Contributions in Mathematical and Computational Sciences 3

Hans Georg Bock
Willi Jäger
Michael J. Winckler *Editors*

Scientific Computing and Cultural Heritage

Contributions in Computational Humanities

Springer

Contributions in Mathematical and Computational Sciences • Volume 3

Editors

Hans Georg Bock
Willi Jäger
Otmar Venjakob

For other titles published in this series, go to
http://www.springer.com/series/8861

Hans Georg Bock • Willi Jäger •
Michael J. Winckler
Editors

Scientific Computing and Cultural Heritage

Contributions in Computational Humanities

Springer

Editors
Hans Georg Bock
Willi Jäger
Michael J. Winckler
University of Heidelberg
Heidelberg
Germany

ISSN 2191-303X ISSN 2191-3048 (electronic)
ISBN 978-3-642-28020-7 ISBN 978-3-642-28021-4 (eBook)
DOI 10.1007/978-3-642-28021-4
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012951147

Mathematics Subject Classification (2010): 65-06, 68-06

© Springer-Verlag Berlin Heidelberg 2013
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Scientific Computing and Cultural Heritage: Contributions in Computational Humanities

Modern information technologies are going to change substantially research in humanities and cultural sciences as well as the presentation of its results. During recent decades mathematical and computational methods became important tools also in these sciences. Computers are no longer used just as elaborate typewriters and machines for traditional text processing, but offer an enormous potential in creating, collecting, storing, retrieving, processing and disseminating of scholarly information. Whereas in disciplines like linguistics, computer and software systems are well established as tools, many disciplines are still just attempting to exploit the possibilities and opportunities scientific computing is offering. Here scientific computing is conceived as a field integrating all sub-disciplines of mathematics and information sciences needed for problem solving with the help of computers.

Computational humanities are an emerging discipline, following concepts like the computational sciences in other fields. Humanities represent a broad spectrum of disciplines ranging from historical to cultural and social sciences. The term *computational* is chosen instead of *digital,* used in the name *digital humanities*, since the spectrum of concepts and methods applied is broader and not focused mainly on information sciences.

New tools adjusted to the special needs in the different disciplines have to be developed. Here is a shortlist of topics, where mathematical and computational methods are in demand and can contribute to the progress of research in the humanities as well as to preserving cultural heritage:

1. Collecting, storing, retrieving and processing complex data, which may be digitized.
2. Processing of information in text and images.
3. Information retrieval from hybrid data sets.
4. Reconstruction or construction of objects, buildings and historical monuments.
5. Preservation and restoration of cultural heritage.
6. Establishing, presenting and providing information systems (virtual archives, libraries and museums) allowing in particular a user-friendly internet access.

7. Modeling, simulations and tests of possible scenarios of courses and processes, important for the evolution of civilization.
8. Developing computer aided tools for a workplace in humanities.
9. Initiating and maintaining virtual co-laboratories.

New challenges are demanding new concepts, methods and computational techniques, e.g.

- Improved methods of image processing and filtering, colour based image processing,
- Content based searching and image retrieval methods,
- Character and word searching in imaged documents,
- Algorithms for restoration of documents and monuments,
- Development of computational methods to analyse and characterize huge data sets and to optimize processes involving large networks,
- Modelling and simulation of physical, chemical and biological processes, relevant in the evolution of civilization, in preservation and restoration,
- Representation and visualisation of multidimensional information.

Preserving the cultural heritage includes also technological problems like the physical preservation of the monuments suffering e.g. from weathering or incompetent restorations. Research into methods of how to protect buildings and monuments is urgently needed. An efficient solution is also highly important for the economic and financial aspects of preservation.

As usual, success stories and prime examples are the best promotion also for *computational humanities*. This volume contains a collection of articles on topics presented at a workshop of the Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS) in November 2009. Humanities and cultural sciences are the topic of a pioneering project of the Interdisciplinary Center for Scientific Computing. IWR is also one of the partners in the project *Transcultural Studies and Scientific Computing in the Humanities* at Heidelberg University, initiated and funded within the Excellence Initiative in Germany. Since its start 25 years ago, IWR has the graphic representation and analysis of complex data as a main aim in its program and has contributed successfully to visualization of historic architecture and monuments.

The papers in this volume are grouped in 3 sections:

Section 1 deals with topics whose focus is more on the side of mathematical and computational methods, e.g. in generation and processing of images. Their performance is illustrated in special applications mainly in archeology and history.

Section 2 deals with research in information systems, contributing to topic (6) in the shortlist.

Section 3 gives an insight into the ongoing work on documenting, restoring and presenting cultural monuments ranging from the temples in Pompeii to the Banteay Chhmar temples of the Angkorian period.

In order to illustrate the challenges to mathematics and information sciences, we just pick out the contribution on computer-based understanding of medieval images. In this case the central problem is to develop algorithms for an object-oriented search in images, a problem highly important by itself and arising in many situations. Decisive steps are a proper characterization of the objects of interest and an adapted search algorithm. In general, the storing and searching of images in a data base needs to be improved.

