JOURNAL OF THE AMERICAN INSTITUTE FOR CONSERVATION
https://doi.org/10.1080/01971360.2019.1703483

Routledge
Taylor & Francis Group

TECHNICAL NOTE

Computational Watermark Enhancement in Leonardo's *Codex Leicester*

William A. Sethares [a], Margaret Holben Ellis[b], and C. Richard Johnson, Jr.[c]

[a]Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI, USA; [b]Conservation Center, Institute of Fine Arts, New York University, New York, NY, USA; [c]School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA

ABSTRACT

A better understanding of the inner structural features (watermarks and chain line/wire intervals) of the manuscript papers of Leonardo da Vinci (1452–1519) can provide valuable information regarding chronology, geographic origins, and studio practise. For valid conclusions, however, clear representations of watermarks and chain lines/wires are required. While the easiest way to capture these unique physical characteristics of paper is via transmitted light, they can be difficult to decipher when writing and other surface marks found on the recto (front/obverse) and the verso (back/reverse) of the sheet obscure them. Using Leonardo's *Codex Leicester* (Bill Gates Collection), compiled circa 1508–1510, as a case study, a computational approach was designed to minimize surface writing to enhance the watermarks and chain lines/wires detectable in the paper, by subtracting out diffuse specular (normal) light images of the recto and the verso sides from the transmitted light image. This is accomplished by posing the problem as a mathematical optimization problem, and solving for the optimal weights is fast and robust.

RÉSUMÉ

Une meilleure compréhension des caractéristiques structurelles intrinsèques (filigranes et écartements des fils de chainette/vergeures) des papiers manuscrits de Léonard de Vinci (1452-1519) peut fournir des informations précieuses au sujet de la chronologie, les origines géographiques, et la pratique d'atelier de ce dernier. Cependant, pour obtenir des conclusions solides, une représentation précise des filigranes et des fils de chainette/vergeures est nécessaire. Si la manière la plus simple d'enregistrer ces caractéristiques physiques uniques du papier est par lumière transmise, ces dernières peuvent cependant être difficiles à déchiffrer lorsque des inscriptions ou autres marques présentes à la surface au recto (endroit/avers) et au verso (dos/revers) les obscurcissent. En utilisant le Codex Leicester de Léonardo (Collection Bill Gates), compilé vers 1508-1510, comme cas d'étude, une approche informatique a été conçue visant à affaiblir les inscriptions en surface et intensifier les filigranes et fils de chainette/vergeures discernables dans le papier, en soustrayant les images en lumière spéculaire diffuse (normale) du recto et du verso de l'image en lumière transmise. Cela est réalisé en posant le problème comme un problème d'optimisation mathématique où la détermination du degré auquel les marques en surface de chaque côté doivent être atténuées est rapide et solide. Traduit par Elsa Thyss.

RESUMO

Uma maior compreensão das características estruturais internas (marcas d'água e cadeia de linhas/intervalo de fios) dos manuscritos em papel de Leonardo da Vinci (1452–1519) pode fornecer informações valiosas sobre a cronologia, origens geográficas e práticas de estúdios. Para conclusões válidas, entretanto, são necessárias representações claras de marcas d'água e rede de linhas/fios. Embora a maneira mais fácil de capturar estas características físicas únicas do papel seja através da luz transmitida, pode ser difícil decifrá–las quando escritos e outras marcas de superfície encontradas no recto (frente/anverso) e no verso (a parte de trás/reverso) da folha obscurecem. Utilizando o Códex Leicester de Leonardo (Coleção Bill Gates), compilado entre 1508–1510, como um estudo de caso, uma abordagem computacional foi projetada para minimizar a escrita na superfície para realçar as marcas d'água e redes de linhas/fios detectáveis no papel, pela subtração de imagens difusas por luz especular (normal) do lado da frente e verso da imagem por luz transmitida. Isso é conseguido, colocando–se o problema como um problema de otimização matemática, no qual a determinação da extensão em que as marcas de superfície de cada lado devem ser reduzidas é rápida e robusta. Traduzido por Sandra Baruki.

