

### AUTHENTIZITÄT DIGITALER BILDER VOR UND NACH DER MIGRATION

Ein Semesterprojekt in Kooperation zwischen Bauhaus-Universität Weimar und Thüringischem Hauptstaatsarchiv Weimar

Im Rahmen des Aufbaus eines Digitalen Magazins für den Freistaat Thüringen zur Archivierung archivwürdiger, elektronischer Unterlagen der Behörden, Gerichte und sonstigen Einrichtungen des Landes sowie der nachgeordneten Bundesbehörden auf dem Territorium des Freistaats, berät der Lehrstuhl für Mediensicherheit der Bauhaus-Universität Weimar seit Ende 2012 das Thüringische Hauptstaatsarchiv Weimar in Fragen der IT-Sicherheit und kryptografischen Maßnahmenplanung und begleitet die Softwareentwicklung. So entstand im Wintersemester 2015/16 ein Semesterprojekt mit Masterstudenten des Studienfachs *Computer Science and Media*. Dabei befassten sich die Studierenden mit der Echtheit (Authentizität) digitaler Bilder, deren ursprüngliche Dateiformate in genormte, langfristig unterstützte Archivierungsformate konvertiert wurden. Diese so genannten Transformationen bzw. Migrationen von Dateiformaten werden in der elektronischen Archivierung eingesetzt, um die Inhaltsinformation und Aussagekraft eines elektronischen Dokumentes langfristig zu erhalten, auch wenn die ursprünglichen Dateiformate bereits nicht mehr lesbar sind.

#### Bestimmung signifikanter Eigenschaften

Zur Feststellung der Authentizität eines Bildes nach der Formatmigration werden sogenannte signifikante Eigenschaften zu Rate gezogen. Diese werden anhand des Originalbildes festgelegt und müssen nach der Formattransformation in dem konvertierten Bild unverändert nachweisbar sein. Hierfür können Metadaten, die das Bild näher beschreiben, wie bspw. Dateigröße, Pixel und verwendeter Farbraum, herangezogen werden. Allein auf Basis dieser Angaben kann jedoch nicht sichergestellt werden, dass im Rahmen einer Dateiformatänderung keine Bildinhalte verloren gehen bzw. verändert werden. Das Semesterprojekt befasste sich daher vorrangig mit der Identifizierung von signifikanten Bildeigenschaften, die sich direkt auf die Bildinhalte beziehen.

Signifikante Eigenschaften sollten vollständig und unvoreingenommen sein. Das Vorselektieren wichtiger Eigenschaften kann dazu führen, dass das Bild in Zukunft nicht mehr uneingeschränkt verwendet oder interpretiert werden kann, da die

Interessen der Betrachter eines Bildes zum Zeitpunkt der Archivierung nicht mit denen zum Zeitpunkt des Abrufs (zum Teil erst nach mehreren Jahrzehnten) übereinstimmen müssen. Wurden Bildinformationen während der Archivierung als nicht wichtig eingestuft und folglich nicht erhalten, sind sie eventuell nachträglich nicht mehr abrufbar und daher auch für zukünftige Benutzerinteressen verloren.

#### Authentifizierung basierend auf Bildinhalten: Getrennte Farb- und Kantenevaluierung

Das Projekt fokussierte sich auf die Klassifizierung von Manipulationen, die eingeschränkt akzeptiert werden können oder gar nicht zulässig sind. Eingeschränkt zugelassene Veränderungen, sogenannte inhaltserhaltende Manipulationen, sind zum Beispiel Farbkorrekturen oder die Dateiformatmigration selbst, während geometrische Veränderungen (z.B. Rotation, Spiegelung, etc.) und das Abändern des Bildinhaltes (z.B. Hinzufügen, Verschieben oder Löschen von Bildinhalten) strikt ausgeschlossen werden müssen, da sie die Aussage des Bildes verändern können. Deshalb wurde ein Ansatz gewählt, der Kanten- und Farbvalidierung getrennt voneinander betrachtet (Abbildung 1): Ein Großteil der unerwünschten Manipulationen lässt sich erkennen, wenn man die Unversehrtheit der Kanten in der Original- und Archivierungsdatei nachweisen kann. Farben hingegen erlauben die Evaluierung von inhaltserhaltenden Manipulationen.