The editors and the authors of this volume hope to attract the interest especially of young researchers to this young discipline *computational humanities.* Bridging disciplines demands disciplinary strength and, depending on the problem, a good understanding of the partners. Research in *computational humanities* is a challenge, offering many perspectives.

Heidelberg
November 2011

Willi Jäger
Michael J. Winckler

Preface to the Series

Contributions to Mathematical and Computational Sciences

Mathematical theories and methods and effective computational algorithms are crucial in coping with the challenges arising in the sciences and in many areas of their application. New concepts and approaches are necessary in order to overcome the complexity barriers particularly created by nonlinearity, high-dimensionality, multiple scales and uncertainty. Combining advanced mathematical and computational methods and computer technology is an essential key to achieving progress, often even in purely theoretical research.

The term mathematical sciences refers to mathematics and its genuine sub-fields, as well as to scientific disciplines that are based on mathematical concepts and methods, including sub-fields of the natural and life sciences, the engineering and social sciences and recently also of the humanities. It is a major aim of this series to integrate the different sub-fields within mathematics and the computational sciences, and to build bridges to all academic disciplines, to industry and other fields of society, where mathematical and computational methods are necessary tools for progress. Fundamental and application-oriented research will be covered in proper balance.

The series will further offer contributions on areas at the frontier of research, providing both detailed information on topical research, as well as surveys of the state-of-the-art in a manner not usually possible in standard journal publications. Its volumes are intended to cover themes involving more than just a single "spectral line" of the rich spectrum of mathematical and computational research.

The Mathematics Center Heidelberg (MATCH) and the Interdisciplinary Center for Scientific Computing (IWR) with its Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS) are in charge of providing and preparing the material for publication. A substantial part of the material will be acquired in workshops and symposia organized by these institutions in topical areas of research. The resulting volumes should be more than just proceedings collecting papers submitted in advance. The exchange of information and the discussions during the meetings should also have a substantial influence on the contributions.

This series is a venture posing challenges to all partners involved. A unique style attracting a larger audience beyond the group of experts in the subject areas of specific volumes will have to be developed.

Springer Verlag deserves our special appreciation for its most efficient support in structuring and initiating this series.

Heidelberg University
Germany

Hans Georg Bock
Willi Jäger
Otmar Venjakob

Contents

Part I Development of Scientific Computing Methods

Part II Information-Based Research

Part III Case Studies and Applications

Part I

Development of Scientific Computing Methods

Chapter 1
Mathematical Methods for Spectral Image Reconstruction

Wolfgang Baatz, Massimo Fornasier, and Jan Haskovec

Abstract We present a method for recovery of damaged parts of old paintings (frescoes), caused by degradation of the pigments contained in the paint layer. The original visible colour information in the damaged parts can be faithfully recovered from measurements of absorption spectra in the invisible region (IR and UV) and from the full spectral data of the well preserved parts of the image. We test algorithms recently designed for low-rank matrix recovery from few observations of their entries. In particular, we address the singular value thresholding (SVT) algorithm by Cai, Candès and Shen, and the iteratively re-weighted least squares minimization (IRLS) by Daubechies, DeVore, Fornasier and Güntürk, suitably adapted to work for low-rank matrices. In addition to these two algorithms, which are iterative in nature, we propose a third non-iterative method (which we call block completion, BC), which can be applied in the situation when the missing elements of a low-rank matrix constitute a block (submatrix); this is always true in our application. We shortly introduce the SVT and IRLS algorithms and present a simple analysis of the BC method. We eventually demonstrate the performance of these three methods on a sample fresco.

Keywords Image recovery • UV and IR absorption spectra • Matrix completion

---

W. Baatz (✉)
Akademie der Bildenden Künste, Wien, Austria
e-mail: w.Baatz@akbild.ac.at

M. Fornasier • J. Haskovec
Johann Radon Institute for Computational and Applied Mathematics, Linz, Austria
e-mail: massimo.fornasier@oeaw.ac.at; jan.haskovec@oeaw.ac.at

H.G. Bock et al. (eds.), *Scientific Computing and Cultural Heritage*,
Contributions in Mathematical and Computational Sciences 3,
DOI 10.1007/978-3-642-28021-4_1, © Springer-Verlag Berlin Heidelberg 2013

1.1 Introduction

In old frescoes, the visible colour information might be completely or partially lost in some parts of the painting. This is due to mechanical or (photo) chemical degradation or other changes of the paint layer. In particular, pigments, but also dyes might undergo modifications, which change their light absorption properties [8]. The influence of ancient binding medium augmenting or diminishing the fluorescence intensities in certain regions of the spectrum has to be taken into consideration as well. However, if these effects do not largely influence the absorption of pigments in invisible parts of the spectrum (UV and IR), there is a hope that the original colour information can be faithfully recovered, using the information from the well conserved parts of the painting.