RESUMEN

Un mejor entendimiento de las características estructurales internas (filigranas o marcas de agua, corondeles y puntizones) de los papeles del manuscrito de Leonardo da Vinci (1452–1519)

ARTICLE HISTORY
Received 13 June 2019
Accepted 9 December 2019

KEYWORDS
Leonardo da Vinci; watermark; optimization; Codex Leicester; Codex Arundel; chain wire/line; laid wire/line; transmitted light images; digital art history; computational art history

CONTACT Margaret Holben Ellis mhe1@nyu.edu Conservation Center, Institute of Fine Arts, New York University, New York, NY, USA

puede ofrecer información de mucho valor con respecto a la cronología, orígenes geográficos, y prácticas de estudio. Sin embargo, con el objeto de obtener conclusiones válidas, se requiere de una representación clara de las filigranas o marcas de agua, puntizones y corondeles. A pesar de que la manera más fácil de capturar estas características físicas únicas del papel es a través de la luz transmitida, puede ser difícil de descifrar con la presencia de escritura y otras marcas superficiales en el anverso y reverso de la hoja, que las oscurecen. Usando como caso de estudio el Códice Leicester de Leonardo (Colección Bill Gates), recopilado hacia 1508 –1510, se diseñó un enfoque utilizando el computador para minimizar la escritura en la superficie y destacar las filigranas o marcas de agua, corondeles y puntizones detectables en el papel, al sustraer imágenes difusas de luz especular (normal) en el anverso y reverso de la imagen obtenida con luz transmitida. Esto se logra planteando el problema como un problema de optimización matemática por el cual la determinación del grado en que las marcas de superficie en cada lado deben reducirse es rápida y robusta. Traducción: Ramón A. Sanchez Chapellín y Vera de la Cruz Baltazar. Revisión: Amparo Rueda.

1. Introduction

Neglected, separated, dispersed, and recombined after his death, the 4,000+ sheets of notes, observations, and drawings by Leonardo da Vinci (1452–1519) present a fascinating puzzle for scholars who aim to recreate their original formats and sequence or simply to better understand the materials and processes behind their fabrication. Leonardo's notebooks or *codices*, whether compiled intentionally during his lifetime or arbitrarily after his death, have long been the focus of art historical, science, literary, and codicological studies. Many efforts to reorder folios and signatures (several folios folded together), produce collation charts, and index and cross-reference individual sheets based on subject matter, page layout, penmanship, continuity of discourse– or lack thereof– and scattered datable references found in the texts have been made (Calvi and Marinoni 1925; Pedretti 1962; Pedretti 1970, 1977; Pedretti 1978). In the case of the *Codex Leicester* (Bill Gates Collection), compiled circa 1508 - 1510, scholarship has been conducted principally by Favaro (1918), Pedretti (1964), Pedretti and Roberts (1981), Veltman and Keele (1986), Farago (1996), Barone (British Library 2019), and Laurenza (2019).

Physical testimony has also been taken into account in theoretical reconstructions of the codices. Paleographical observations include offset writing inks, red chalk transfers, matching fold lines, sewing holes, remnants of blue cartonnage, and the determination of original paper formats and orientation. Watermark identification using the standard compendia of Briquet (1907) and Zonghi and Zonghi (1953) has also promised tantalizing, if highly problematic, clues for questions of dating and geographic origins. Thanks to the Bernstein Consortium (2019) and the Fondazione Fedrigoni Fabriani (2019), these historic tracings made by hand are now available online. Interest in materiality and facture remains strong as demonstrated by Leonardo scholars, including most recently Clayton (2018), Bambach (2019), Barone (British Library 2019), and Donnithorne (2019), whose investigative approach is distinctly "archaeological."

Any program to characterize the physical properties of the papers of Leonardo will require clear representations of watermarks and chain line/wire intervals. Other inner structural features of paper, such as formation patterns (the unique pattern in which pulp is deposited across the mold), and laid line/wire frequencies (the number of laid lines/wires that occur per centimeter), also provoke curiosity about the potential for their use in paper investigations. Because different computational tools would be required, however, these alluring avenues are beyond the scope of this investigation. The study of Leonardo's notebooks and manuscripts is increasingly dependent upon the availability of high-resolution diffuse specular (normal) light images of the folios, and several collecting institutions have already produced transmitted light images of their manuscript pages.

Each sheet of handmade paper contains idiosyncratic physical features that can shed light on its geographic origin and approximate date. This study attempts to record only the presence of watermarks and chain lines/wires. Their presence is a result of the technology of making paper by hand. In the 15th and 16th centuries, paper was made by scooping up macerated and suspended pulp, typically made from processed linen and hemp rags, from a vat using a rectangular mold comprised of a porous screen surrounded by a removable wooden frame. The screen was fabricated from fine densely spaced horizontal rows of "laid" wires tautly held in position by widely spaced vertical "chain" wires. When the mold was plunged into the vat and lifted out, the wires acted as a sieve, retaining the pulp in thinner and thicker accumulations depending upon the degree of interference they produced as the water drained away, as well as the density and quantity of the pulp scooped up. The grid-like configuration of the chain and laid

lines/wires is replicated in the final sheet of paper (Johnson et al. 2015) and can be clearly seen via transmitted light. A watermark, made from a wire bent into a simple shape that was secured to one or two chain or laid lines/wires, can be similarly detected when a sheet of paper is held up to the light (Hunter 1930; Loeber 1982) or illuminated by a strong raking light.