Abbildung 1: Darstellung der Trennung eines Bildes (links) in Kanten (Mitte) und die drei RGB-Farbkanaäle Rot, Grün und Blau (rechts).

Quelle Originalbild: StockSnap.io. Beautiful Free Stock Photos (CC0). URL: <https://stocksnap.io/photo/IHPVE3XLUE> (letzter Zugriff am 16.06.2016).

Auf Basis dieser Annahmen experimentierte die Projektgruppe mit unterschiedlichen Algorithmen zur Extrahierung von Eigenschaften, um die Entwicklung eines eigenen Authentifizierungstools anzustoßen. Dieses Tool soll zukünftig automatisch Eigenschaften aus Bildern extrahieren und daraus Signaturen erstellen. Diese Signaturen sind ähnlich wie Fingerabdrücke für jedes Bild eindeutig. Wenn die Signatur einer Originaldatei mit der Signatur der zugehörigen migrierten Datei übereinstimmt, ist das Bild im Regelfall authentisch und der Inhalt somit nicht verändert worden.

Zur Kantvalidierung wurden mehrere Ansätze getestet, unter anderem auch der Edge Histogram Descriptor (Kanten-histogramm Deskriptor) des MPEG-7-Standards, welcher zur vereinheitlichten Beschreibung unterschiedlicher Eigenschaften von Film- und Bilddaten herangezogen wird. Aufgrund der geringen Größe der MPEG-7-Deskriptoren werden diese häufig eingesetzt, wenn eine große Menge an Daten in möglichst kurzer Zeit nach bestimmten Eigenschaften durchsucht werden muss. Der Edge Histogram Descriptor beschreibt dabei die Anzahl bestimmter Kantentypen, die in einzelnen Bereichen eines Bildes auftreten. Unterschieden werden hierbei fünf Kantentypen, deren Anzahl dann pro untersuchtem Bildbereich angegeben und in Form eines Histogramms (Abbildung 3) ausgedrückt wird. Zusätzlich wurde mit der Fourier-Transformation experimentiert, die ein Bild als Verbund aus wellenförmigen Signalen interpretiert und in Form von Phasen und Amplituden darstellt (Abbildung 2).

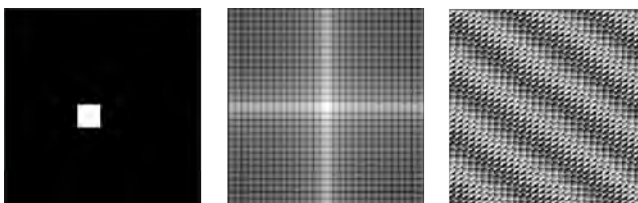


Abbildung 2: Interpretation des Ausgangsbildes (links) als Signale: Amplitude (mittig) und Phase (rechts).

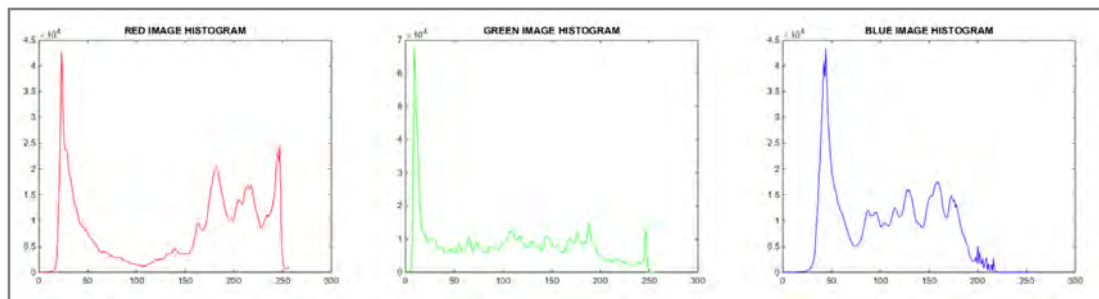


Abbildung 3: Repräsentation eines Bildes als Histogramm der drei Farbkanäle Rot, Grün und Blau im RGB-Farbraum.