In this paper, we show that the problem of image recovery can be mathematically formulated as a "matrix completion problem" [2, 9] and consider three algorithms for its solution: The singular value thresholding (SVT) algorithm [1], the iteratively re-weighted least squares minimization (IRLS) [5] and, finally, the block completion (BC) method, which is the original contribution of this paper in the context of completion of low-rank matrices where the missing elements form a sub matrix (block). We present the results of these three methods for a sample fresco with parts which has been fictitiously removed, because assumed corrupted.

1.2 Main Idea of Spectral Recovery

The fresco under consideration is divided into N possibly small surface portions, which we call pixels in this context. In each pixel, we measure the absorption spectrum in the interval of wavelengths 307–1,134 nm with the resolution of M equidistant samples (Fig. 1.1). A typical human eye will respond to wavelengths from about 380 to 750 nm, wavelengths below 380 nm and, resp., above 750 nm correspond to UV, resp. IR. The measurements are collected in a matrix $A \in \mathrm{R}^{N \times M}$, where each row corresponds to one pixel and each column to one measured wavelength. We denote by Λ_{vis} the subset of indices of $\{1,\ldots,M\}$ corresponding to the measurements in the region of visible wavelengths. By $\Theta \subset \{1,\ldots,N\}$ we denote the set of non-damaged pixels, i.e., those where the colour information is considered authentic (in the language of machine learning, Θ is the *training set*); Θ^c is its complement in $\{1,\ldots,N\}$. Consequently, the indices in $\Omega^c = \Theta^c \times \Lambda_{\mathrm{vis}}$ correspond to the damaged visible parts of the fresco and are to be reconstructed. In fact, we completely neglect the measurements taken in Ω^c and the reconstruction will be based purely on the information contained in the complement Ω.

The main idea of our spectral recovery method is to look for a *low-rank* matrix Z which coincides with A on the index set Ω. This is a reasonable thing to do if the following assumptions are met:

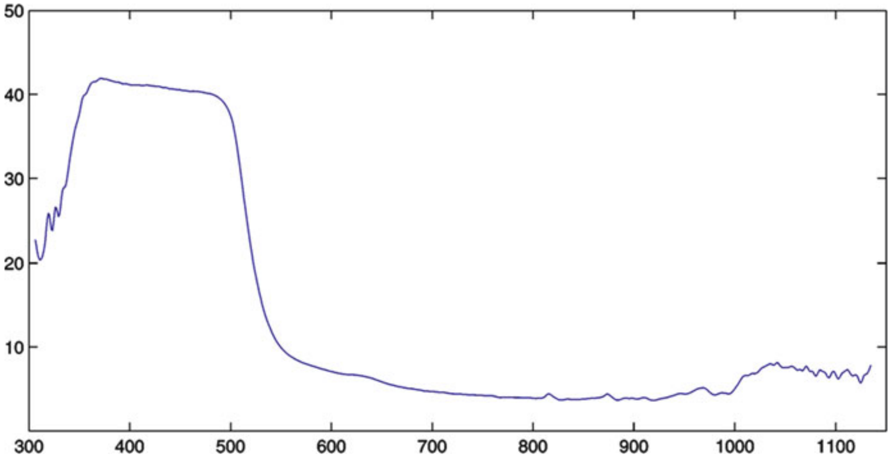

Fig. 1.1 A sample spectrum: absorption [%] vs. wavelength [nm]

1. The invisible parts (UV, IR) of the absorption spectra in the damaged regions are not largely influenced.
2. Mixing of pigments affects the absorption in a linear manner (i.e., the spectrum corresponding to a mixture of pigments is a linear combination of the spectra produced by the single components).
3. Pigments of different colours have different invisible absorption spectra.

While assumptions 2 and 3 seem to be plausible, the first one is more questionable. Apparently, its validity will be largely influenced by the kind and degree of damage the painting is subject to, and, therefore, it will depend on the particular case at hand. In the examples presented in Sect. 1.5, sample frescoes are used which were not subject to any real damage, and the region Ω^c of missing information is introduced fictitiously. Consequently, the data in the invisible regions are untouched and clearly the first assumption holds.

From the mathematical point of view, we are about to solve the following *low-rank matrix completion problem*

$$\begin{aligned} &\text{minimize} && \operatorname{rank}(X) \\ &\text{subject to} && X_{ij} = A_{ij} \quad \text{for } (\mathrm{I}, \mathrm{j}) \in \Omega. \end{aligned} \tag{1.1}$$

However, this is known to be NP-hard and all known algorithms which provide exact solutions require time doubly exponential in the dimension of the matrix in both theory and practice [4]. As a remedy, it was proposed to consider the convex relaxation problem

$$\begin{aligned} &\text{minimize} && \|X\|_* \\ &\text{subject to} && X_{ij} = A_{ij} \quad \text{for } (\mathrm{I}, \mathrm{j}) \in \Omega, \end{aligned} \tag{1.2}$$

where $\|\cdot\|_*$ denotes the *nuclear norm* defined by

$$\|X\|_* \sum_{k=1}^{\dim(X)} \sigma_{\mathrm{k}}(X) \tag{1.3}$$

and σ_{k} (X) is the k-th largest singular value of X. The main point here is that the nuclear norm is a convex function and in principle (1.2) can be solved via semidefinite programming. The corresponding theory has been developed recently, [2, 3], and the main result can be roughly summarized as follows:

Let A be an $N \times M$ matrix of rank r sampled randomly from a certain distribution, and put $n = \max(N,M)$. Suppose we observe m entries of A with locations sampled uniformly at random. Then, if m is of the order $nr \log(n)$, with very high probability the minimizer of problem (1.2) *is unique, and it coincides with A.*

1.3 Algorithms

Several methods have been proposed for the numerical solution of the semidefinite program (1.2). We will consider two of them:

1. The singular value thresholding (SVT) algorithm by Cai et al. [1]. The algorithm is iterative and produces a sequence of matrices $\{X_k, Y_k\}$; at each step, it mainly performs a soft-thresholding operation on the singular values of the matrix $\{Y_k\}$. There are two remarkable features making this attractive for low-k rank matrix completion problems. The first is that the soft-thresholding operation is applied to a sparse matrix; the second is that the rank of the iterates $\{X_k\}$ is empirically non-decreasing. Both these facts allow the algorithm to make use of very minimal storage space and keep the computational cost of each iteration low. On the theoretical side, a convergence analysis is provided in [1] showing that the sequence of iterates converges to the (unique) minimizer of (1.2).
2. The iteratively re-weighted least squares minimization (IRLS) by Daubechies, DeVore, Fornasier and Güntürk [5], is an iterative algorithm, which in each step for computing X_{k+1} minimizes a weighted Forbenius norm and updates the matrix of weights W_{k+1} appropriately. The method has been analyzed in [5] in the k +1 context of compressed sensing; an analysis in the context of the low-rank matrix completion problem is a work in progress [7].

Finally, we propose a new third method, which we call *block completion (BC)*; although it is relatively simple, it gives very competitive results, as shown in the next section. It is based on the assumption that the matrix A can be put in a form consisting of four blocks, where the unknown elements a_{ij} with $(i,j) \in \Omega^c$ are collected in the lower right block, denoted by $\tilde{\mathrm{B}}$. In our application the matrix A

indeed fulfills this assumption after a suitable permutation of rows and shift of columns:

$$A = \left(\begin{array}{c|c} \bar{A} & \tilde{A} \\ \hline \bar{B} & \tilde{B} \end{array}\right). \tag{1.4}$$

Given that the solution Z of the rank minimization problem (1.1) is unique, we make the fundamental assumption

$$\operatorname{rank}(Z) = \operatorname{rank}(\bar{A}). \tag{1.5}$$

In our work [6] we show that this assumption is fulfilled with very high probability for random matrices A as soon as the dimension of $\bar{A}$ is bigger or equal to rank(A). Then, we can reconstruct the missing information in $\tilde{B}$

- either by expressing the column s of $\tilde{A}$ as linear combinations of the columns of $\bar{A}$, i.e.,

  $$\tilde{A} = \bar{A}C \qquad \text{for a suitable matrix } C \tag{1.6}$$

  and establishing the same link between $\tilde{B}$ and $\bar{B}$, i.e., $\tilde{B} = \bar{B}C$;
- or by expressing the rows of $\bar{B}$ as a linear combination of the rows of $\bar{A}$, i.e.,

$$\bar{B} = D\bar{A} \qquad \text{for a suitable matrix } D, \tag{1.7}$$

and establishing the same link between $\tilde{B}$ and $\tilde{A}$, i.e., $\tilde{B} = D\bar{A}$.

In our context, the first option consists in "explaining" the visible spectra in the well preserved parts of the fresco by the corresponding IR and UV data, and recovering the missing visible information by "induction" from the invisible data. The second option establishes a link between the invisible spectra in the preserved and destroyed parts of the fresco, and reconstructs the painting by establishing the same link between the visible parts.

The following simple lemma shows that the results obtained by both options coincide:

Lemma 1. *Let A be of the block shape* (1.4) *and let rank(A)* $=$ *rank* ($\bar{A}$)*. Then for any matrix C satisfying* (1.6) *and any matrix D satisfying* (1.7)*, it holds*

$$\bar{B}C = D\bar{A}. \tag{1.8}$$

In particular, the result $\tilde{B}$ *does not depend on the choice of C or D.*

Proof. We have $\bar{B}C = D\bar{A}C = D\tilde{A}$.