When viewed via transmitted light, watermarks and chain lines/wires, the two features of this study, are often obscured by distracting marks on the recto (front/obverse) and the verso (back/reverse) of the paper, such as drawing media, printing ink, old repairs, and damages. Figure 1 contains details of the watermarks found in folios 1, 2, 3, and 5 of the *Codex Leicester* photographed via transmitted light. Each watermark is located in the right or left center of the sheet, which all measure approximately 300 mm x 450 mm. The watermark is discernable, however, its details are impossible to decipher precisely because of the interference caused by Leonardo's handwriting found on both sides of the sheet. Three of the four watermarks are variants of an eagle with outstretched wings, while the fourth (bottom right) is difficult to identify. Furthermore, the vertical chain lines/wires are imperceptible in all four images.

In addition to viewing differences in paper translucency via transmitted light, there are other means of detecting and recording the inner structural features of paper with varying degrees of accuracy (Ellis and Johnson 2019). Harris (2017) has written extensively about the history of reproducing watermarks and the advantages and disadvantages of each method. Methods dependent upon differences in paper density include tracing by hand or assisted by auto CAD (Calì 2018) over a light source; exposure to a photo-sensitive proofing paper such as DYLUX, largely discontinued; β-radiography; low energy χ-radiography, and reflected infrared (IR) and various multispectral imaging approaches (Harris 2017). Methods dependent upon differences in paper topography include rubbing, ideally on the wire or mold side of the paper, with a soft pencil, largely discontinued; and Reflectance Transformation Imaging (RTI) (Heumiller et al. 2016).

Computational manipulation to increase clarity is required for radiographic, multispectral, and RTI images. Van Staalduinen et al. (2006) compared several such techniques and concluded that transmitted light images can often be made comparable with the more sophisticated techniques by using suitable post processing. Because transmitted light photography is simple, familiar, and easily reproducible, and because transmitted light images are more readily available, the focus of this work is on developing computational means to

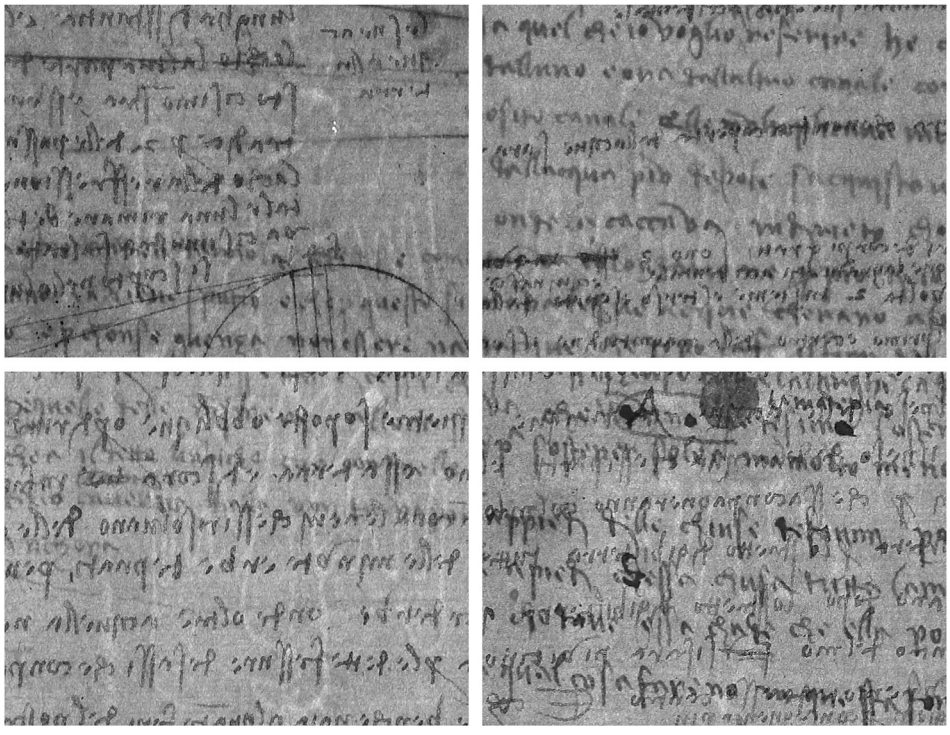

Figure 1. Details of transmitted light images depict four watermarks found in folios 1, 2, 3, and 5 of the *Codex Leicester*. A watermark of a crowned eagle with outstretched wings is partially visible in the top two images, facing in opposite directions. An eagle is also visible (though upside down) on the bottom left, and a watermark is discernable, but not identifiable, in the bottom right. Leonardo's handwriting on both sides of the sheets obscures all the watermarks and other inner structural paper features. (Image: Courtesy Bill Gates).