Die Phase steht hierbei für die Höhe der Welle, die Amplitude gibt die Position selbiger an. Im Fall der Fourier-Transformation wird jeder Wechsel der Helligkeit (= Intensitätswechsel) in einem Bild als ein solches Signal interpretiert. Wechsel von Schwarz nach Weiß verursachen hierbei größere Amplituden als Intensitätsänderungen in den Graubereichen.

Für die Farbevaluierung wurden klassische Ansätze wie globale und lokale Farbhistogramme verwendet (Abbildung 3). Hierbei wird die Anzahl farbiger Pixel jeweils der entsprechenden Intensität zugeordnet. Im 8-Bit-RGB-Farbraum ergibt sich daraus jeweils ein Histogramm für jeden der drei Farbkanäle Rot, Grün und Blau mit einer Intensitätsspanne von 0 (minimale Intensität) bis 255 (maximale Intensität). Zusätzlich wurden zwei Farbdeskriptoren des MPEG-7-Standards evaluiert: Der Dominant Color Descriptor beschreibt die acht repräsentativsten Farben eines Bildes, während der Color

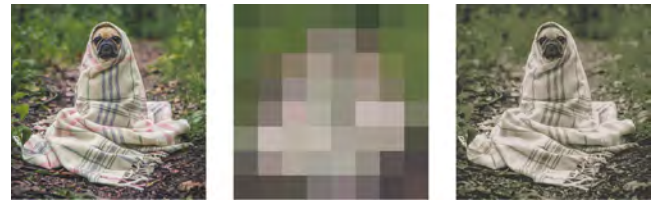


Abbildung 4: Simulation des Color Layouts (Mitte) und des Dominant Color Descriptors (rechts).

Quelle Originalbild: StockSnap.io. Beautiful Free Stock Photos (CC0). URL: <https://stocksnap.io/photo/IHPVE3XLUE> (letzter Zugriff am 16.06.2016).

Layout Descriptor ein Bild in 64 farbige Kacheln unterteilt, die sich aus den Farb-Mittelwerten der umliegenden Pixel zusammensetzen (Abbildung 4).

In einem Testsetup mit über 100 manipulierten Testbildern konnte bereits bestätigt werden, dass sich die oben beschriebenen Ansätze zur Erkennung von Veränderungen von Bildern eignen. Insbesondere die Vereinigung verschiedener Ansätze versprach ein zufriedenstellendes Resultat.

Erste Ergebnisse des Semesterprojektes konnten auf der 20. Tagung des Arbeitskreises ›Archivierung von Unterlagen aus digitalen Systemen‹ Anfang März 2016 in Potsdam vorgestellt werden und fanden großen Zuspruch. Das Projekt wird voraussichtlich im Wintersemester 2016/17 fortgeführt.

Veronika Krauß  
Arefeh Bahrami  
Prof. Dr. Stefan Lucks

*Frau Krauß ist Absolventin des Bachelorstudiengangs ›Medieninformatik‹ der Hochschule für angewandte Wissenschaften Hof. Zurzeit studiert sie im 4. Fachsemester den Masterstudiengang ›Computer Science and Media‹ an der Bauhaus-Universität Weimar. Sie wird ihr Studium demnächst abschließen.*

*Frau Bahrami ist Absolventin des Bachelorstudiengangs ›Computer Software Engineering‹ der Najafabad University in Esfahan (Iran). Zurzeit studiert sie im 3. Fachsemester den Masterstudiengang ›Computer Science and Media‹ an der Bauhaus-Universität Weimar.*

*Herr Prof. Dr. Lucks ist seit 2007 Inhaber des Lehrstuhls Medien-sicherheit der Bauhaus-Universität Weimar. In Forschung und Lehre beschäftigt er sich insbesondere mit kryptografischen Algorithmen sowie Sicherheitsprotokollen für sichere Medien und Kommunikationssysteme.*