Finally, let us remark that the block completion method is robust if the measured data are subject to noise. In this case, (1.6) and (1.7) might not be exactly solvable,

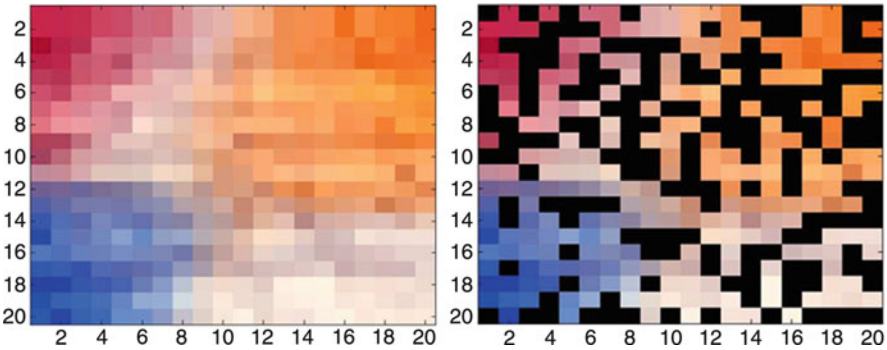

Fig. 1.2 The sample fresco, as obtained from the spectral measurements (*left*) and random deletion of the visible information of 50% of the pixels (*right*)

and the solution in the sense of least square minimization is taken instead. For details we refer to the manuscript in preparation [6].

1.4 Numerical Results

To test the performance of the aforementioned methods in "laboratory conditions", a sample fresco has been painted, using three pigments only – yellow, red, and blue. Measurements of absorption spectra have been taken on a grid of 20 × 20 pixels in the range 307–1,134 nm with the resolution of $M = 830$ equidistant samples per spectrum; i.e., the matrix $A \in \mathrm{R}^{N \times M}$ has dimensions $N = 400$ and $M = 830$. The measured data are transformed into RGB values by standard methods [11] for the sake of visualization, see Fig. 1.2 *left*.

Then, 200 points are chosen randomly where the visible information (380–750 nm) is deleted, Fig. 1.2 *right*. The so-created matrix A with the set of missing elements Ω^c is used as the input of the semidefinite program (1.1) and the SVT, IRLS, and BC algorithms are applied for its solution. Since the first two are of iterative nature, a suitable stopping criterion has to be imposed. We define the residual error at k-the iteration as the measure of the quality of recovery.

$$e^{\,k}_{\,res} \frac{\|(A - Xk)\Omega c\|F}{\|A|\Omega c\|F}, \tag{1.9}$$

where $A|_{\Omega}{}^{\mathrm{c}}$ denotes the matrix created by setting all elements from Ω to zero, and $\|\cdot\|_{\mathrm{F}}$ is the Frobenius matrix norm. X^k denotes the matrix produced by the algorithm after k-th iteration. The stopping criterion is that the residual error ceases to decrease, i.e., when $\left|e^{k+1}{}_{res} - e^{k}{}_{res}\right| / e^{k}{}_{res} < \varepsilon$, for a given $\varepsilon > 0$; we used $\varepsilon = 10^{-4}$. The BC method is non-iterative and the running time is negligible, which is a big advantage over the concurrents.

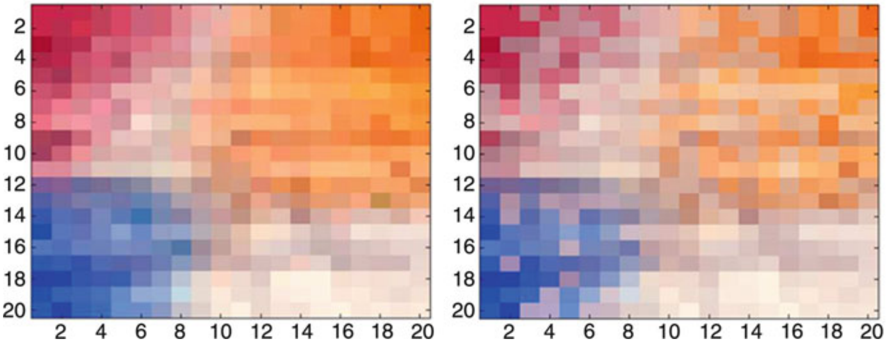

Fig. 1.3 Reconstruction by the IRLS and BC methods (visually indistinguishable; *left*) and by the SVT method (*right*)

As showed in the table below, the IRLS algorithm provided the best result in terms of the residual error, but it took more than 10 min of CPU time. The result achieved by the BC method is only slightly worse, but is obtained almost immediately; the two resulting pictures are visually indistinguishable from each other and shown as one picture in Fig. 1.3 *left*. Finally, the SVT algorithm gave the worst result (Fig. 1.3 *right*); we should note that after a few initial iterations, the SVT method turned into a regime of very slow convergence. Therefore, if we let the algorithm run for a sufficiently long time, a better result probably could have been achieved. However, from a practical point of view, the large computational cost is a very prohibitive issue. Finally, we must point out that the speed of convergence of the iterative methods is largely influenced by the setting of their internal parameters. Further experiments would be necessary, but since this is out of scope of this short paper, we used the parameter values suggested in the corresponding references, [1] and [5].