produce clear images of the inner structure of double-sided works on paper using diffuse specular and transmitted light images.

2. Related work

The general problem of removing visual interference from images has a long history in signal processing (Gonzalez and Woods 2016). Most methods attempt to exploit the known or assumed properties of the obscuring signal. For example, when the desired image occupies a different region of the spectrum from the interference, linear filters can be effectively applied. When the interference can be characterized statistically, for example as impulsive noise, nonlinear filters and optimization techniques can be applied (Rudin, Osher, and Fatemi 1992).

For the problem of removal of obscuring marks from transmitted light images, such general techniques are less useful because it is not easy to characterize mathematically the interference. France et al. (2013) collected a stack of 14 images captured with different wavelengths, combined them via a Principle Component Analysis, and then post processed the aligned images with Photoshop. More recently, Valero (2018) has suggested a simplified technique that uses Photoshop's *color range* tool to select and erase the undesirable features, retouches with a mid-tone color selected from the image to blend in the area, and then adjusts the contrast. For certain images this can effectively remove overlaid writing and the process has the advantage of using familiar tools. Ruiz et al. (2019) implemented a surface image removal technique that divides the pixels of the transmitted light images by the product of the corresponding pixels of the recto and verso images. This approach is analogous to our method, though we use a linear model, which allows for a numerical optimization when combining the three images. The Van Staalduinen et al. (2006) method subtracts the intensity channel of the surface image from the intensity channel of the transmitted light image. The next section presents a method that can be viewed as a generalization that takes weighted versions of two surface images and subtracts them from the transmitted light image. The weights can be determined as the solution to an optimization problem.

A more elaborate approach is presented in recent work by Sabetsarvestani et al. (2019), which examines x-rays of the complex Ghent Altarpiece (circa 1425–1432) by Hubert (1385/90–1426) and Jan van Eyck (c. 1390–1441) painted on both sides of its wooden panels. A convolutional neural network architecture was used to separate the image content from the two sides of the panel using x-rays that were already registered with pixel level accuracy. The neural network model allowed more flexibility in the "subtracting out" process than a linear model.

3. Eliminating interference caused by Leonardo's writing

Imagine that diffuse specular light photographs (Image A and Image B) of the recto and the verso of a folio are available as well as a transmitted light photograph (Image T) of the same folio. Figure 2 shows two diffuse specular light photographs of the recto and the verso of folio 1r of the *Codex Leicester* (Image A and Image B) in the top row. The bottom left image is the transmitted light photograph (Image T), which contains a barely discernable watermark of an eagle.

Now imagine that Images A, B, and T can be aligned so that the locations of the watermarks are superimposed. Observe that full alignment cannot be achieved by changes in lighting because both sides of the paper must be imaged. Once aligned, it possible to "subtract out" the surface writing from the transmitted light image. This can be translated into the optimization problem of finding weights x, y, and z, so that

$$||T - xA - yB - zC|| \tag{1}$$

is as small as possible, where C is a matrix of ones, and where the double bars $||\ .\ ||$ represent a numerical measure such as the two norm. In this way, the weights are set to remove, as much as possible, the obscuring marks of Images A and B from Image T, leading to an enhanced processed image (Image P). The weight z and the corresponding matrix control the overall grey level of the resulting image, essentially adjusting the brightness of the output. The Appendix describes in detail how this problem can be formalized and solved in a computationally efficient manner.