## TIPPS UND TERMINE

### 20. TAGUNG DES ARBEITSKREISES ›ARCHIVIERUNG VON UNTERLAGEN AUS DIGITALEN SYSTEMEN‹ IN POTSDAM

Das Projektteam hat auch in diesem Jahr an der Tagung des Arbeitskreises teilgenommen, die am 2. und 3. März in Potsdam ihr 20jähriges Jubiläum feierte. Nachdem das Thüringische Hauptstaatsarchiv Weimar 2014 selbst Gastgeber der Tagung sein durfte und das Projekt ›Digitales Magazin des Freistaats Thüringen‹ seit 2013 mit Vorträgen zu verschiedenen Themen im Arbeitskreis vertreten war, wurde in diesem Jahr der Fortschritt des Projekts in kompakter Weise präsentiert. Im Rahmen einer Posterausstellung konnten sich Interessierte über den Fortgang von Projekten informieren, die bereits durch Vorträge früherer Tagungen bekannt waren. Herr Filtz stand am Plakat des Thüringischen elektronischen Magazins (ThELMA) in den Kaffeepausen Rede und Antwort. Der Lehrstuhl für Mediensicherheit der Bauhaus-Universität Weimar war mit einem Vortrag der Studentin Veronika Krauß vertreten, die über ein Seminarprojekt des Wintersemesters berichtete. Der spannende Vortrag drehte sich um die technischen Herausforderungen, denen sich die digitale Archivierung in Bezug auf Bilddateien stellen muss. Für genauere Informationen zu diesem Projekt, seinen Ergebnissen und der Kooperation des Lehrstuhls mit dem ThHStAW werfen Sie gerne einen Blick in den Leitartikel dieser Newsletter-Ausgabe.

Das Staatsarchiv St. Gallen informiert auf seiner Tagungsseite über das Programm. Sie finden dort auch die Vortragspräsentationen:

[www.staatsarchiv.sg.ch/home/auds/20.html](http://www.staatsarchiv.sg.ch/home/auds/20.html)

Als diesjähriger Gastgeber bietet die Fachhochschule Potsdam auf ihrer Webseite mit einem Bericht über den Ablauf der Tagung und Videos, die die Vorträge in direkten Zusammenhang mit den Präsentationen stellen, für Interessierte einen besonderen Mehrwert:

[www.fh-potsdam.de/studieren/informationwissenschaften/forschung-und-entwicklung/tagungen/auds-tagung-2016/tagungsbericht/vortraege/](http://www.fh-potsdam.de/studieren/informationwissenschaften/forschung-und-entwicklung/tagungen/auds-tagung-2016/tagungsbericht/vortraege/)

### VOM ZAUBER DES WANDELS – NESTOR-PRAKTIKERTAG 2016 IN DRESDEN

Am 14. Juni fand an der Staats-, Landes- und Universitätsbibliothek Dresden der diesjährige nestor-Praktikertag zum Thema Migration und Emulation statt. Vertreter verschiedener

Institutionen aus dem deutschsprachigen In- und Ausland kamen zusammen, um aktuelle Projekte vorzustellen, Erfahrungen auszutauschen und ins Gespräch zu kommen. Besondere Freude machte allen Teilnehmerinnen und Teilnehmern der Einführungsvortrag ›Migration und Emulation – Angewandte Magie?‹ von Stefan Funk von der Staats- und Universitätsbibliothek Göttingen, wobei der Begriff ›Vortrag‹ der tatsächlich magisch anmutenden Darbietung nicht gerecht wird: Herr Funk war mit mehreren Kisten veralteter Hardware angereist, mit der er in der Manier eines Alleinunterhalters verschiedene Aspekte von Migration und Emulation live vorführte. Schlussendlich konnte er die Daten, die auf dem Magnetband einer Audiokassette gespeichert waren, auf seinem modernen MacBook in der emulierten Systemumgebung eines Commodore 64 wiedergeben. Sowohl für Einsteiger in die digitale Archivierung wie auch für Fortgeschrittene war dies ein eindrucksvoller Beginn der Tagung, die mit interessanten Einblicken in die Praxis und regen Diskussionen erfolgreich verlief.

Über die weiteren Themen der Vorträge und die Referentinnen und Referenten können Sie sich auf der Webseite von nestor informieren:

[www.langzeitarchivierung.de/Subsites/nestor/DE/Veranstaltungen/TermineNestor/praktikertag2016.html](http://www.langzeitarchivierung.de/Subsites/nestor/DE/Veranstaltungen/TermineNestor/praktikertag2016.html)

Dort sind auch die Präsentationen veröffentlicht.