Method	# iter	CPU time (s)	res. error (%)
SVT	5190	1294	32.27
IRLS	1589	637	5.91
BC	—	0.1	8.64

1.5 Conclusion and Outlook

We showed that the mathematical methods for the matrix completion problem could be a powerful tool for image recovery in art restoration. In particular, the surprising hypo thesis of possible recovery of visible colour information from incorrupted invisible light spectrum has been successfully tested by means of low-rank matrix completion strategies. It seems that for our particular “laboratory experiment”, the simple block completion algorithm would be the method of choice, due to its competitive results and very low computational cost. However, for data measured on real frescoes, where also the invisible parts of the spectra might be influenced substantially by the aging process, more robust methods might be necessary, like the iteratively re-weighted least square minimization. Alternatively, one can choose even more elaborate approaches, for instance the "robust principal component analysis" [10], which is able to detect and filter out large errors in the input data. These and related issues will be subject to further intensive research.

References

1. Cai JF, Candès EJ, Shen Z (2008) A singular value thresholding algorithm for matrix completion. October. arXiv:0810.3286
2. Candès EJ, Recht B (2009) Exact matrix completion via convex optimization. Foundations of Computational Mathematics. arXiv:0805.4471v1
3. Candès EJ, Tao T (in press) The power of matrix completion: near-optimal convex relaxation. IEEE Inf Theory
4. Chistov AL, Grigoriev Dyu (1984) Complexity of quantier elimination in the theory of algebraically closed elds. In: Proceedings of the 11th symposium on mathematical foundations of computer science. Lecture notes in computer science, vol 176. Springer, Berlin, p 1731
5. Daubechies I, DeVore R, Fornassier M, Güntürk CS (2009) Iteratively re-weighted least squares minimization for sparse recovery. Commun Pure Appl Math 35
6. Fornasier M, Haskovec J, Vybiral J (in preparation) The block completion method for matrix reconstruction
7. Fornasier M, Rauhut H, Ward R (in preparation) Iteratively re-weighted least squares for low-rank matrix completion
8. Levinson R, Berdahl P, Akbari H (2005) Solar spectral optical properties of pigments–Part II: survey of common colorants. Sol Energ Mat Sol C 89:351–389
9. Recht B, Fazel M, Parrilo P (2007, submitted) Guaranteed minimum rank solutions of matrix equations via nuclear norm minimization. SIAM Rev. arXiv: 0706.4138
10. Wright J, Ma Y, Ganesh A, Rao S (2009) Robust principal component analysis: exact recovery of corrupted low-rank matrices via convex optimization. J ACM (to appear)
11. Wyszecki G, Stiles WS (1982) Color science: concepts and methods, quantitative data and formulae. Wiley, New York

Chapter 2
3D Modeling: New Method for Quantifying Post-depositional Damages

Leore Grosman, Gonen Sharon, Talia Goldman-Neuman, Oded Smikt, and Uzy Smilansky

Abstract We discuss the results of an experiment which was designed to explore post-depositional damage observed on prehistoric bifaces. Often, post-depositional damages are inflicted when the artifacts are rolled by rivers or waves, and their edges are chipped off by battering. This process was simulated in the experiment, where the damage history was recorded by 3D scanning of the artifacts. The resulting data set was analyzed and the damage patterns were established. To anchor our findings in the Archaeological context, we scanned an assemblage of lithic tools which are known to have undergone battering and damage in a river bed. The implications of our findings to the study of the morphology of lithic tools and their typology are discussed.

Keywords Post deposition damage • Handaxes morphology

2.1 Introduction

The morphology of Acheulian bifaces provided invaluable evidence about human behavior, dispersal and site chronology. Bifaces are classified and compared according to descriptive and metric measures (e.g., [1, 2]; Tixier 1956), tacitly

---

L. Grosman (✉)
Department of Physics of Complex Systems, The Weizmann Institute of Science, Rehovot, Israel

Institute of Archaeology, The Hebrew University of Jerusalem, Mount Scopus, Jerusalem, Israel
e-mail: lgrosman@mscc.huji.ac.il

G. Sharon • T. Goldman-Neuman • U. Smilansky
Institute of Archaeology, The Hebrew University of Jerusalem, Mount Scopus, Jerusalem, Israel
e-mail: gonen.sharon@mail.huji.ac.il; tali.neuman@mail.huji.ac.il; uzy.smilansky@weizmann.ac.il

O. Smikt
passed away on 28th of November 2009, ∎, ∎

H.G. Bock et al. (eds.), *Scientific Computing and Cultural Heritage*,
Contributions in Mathematical and Computational Sciences 3,
DOI 10.1007/978-3-642-28021-4_2, © Springer-Verlag Berlin Heidelberg 2013

assuming that the tool shapes observed today portray accurately the original tools as produced by the knappers. Yet, during the long time span between the production and observation, a broad spectrum of processes could have altered the original shapes. Clearly, the older the assemblage, the harder it is to assess the impact of post depositional processes on the morphology of tools. Several experimental studies have address the issue (Hosfield and Chambers 2004 and references therein) but did not reach the level of heavy damage observed on Lower Paleolithic tools.