The complete procedure consists of image extraction, alignment, and cropping, followed by a calculation of the optimal parameters and the output image. The extraction of the region about the watermark in the transmitted light image T and the extraction of the corresponding regions from the two surface images A and B can be readily accomplished using available tools (we use Photoshop for these steps). The alignment may be made using Photoshop's Auto-Align Layers functionality, though we use Mathematica's *ImageAlign* function, which allows more flexibility in methods of alignment via the *ImageCorrespondingPoints* options. (The examples of Figures 2–5 were all aligned using the default parameters in *ImageAlign*.) The three images T, A, and B are then cropped to a common size. The optimal parameters x, y, and z, can then be calculated according to

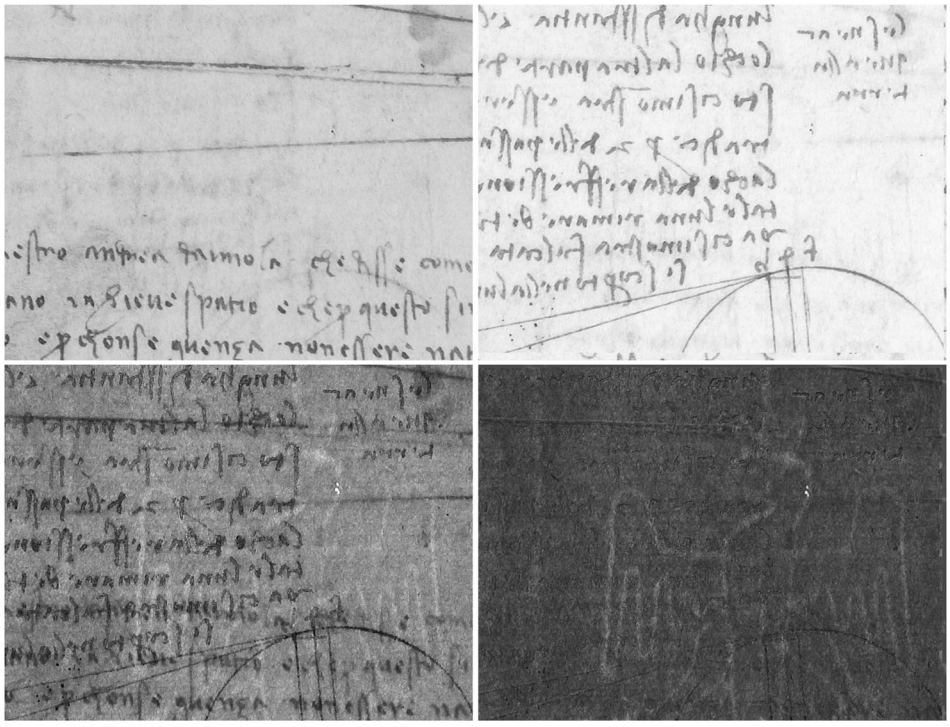

Figure 2. In the top row are two diffuse specular light photographs of the recto and the verso of a portion of folio 1 of the *Codex Leicester* (Image *A* and Image *B*) where the watermark is located. The bottom left image is the transmitted light photograph of the same area (Image *T*), which contains a barely discernable watermark of an eagle. The processed image (Image *P*) in the bottom right reveals a crowned eagle watermark and two chain lines/wires. (Image: Courtesy Bill Gates).

