale 2015) or the recent work of Lenardo et al (2016) on visual pattern discovery in large databases of paintings as presented at the DH 2016 conference. Others focus on the identification of finding recurring images in visual collections. Terras & Kirton (2013) analyse the reuse of digital images of cultural and heritage material on the internet while Resig (2014) and Reside (2015) follow a similar approach using a closed dataset.

A similar trend can be observed at the ADHO Digital Humanities Conferences. In 2014 the special Interest group Audiovisual Materials in Digital Humanities was founded and at the 2015 and 2016 conferences, several papers were presented using (audio) visual sources (see e.g. Kleppe 2015, Lenardo et al. 2016, Lincoln 2016). Based on an analysis of the submissions for the 2017 conference, Scott Weingart observes that an increasing number of papers are based on research concerning non-textual sources (Weingart 2017). The use of non-textual sources challenges established research practices of Digital Humanities, leading to new research questions and the need for transformation of existing approaches, and the development of new methodologies. For example, how can we analyse the characteristics of images in large visual datasets? Some of these questions have also been raised in the analogue era. See, for example, the work of Barry Salt (1974) on the characteristics of opening shots of twentieth-century films, Scott McCloud (1994) on the visual language of Japanese manga and comics, or Peter Burke’s take on using visual materials as historical evidence (2001).

Due to the digital turn, we now see the application of new techniques to answer similar and new types of research questions. One of the dominant techniques to analyse large-scale visual datasets is computer vision: a field that deals with how computers can generate a high-level of understanding of visual material (Smeulders 2000). This field has been seeing unprecedented improvements as part of the “Deep Learning Revolution” since the first deep Convolutional Neural Network (Krizhevsky 2012), unleashing new possibilities. This workshop will focus on how computer vision can be applied within the realm of Audiovisual Materials in Digital Humanities. During the workshop, attendees will both present (ongoing) work on applying computer vision and experiment with computer vision in their own work in a hands-on session. Possible questions include:

- What are state-of-art techniques in computer vision?
- What are possibilities and challenges in the application of computer vision in Digital Humanities research?
- How can non-expert researchers employ computer vision?
- How can cooperation between Humanities researchers and computer vision experts be improved so both can benefit of each other's expertise?

## Workshop outline

The workshop will consist of four pillars:

- **A keynote** by [Lindsay King](#) (Associate Director for Access and Research Services Robert B. Haas Family Arts Library, Yale University) & [Peter Leonard](#) (Director Digital Humanities Lab, Yale University Library) & on “Processing Pixels: Towards Visual Culture Computation”
- **Paper presentations** with results and ongoing work on applying computer vision in DH research projects. Papers will be selected by a review commission. All details on the Call for Abstracts can be found at [the workshop website](#).
- **A hands-on session** in which participants will be able to experiment with open source Computer Vision tools. This session will be led by [Benoit Seguin](#) of École Polytechnique Fédérale de Lausanne, (EPFL). Benoit Seguin is a PhD Student in Computer Science from the DHLAB at EPFL. His work focuses on automatic patterns detection across iconographic collections. Before starting his PhD, he worked with microscope images for biomedical applications or electronic manufacturing problems.
- **Lightning Talks** allowing participants to share their ideas, projects or ongoing work in a short presentation of two minutes. Interested researchers who want to present work will be able to submit their idea in a later stage.

## SIG Endorsement

This workshop proposal is endorsed by the steering committee of the Special Interest Group [Audiovisual Materials in the Digital Humanities \(AVinDH\)](#). This will be the third workshop that will be organised by the SIG. The [first workshop](#) on audiovisual material in digital humanities took place at the 2014 DH Conference in Lausanne. This workshop paved the way for the foundation of the [AVinDH Special Interest Group](#) that was officially launched during the DH2015 Conference in Sydney. [The second workshop](#) took place during the DH2016 Conference in Cracow and dealt with audiovisual material and dedicated analysis tools in DH.

While the previous SIG-endorsed workshops were organised by the SIG’s steering committee, this proposal on Computer Vision in Digital Humanities is

mainly organised by members of the SIG and one representative of the Steering Committee. The steering committee applauds the initiative of the organisers since the theme of the workshop fits perfect in both the SIG’s aims as well as the current interest in non-textual sources as observed by [Scott Weingart in his first analyses of the DH2017 submissions](#). Furthermore, the steering committee is enthusiastic about providing a platform for members of the SIG and the complete DH community to present high-quality work that will be selected by a review committee as well as ongoing work that will be presented in lightning talks. Both formats will contribute to the community and network building function that the SIG AVinDH promotes.