Given the difficulties of studying processes that damage ancient tools, we compare experimental and archaeological handaxes to characterize, quantify, and describe post-depositional damages. Assuming that damage is mainly caused due to battering of the tools by random collisions with other stones, we devised an experiment where handaxes were randomly battered by basalt cobbles in a way which simulated the natural processes. The damage history of the experimental handaxes was recorded by scanning their shape in 3D at selected intervals. The resulting digitized 3D images give a precise description of the tool surface along the history of damage.

To anchor the experiment in the Paleolithic context, the artificially induced breakage patterns were compared with the environmental damages observed in the assemblage of flint handaxes from the NBA Acheulian site in the Northern Jordan valley [10, 11]. The NBA assemblage comprised of 290 handaxes and cleavers the majority of them made of basalt. In this area, the Jordan River overflows frequently, and the floods are sometimes sufficiently powerful to change the Jordan course and transports boulders weighing up to 3–4 t [7]. The NBA handaxe assemblage offers therefore an excellent case study for the present investigation. This assemblage has another advantage for the present point of view – it consists of two groups which represent the two extreme points in the life span of a tool: The initial stage is represented by a group found in situ which shows no damage patterns. The end product of the environmental battering process is represented by other tools collected from the site. A sample ($N = 30$) of the complete flint handaxes showing pronounced damage marks was scanned in 3D. The extracted shape parameters were compared to their counterparts from the experimental damage repertory.

2.2 The Simulation Experiment

The experimental setup consisted of a *ca* 100 l cylindrical barrel which could be turned about its horizontal axis with the help of an external shaft. Eight flint replicates of handaxes were selected for the experiment [8]. The handaxes were weighted and 3-D scanned in the same way as the NBA sample before starting the experiment. Fifteen basalt cobbles were placed in the barrel ranging in weight from 250 to 2.920 Kg. All well rounded basalt cobbles originated from a stream running into Lake Kinneret from the Golan and are well rounded dense basalt. The handaxes and the basalt cobbles were placed in the barrel and battering was induced by turning the barrel on its axis. The impact of the container was minimized by using a barrel made of softer plastic material. The turning of the barrel was interrupted after 5, 10, 20, 40, 60, 100 and 200 complete turns, and the handaxes were removed from the barrel to be weighted, 3-D scanned and returned for the next rolling and tumbling interval.

The further analysis of the recorded shapes was carried out off-line, using both available and specially developed software. The first stage of the analysis consisted of positioning the handaxes using a computerized algorithm recently developed [5]. The program computes the inertia tensor which is diagonalized to provide the three principal inertia directions and principal inertia values. The principal directions form a coordinate frame with origin which is placed at the center of mass. The model is then rotated so that the direction of the eigenvector with the smallest principal inertia corresponds to the new z-axis and the maximum inertia to the x-axis. The volume is also computed and recorded. It is important to note that the positioning of the same handaxe somewhat changes between experimental rolling stages due to the mass loss that may change the location of the center of mass and the principal inertia directions. This effect is taken into account when successive models of the rolled handaxes are compared as will be explained in the sequel.

The positioning algorithm provides several views (projections) of the object in the directions of its main axes (Fig. 2.1 shows three views). One can obtain a vivid view of the damage history by following the images of a handaxe as it is subjected to an increasing number of rolls (Fig. 2.2). It is instructive to compare the images extracted for the experimental handaxes and the images obtained by scanning the NBA assemblage, and note the similarity in the damage characteristics, such as the concentration of damage locations and damage intensities.

In order to study the damage history in a quantitative way, one should compare 3-D models of the same handaxe taken before and after an interval of rolling in the barrel. To do so, one should be able to identify in these successive 3-D models, which part of the "parent" volume was removed, or, in other words, how to place the resulting (and smaller) volume within the parent. Luckily, the amount of material lost between successive intervals is limited, and one can always find parts of the boundary which were not affected by the battering during the relevant rolling interval. Fitting these parts and requiring that every point in the model is strictly inside its parent was the main challenge that the algorithm has to face. Here we report on a somewhat simplified version where we fitted the successive 2-D *profiles* of the handaxes within their predecessors. The profile (contour) of a bifacial tool is defined as the boundary of the artifact projection on its ventral plane, here, the (y,z) $(x = 0)$ plane. One can visualize the profile as the curve separating between the illuminated and the shadowed regions when the object is illuminated in the x direction, and the shadow is viewed on the (y,z) plane (Fig. 2.3). Because of the bifacial symmetry, the extreme points on the object whose projections define the profile are likely to belong to the sharp and most fragile edge of the artifact (Fig. 2.3). Hence, most of the damaging events are likely to leave their mark on the profiles.