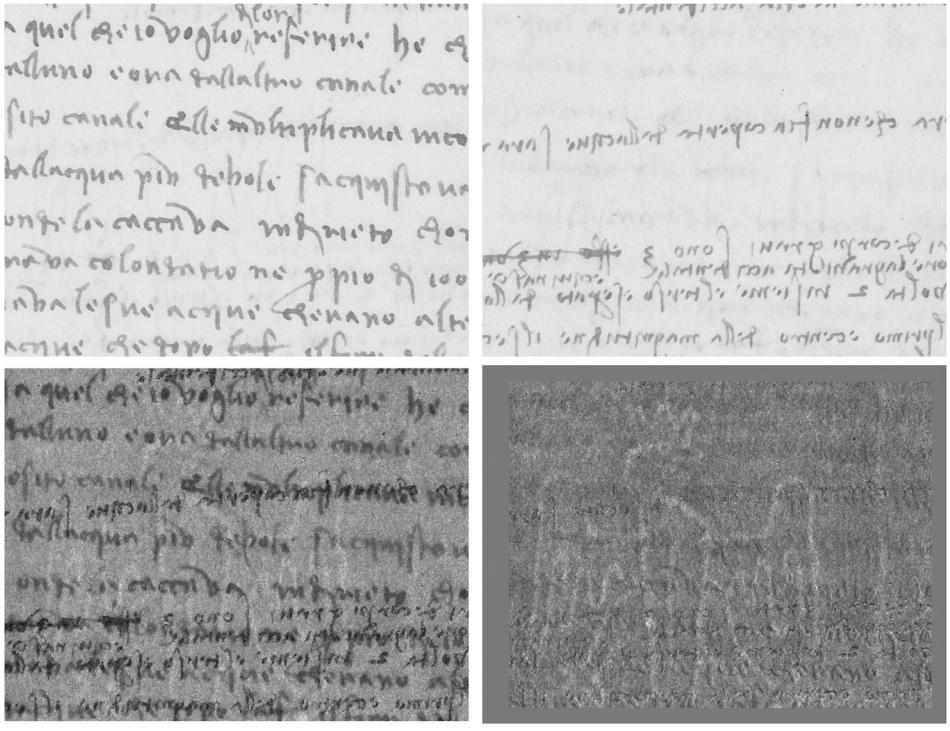

Figure 3. In the top row are two diffuse specular light photographs of the recto and the verso of a portion of folio 2 of the *Codex Leicester* (Image *A* and Image *B*) where the watermark is located. The bottom left image is the transmitted light photograph of the same area (Image *T*), which contains a barely discernable watermark of an eagle. The processed image (Image *P*) in the bottom right reveals a crowned eagle watermark and two chain lines/wires. (Image: Courtesy Bill Gates).

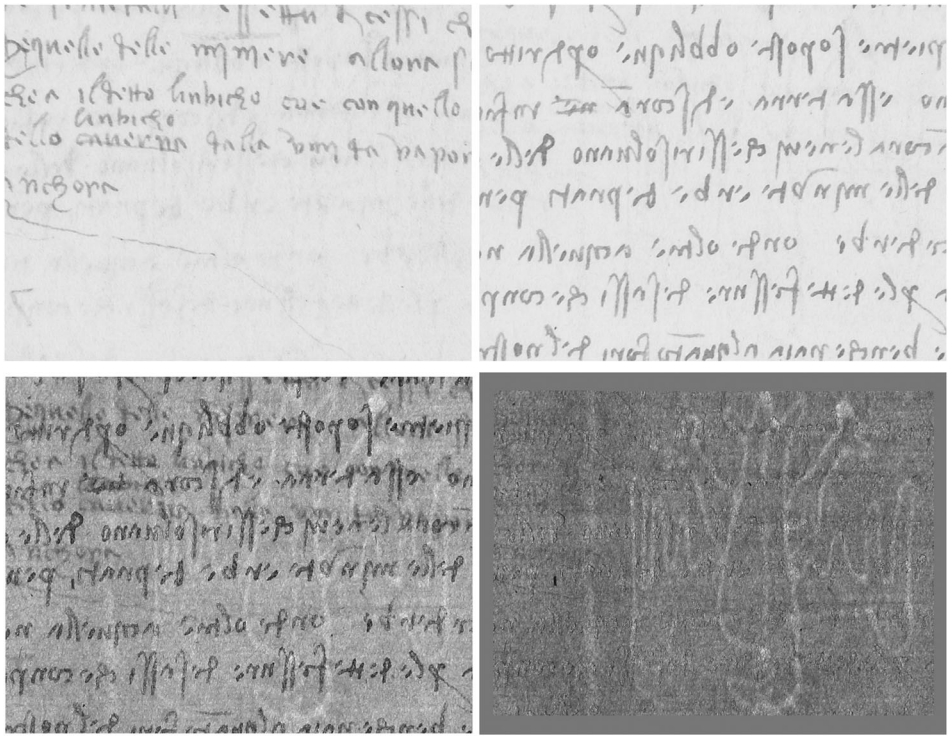

Figure 4. In the top row are two diffuse specular light photographs of the recto and the verso of a portion of folio 3 of the *Codex Leicester* (Image *A* and Image *B*) where the watermark is located. The bottom left image is the transmitted light photograph of the same area (Image *T*), which contains a barely discernable watermark of an eagle. The processed image (Image *P*) in the bottom right reveals a crowned eagle watermark and two chain lines/wires. (Image: Courtesy Bill Gates).