## Programme Committee

The programme committee (PC) for this workshop consists of the workshop proposers (see list with names and bios below). For the paper selection process, the PC will be extended with at least 10 experts. The selection process will be chaired by Prof. Franciska de Jong, Utrecht University/CLARIN ERIC and chair of the SIG’s Steering Committee ([f.m.g.dejong@uu.nl](mailto:f.m.g.dejong@uu.nl))

## Composition of the Programme Committee (workshop proposers only)

### *Dr. Martijn Kleppe*

Martijn Kleppe is a researcher at the Research Department of the National Library of the Netherlands (KB) where he works on several Digital Humanities Projects as part of the KB’s [Digital Humanities Team](#). Before he worked as academic researcher at the Erasmus University Rotterdam and Vrije Universiteit Amsterdam. He wrote his dissertation on ‘Canonical Iconic Photographs’ and was involved in several Digital Humanities projects focussing on opening up (audio)visual archives.

### *Melvin Wevers*

Melvin Wevers is in the final stage of his PhD research at Utrecht University on the role of the United States as a reference culture in twentieth-century Dutch newspaper debates on consumer goods. In his research, he primarily focuses on text analysis but he is moving in the direction of computational analyses of visual materials. He also worked on the eScience project *ShiCo: Shifting Concepts Over Time* in which he used word embeddings to study conceptual change in

large historical corpora. In Spring 2016, he was a research fellow of the 'culture analytics' program hosted by the UCLA's Institute of Pure and Applied Mathematics. In March 2017, he will be researcher-in-residence at the National Library of the Netherlands (KB) to work on the computational analyses of images in newspaper advertisement.

### Mark Williams

Mark Williams is Associate Professor of Film and Media Studies at Dartmouth College. He received both of his graduate degrees in Critical Studies from The School of Cinema-Television at The University of Southern California. He has published in a variety of journals and anthologies, including *Télévision: le moment expérimental (1935-1955)*; *Convergence Media History*; *New Media: Theories and Practices of Digitextuality*; *Collecting Visible Evidence*; *Dietrich Icon*; *Television, History, and American Culture: Feminist Critical Essays*; and *In Living Color: Race, Feminism, and Television*. He directed the Leslie Center Humanities Institute entitled [Cyber-Disciplinarity](#). In conjunction with the Dartmouth College Library, he is the founding editor of an e-journal, [The Journal of e-Media Studies](#). With Adrian Randolph, he co-edits the book series [Interfaces: Studies in Visual Culture](#) for the University Press of New England. With Michael Casey, he received an NEH Digital Humanities Start-Up Grant to build the [ACTION](#) toolset for cinema analysis. He received an award for Scholarly Innovation and Advancement at Dartmouth for directing [The Media Ecology Project \(MEP\)](#). He has published about MEP in [The Arclight Guidebook to Media History and The Digital Humanities](#) (2016), and also in *The Moving Image* (2016). Most recently Williams has received with John Bell (MEP architect) an NEH grant to build a Semantic Annotation Tool (SAT). With Lorenzo Torresani (Dartmouth Computer Studies) he has received a Knight Foundation grant to develop computer vision and machine learning capacities for moving images.

### Thomas Smits

Thomas Smits is completing a PhD on the transnational trade in illustrations of the news and the production of identity by mid-nineteenth century European illustrated newspapers at the Radboud University in the Netherlands. He is an editor of the *Journal for European Periodicals Research (JEPS)* and a PhD-board member of the Royal Netherlands Historical Society. In May 2017 he will start a researcher-in-residence project at the National Library of the Netherlands (KB) entitled *Illustrations to Photographs: using*

*computer vision to analyse news pictures in Dutch newspapers, 1860-1940.*

### Benoit Seguin

Benoit Seguin is a PhD student at the Digital Humanities Laboratory of EPFL. His work focuses on using modern computer vision algorithms to navigate large iconographic collections, and he recently showed how Neural Networks could be trained to recognize similar patterns in different artworks. He is part of the bigger REPLICA project which aims at digitizing and making searchable the 1M photos of the Cini collection in Venice. He took part of the organization of the recent "New Methods and Technologies for Art History" Summer School between ETHZ and EPFL. Before starting his PhD, he got his MSc in Computer Science with a thesis about SEM Image Analysis from IBM Research Zurich.