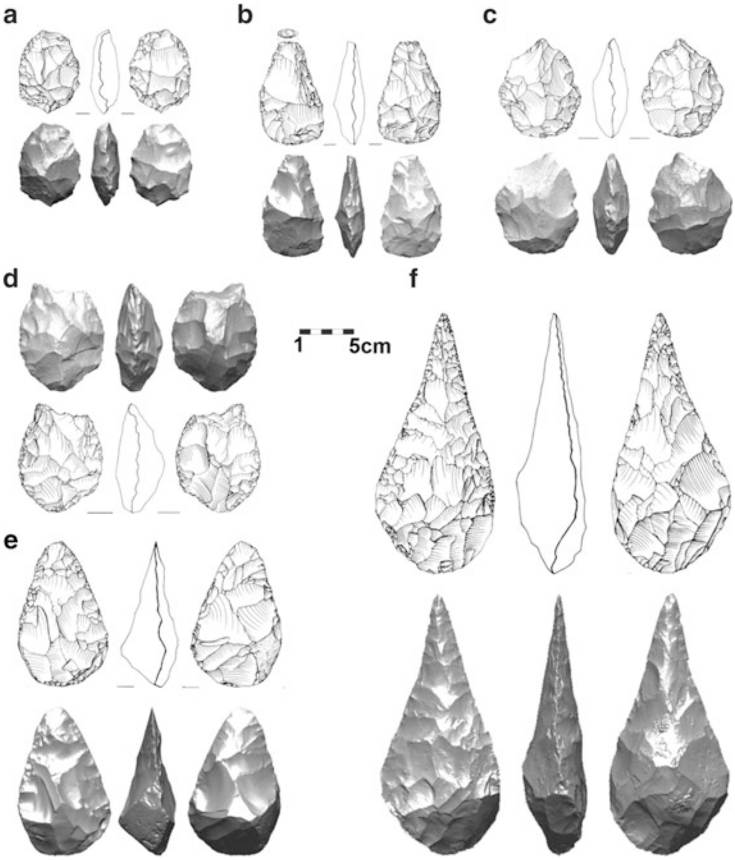

Fig. 2.1 Broken (**a, b, c, d**) and fresh (**e, f**) bifacial tools from NBA site drawn by a draft person and automatically positioned after 3-D scan

Fitting the profiles from each of the scanning intervals within their predecessors results in sequences of encapsulated profiles for each rolled handaxe as shown in Fig. 2.4. The outward profile belongs to the original handaxes before the experiment. This presentation allows us to examine the accumulation of damage on the profiles and to provide quantitative measures to the damage intensity and distribution.

Several characteristic features can be observed by just looking at the complete damage history portrayed by the overlapping 2-D profiles (Fig. 2.4):

1. The inner profiles, which represent the last stage of damage, show the least similarity to the original handaxe since no part of the original lateral edge of any of the handaxes survives after 200 rolls in the barrel.

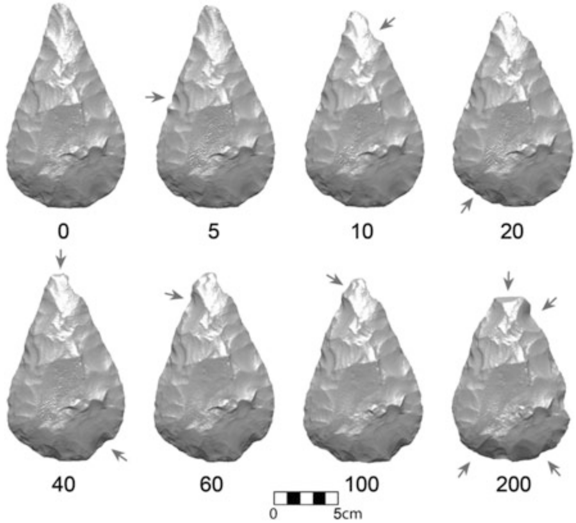

Fig. 2.2 A handaxes drawn in each stage of scanning according to the number of rolls. The *arrows* indicate new scars at each stage observed by the naked eye

2. The initial profile is smooth with hardly any concave sections. Concavity becomes increasingly conspicuous as the damage progresses, and concave sections appear randomly along the profile.
3. In contrast with the tip of the handaxes which is considerably damaged, the proximal edge is the least part of the handaxe to be damaged. This can probably be explained by the differences between the thickness and width of the artifact in its extremities.
4. The original profiles display a fair degree of (mirror) symmetry about the z axis. They also tend to be rather convex. The damaging process tends to increase both their asymmetry and concavity (see below).

In the following sections, these observations will be substantiated by formulating the above observations in a quantitative way which will allow a further statistical evaluation.

2.3 Quantitative Analysis

Having the nested sequence of contours, we quantify the damage intensity and distribution by computing the *damage profile* in the following way: Given a pair of successive contours, one measures the distance *d(s)* between contours at each point s on the outer contour. The larger this distance is the deeper the damage at this point. The function *d(s)* provides the damage profile. Figure 2.5 shows the damage profiles obtained along the *via dolorosa* of one of the handaxes. One can see that the initial damage is mostly concentrated at the tip of the profile (at the vicinity of $s = 250$) with marked loss of area in this part of the tool. During the rolling process,