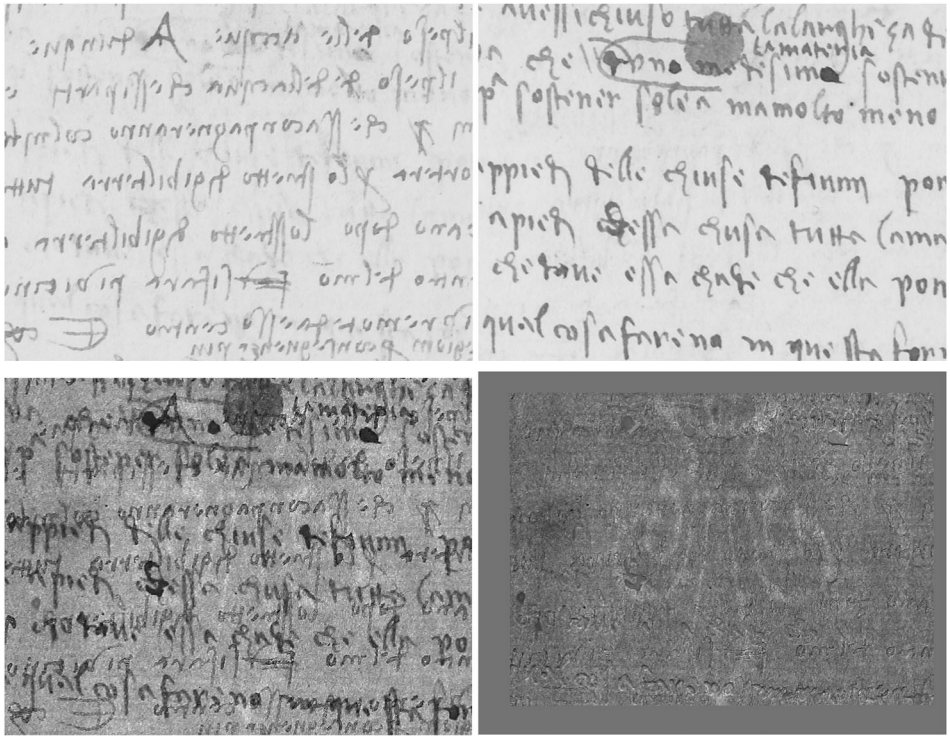

Figure 5. In the top row are two diffuse specular light photographs of the recto and the verso of a portion of folio 5 of the *Codex Leicester* (Image *A* and Image *B*) where the watermark is located. The bottom left image is the transmitted light photograph of the same area (Image *T*), which contains a barely discernable watermark. The processed image (Image *P*) in the bottom right reveals a fleur-de-lis watermark and two chain lines/wires. (Image: Courtesy Bill Gates).

Equation (1), and the output image formed as $P = T - xA - yB - zC$. The final two steps are calculated using a Mathematica notebook which operates using the free downloadable Mathematica player.

Figures 2–5 demonstrate the results of applying this computational enhancement process to each of the remaining three watermarks (Folios 1, 2, 3, and 5) shown in Figure 1. In each case, high quality diffuse specular light or "normal" photographs (Images *A* and *B*) are required of the recto and verso of each folio. Also required is a transmitted light photograph (Image *T*) of the same sheet. Image *P* in the bottom right reveals the enhanced watermark and clearly delineated chain lines/wires. Note that the watermark of Folio 5 is a fleur-de-lis.

4. Discussion

The success of this enhancement procedure is dependent upon achieving good alignment among Images *A*, *B*, and *T* – even a misalignment of a pixel or two can cause disturbing artifacts or even exaggerated image noise. For these examples, *A*, *B*, and *T* were selected as regions around the watermarks and were cut out roughly using Photoshop. Standard image alignment algorithms were run to achieve the pixel-accurate alignment needed for the optimization process. When it is possible to achieve perfect alignment, the optimization does a good job removing the surface features, as suggested by the examples of Figures 2–5.

5. Conclusion

The computational watermark enhancement procedure described above will be part of a package, produced with the help of the Getty Foundation, to be distributed under an open source license. The package will include basic calculations along with a number of features to facilitate its use by researchers with Photoshop-level skills wishing to investigate the inner structural characteristics of any historical laid paper. By developing a successful and easy-to-use computational method of enhancing the watermarks and chain line/wire intervals found in the double-sided folios of the *Codex Leicester*, Leonardo scholars now have access to accurate information about a selection of his papers. Because the procedure uses more readily available diffuse specular and transmitted light images, it is more practical and economical and, ultimately, more feasible than procedures requiring sophisticated imaging such as radiography, reflected IR imaging, or RTI. Based on the success of enhancing watermarks and chain line/wire intervals in the *Leicester Codex*, the analysis of a closely related Leonardo manuscript, the *Codex Arundel* (British Library), has begun.

Acknowledgements

The authors are grateful for the reviewers who contributed valuable feedback and the assistance of Frederick Schroeder and Domenico Laurenza in the preparation of this article. We acknowledge with deep appreciation the financial support from the Getty Foundation (Grant ORG-201943572) and the National Science Foundation (Grant CCF-1822007).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Getty Foundation [Grant Number ORG-201943572]; National Science Foundation: [Grant Number CCF-1822007].