### FORTBILDUNG ZUR DIGITALEN ARCHIVIERUNG IM HERBST

Am 12. und 13. September findet im Rahmen des Jahresprogramms des Thüringer Ministeriums für Inneres und Kommunales erneut eine ressortübergreifende Fortbildung zur digitalen Archivierung statt, die das Thema von Seiten der abgebenden Behörden theoretisch und praktisch beleuchtet. Es werden Kenntnisse zu bestehenden Standards, Fachbegriffen und Arten elektronischer Daten sowie die Herausforderungen im Umgang mit diesen vermittelt. Sie lernen, die elektronischen Daten von ihrer Entstehung, über die Langzeitspeicherung, bis hin zur Anbiertung an die Staatsarchive effektiv zu managen. Ein praktischer Teil mit Übungen im Thüringischen elektronischen Magazin (ThELMA) rundet die Veranstaltung ab.

Bei Interesse an dieser Fortbildung erhalten Sie als Mitarbeiterin bzw. Mitarbeiter des Landes weitere Informationen über Inhalte und Ablauf bei der Redaktion dieses Newsletters. Wir freuen uns über Ihre Anfrage unter:

[digitales.magazin@staatsarchive.thueringen.de](mailto:digitales.magazin@staatsarchive.thueringen.de)

## DIGITALE VERWALTUNG THÜRINGEN

### ARBEITSTREFFEN DER VIS-ANWENDER IN BREMEN

Die Behörden, Gerichte und sonstigen Einrichtungen des Freistaats Thüringen arbeiten seit vielen Jahren mit Dokumentenmanagement- und Vorgangsbearbeitungssystemen (DMS,

VBS). Der Einsatz und Betrieb soll in Vorbereitung der Einführung der Elektronischen Akte in der Thüringer Verwaltung in den nächsten Jahren sukzessive normiert und zentralisiert werden. Auch wenn die Aktenführung im Land im Moment noch nicht verpflichtend elektronisch erfolgt, entstehen durch die Verwendung von DMS und VBS bereits massenhaft Elektronische Akten, die kein bzw. kein vollständiges Papier-

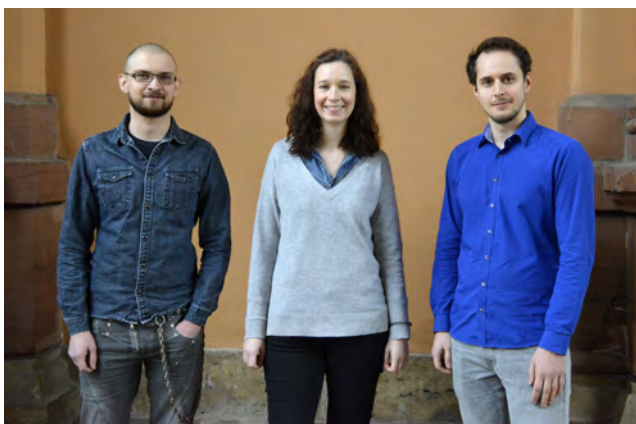
äquivalent mehr besitzen und die den Thüringischen Staatsarchiven nach Ablauf der gesetzlichen Aufbewahrungsfristen anzubieten sind. Die Staatsarchive sind daher bestrebt, die Festlegung von Datenstrukturen und Datenaustauschformaten sowie die technische Entwicklung der Anbietungs- und Aussonderungsprozesse in den behördlichen Fachanwendungen mit ihrem Know-how zu begleiten. Für das von vielen Behörden des Freistaats eingesetzte DMS VIS der Firma PDV-Systeme besteht seit vielen Jahren ein etablierter Kreis bundesweiter Anwender, der sich regelmäßig in Arbeitstreffen zu organisatorischen, funktionalen, ergonomischen, administrativen und sonstigen technischen Fragen des Systems austauscht, berät und gemeinsame Optimierungs- und Fortentwicklungswünsche gegenüber dem Hersteller vertritt. Das Projektteam zum Aufbau des Digitalen Magazins des Freistaats Thüringen ist seit verganginem Jahr mit einem Mitglied in der VIS-Anwendergruppe sowie in der Unterarbeitsgruppe *Langzeitspeicherung und Aussonderung* aktiv vertreten. Das