### Matthew Lincoln

Matthew Lincoln (Ph.D. University of Maryland 2016) is a Data Research Specialist with the Getty Research Institute, working on the Provenance Index Remodeling Project. He specializes in computationally-driven analysis of art history, including complex network analysis and machine learning methods on cultural datasets. He has previously worked as a curatorial fellow with the National Gallery of Art in Washington, DC, and has received grants from both the Getty and Kress Foundations as part of digital art history summer institutes, and in 2016 was an organizer for a jointly-funded conference "Digital Dimensions of Art History" hosted by the University of Maryland and the Maryland Institute for Technology and Humanities. He has published data-driven research articles in the *International Journal of Digital Art History* and *British Art Studies*, and is a contributor to *The Programming Historian*.

### Target audience

This workshop aims to bring together scholars from different fields who have an interest in or actively use computer vision to analyse (large) datasets of digitised audiovisual sources. We have a wide network of peers that can be reached, and will also scan the papers of the conference that have been selected with the aim to invite potential contributors in this realm to present their work in our workshop. We aim at participation of 20 – 30 persons.

### Bibliography

- Burke, P.** (2001). *Eyewitnessing. The uses of images as historical evidence.* London: Cornell University Press
- Kleppe, M.** (2015) Tracing the afterlife of iconic photographs using IPTC. *Digital Humanities 2015*, 29 Juni - 3 Juli 2015, Sydney
- Krizhevsky, A., Sutskever, I. and Hinton, G. E.** (2012). ImageNet classification with deep convolutional neural network
- di Lenardo, I., Seguin, B., Kaplan, F.** (2016). Visual Patterns Discovery in Large Databases of Paintings. In *Digital Humanities 2016: Conference Abstracts*. Jagiellonian University & Pedagogical University, Kraków, pp. 169-172. <http://dh2016.adho.org/abstracts/348>
- Lincoln, M.** (2016). If Paintings were Plants: Measuring Genre Diversity in Seventeenth-Century Dutch Painting and Printmaking. In *Digital Humanities 2016: Conference Abstracts*. Jagiellonian University & Pedagogical University, Kraków, pp. 256-259. <http://dh2016.adho.org/abstracts/133>
- Manovich, L.** (2012). How to compare one million Images? In Berry, D. M., *Understanding Digital Humanities*, pp. 249-78.
- Manovich, L. and Tifentale, A.** (2015). Selfiecity: Exploring Photography and Self-Fashioning in *Social Media*. In Berry, David M., Dieter, M. (eds), *Postdigital Aesthetics: Art, Computation and Design*, pp. 109-22.
- McCloud, S.** (1994). *Understanding Comics: The Invisible Art.* New York: HarperPerenn
- Ordelman, R., Kemman, M., Kleppe, M., de Jong, F., Scagliola, S.** (2014) Sound and (moving) Images in Focus - How to integrate Audiovisual Material in Digital Humanities Research. *Digital Humanities 2014* <http://dhar-chive.org/paper/DH2014/Workshops-914.xml>
- Reser, G. and Bauman, J.** (2012). The Past, Present, and Future of Embedded Metadata for the Long-Term Maintenance of and Access to Digital Image Files. *International Journal of Digital Library Systems (IJDLIS)*, 3(1): 53-64.
- Reside, D.** (2014). Using Computer Vision to Improve Image Metadata. *Digital Humanities 2014*. <http://dhar-chive.org/paper/DH2014/Paper-294.xml>
- Salt, B.** (1974). The Statistical Style Analysis of Motion Pictures. *Film Quarterly*, 28(1): 13-22.
- Smeulders, A. W., Worring, M., Santini, S., Gupta, A., & Jain, R.** (2000). Content-based image retrieval at the end of the early years. *Pattern Analysis and Machine Intelligence, IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12), 1349-138
- Weingart, S.** (2017), *Submission to DH2017 (pt. 1)* <http://scottbot.net/submissions-to-dh2017-pt-1/>
- Terras, M. M. and Kirton, I.** (2013). Where do images of art go once they go online? A Reverse Image Lookup study to assess the dissemination of digitized cultural heritage. *Selected papers from Museums and the Web North America*, pp. 237-48. <http://mw2013.museumsandtheweb.com/paper/where-do-images-of-art-go-once-they-go-online-a-reverse-image-lookup-study-to-assess-the-dissemination-of-digitized-cultural-heritage/>