diesjährige Anwendertreffen wurde am 15. und 16. Juni vom Land Bremen ausgerichtet. Ein Schwerpunkt der Tagung lag im Austausch über den derzeit in VIS implementierten Workflow zur Anbietung und Aussonderung elektronischer Unterlagen, dessen Funktionen und Arbeitsschritte von Maik Hofmann (PDV-Systeme) in einer Präsentation demonstriert wurden. Als Überarbeitungs- und Optimierungsbedarfe wurden u.a. die Benutzungsvereinfachung durch Teilautomation und Zusammenfassen von Arbeitsschritten, die Behebung von Problemen bei der Dateiformatkonvertierung nach PDF/A sowie die sicherheitstechnische Verbesserung des Datenaustauschs zwischen Fach- und Archivbehörde im Rahmen des zwei- und vierstufigen Aussonderungsverfahrens identifiziert. Im Ergebnis konnte Einigung im Anwenderkreis erzielt werden, dass die Anforderungen bundesweit konsolidiert und ein Standard-Aussonderungsmodul für VIS angestrebt werden sollte.

## IN EIGENER SACHE

### NEUIGKEITEN IM PROJEKTTEAM

Zum November 2015 hat das Projektteam eine Frischzellenkur in Form dreier neuer MitarbeiterInnen erfahren:



Die neuen Teammitglieder (v.l.n.r.):  
Daniel Wittmann, Svenia Pohlkamp, Jonathan Siefert.

Nach dem Weggang von Herrn Christian Horn konnte mit Frau Svenia Pohlkamp eine neue Projektmitarbeiterin gewonnen und das dreiköpfige Projektteam wieder vollständig besetzt werden. Frau Pohlkamp ist Buchwissenschaftlerin sowie Bibliotheks- und Informationswissenschaftlerin mit Abschlüssen von der Johannes Gutenberg-Universität Mainz und der Humboldt-Universität zu Berlin. Aus ihrem Studium bringt sie Vorkenntnisse auf dem interdisziplinären Gebiet der elektronischen Archivierung sowie erste Erfahrungen im Umgang mit

heterogenen Sammlungen und ihren Metadaten mit. Nach der erfolgreichen Einarbeitung liegen Schwerpunkte ihrer Tätigkeit in den Bereichen Metadaten, Datenschutz sowie Öffentlichkeitsarbeit.

Im Rahmen der Kooperation mit dem Lehrstuhl für Mediensicherheit der Bauhaus-Universität Weimar hat sich zudem die Möglichkeit ergeben, zwei studentische Hilfskräfte insbesondere für die umfangreichen Systemtests im Digitalen Magazin einzusetzen. Herr Jonathan Siefert und Herr Daniel Wittmann studieren in Weimar Medieninformatik mit dem Schwerpunkt Mediensicherheit. Als angehende Fachleute für Softwareentwicklung und moderne Webtechnologien sowie deren Schutz vor Angriffen haben sie sich schnell mit den Grundlagen der elektronischen Archivierung vertraut gemacht und unterstützen das Projekt in vielen technischen Belangen während des Aufbaus der IT-Infrastruktur für das Digitale Magazin.

### NÄCHSTE NEWSLETTER-AUSGABE

In der nächsten Ausgabe möchten wir Sie über die derzeit laufende Zusammenarbeit zwischen dem Hauptstaatsarchiv und dem Thüringer Landesrechenzentrum zur Integration des Digitalen Magazins in die Systemumgebung des Rechenzentrums informieren. In diesem Projekt werden die technischen und organisatorischen Rahmenbedingungen für den Dauerbetrieb des Digitalen Magazins geschaffen.

Alle Newsletter-Ausgaben unter:  
[www.thueringen.de/th1/tsk/kultur/staatsarchive/fachinformationen/digital/projekt/newsletter/](http://www.thueringen.de/th1/tsk/kultur/staatsarchive/fachinformationen/digital/projekt/newsletter/)