Notes on contributors

William A. Sethares received his PhD degree in electrical engineering from Cornell University, and is currently Professor in the Department of Electrical and Computer Engineering at the University of Wisconsin in Madison. He has held visiting positions at the Australian National University, at the Technical Institute in Gdansk, and at the NASA Ames Research Center. He is a scientific researcher at the Rijksmuseum in Amsterdam and is the Honorary International Chair Professor at the National Taipei University of Technology. His research interests include adaptation and learning in image and signal processing with special focus on the imaging and the interpretation of historical papers and canvases. Dr. Sethares is the author of five books, including *Tuning, Timbre, Spectrum, Scale*, and is co-author of a recent monograph, *Counting Vermeer*, which examines all the extant canvases of Johannes Vermeer. He is grateful to the Getty Foundation for its support of this work. Address: University of Wisconsin, Dept. ECE, 1415 Engineering Drive, Madison, WI 53706, USA. Email: sethares@wisc.edu.

Margaret Holben Ellis is the Chair and Eugene Thaw Professor of Paper Conservation at the Institute of Fine Arts, New York University, where she teaches the conservation treatment of prints and drawings, as well as technical connoisseurship for art historians. She served as founding Director of the Thaw Conservation Center at the Morgan Library & Museum until January 2017. She is a Fellow and current President of the American Institute for Conservation (AIC), a Fellow and past Council member of the International Institute for Conservation of Historic and Artistic Works (IIC), a Fellow of the American Academy in Rome (AAR), and an accredited Conservator/Restorer of the Institute of Conservation (Icon). In the fall of 2015 she was a Getty Conservation Institute Scholar in Residence. She has published and lectured on artists ranging from Raphael, Dürer, and Titian to Pollock, Samaras, Lichtenstein, and Dubuffet. Her research on artists' materials and techniques

is similarly wide-ranging and encompasses Day-Glo colors, Magic Markers, and Crayola crayons. She served as Editor for *Philosophical and Historical Issues in the Conservation of Works of Art on Paper,* published in 2014 by the Getty Conservation Institute. The 2nd edition of her book, *The Care of Prints and Drawings* appeared in 2017. Address: Conservation Center, Institute of Fine Arts, New York University, 14 East 78th Street, New York, NY 10075, USA. Email: mhe1@nyu.edu.

C. Richard Johnson, Jr. is the Geoffrey S. M. Hedrick Senior Professor of Engineering and a Stephen H. Weiss Presidential Fellow at Cornell University. He is also a Research Scientist of the Rijksmuseum and a Senior Research Advisor of the Frick Art Reference Library. His current research focus is computational art history, primarily computer-assisted matching of manufactured patterns in art supports. He has founded several multi-disciplinary, multi-institutional projects, including on automated canvas thread-counting and weave map creation for rollmate hunting among the paintings of Van Gogh and Vermeer and on watermark identification and chain line pattern marking and matching for moldmate hunting among Rembrandt's prints on European laid paper. Professor Johnson's research is currently supported by the National Endowment for the Humanities, the National Science Foundation, the Samuel H. Kress Foundation, and the Getty Foundation. Address: School of Electrical and Computer Engineering, 390 Rhodes Hall, Cornell University, Ithaca, NY 14853, USA. Email: crj2@cornell.edu.

ORCID

William A. Sethares http://orcid.org/0000-0002-0318-7638

References

Bambach, C. 2019. *Leonardo da Vinci Rediscovered. I – IV*. New Haven and New York: Yale University Press.

Bernstein Consortium. 2019. *The Memory of Paper*. Accessed October 16, 2019. www.memoryofpaper.eu.

Briquet, C. M. 1907. *Les Filigranes. Dictionnaire historique des marques du papier des leur apparition vers 1282 jusqu'en 1600*. Paris: Alphonse Picard et Fils.

British Library. 2019. *Leonardo da Vinci: A Mind in Motion*, edited by Juliana Barone. London: The British Library.

Calì, C. 2018. "The Importance of a Project to Enhance the Watermarks of the Codex Atlanticus by Leonardo de Vinci." *5th International Multidisciplinary Scientific Conference on Social Sciences and Arts SGEM* 2018. 23–26 October 2018, 333–338.

Calvi, G., and A. Marinoni. 1925. *I manoscritti di Leonardo da Vinci del punto di vistocronologico, storico e biografico*. Bologna: Zanichelli.

Clayton, M. 2018. *Leonardo da Vinci: A Life in Drawing*. London: Royal Collection Trust.

Donnithorne, A. 2019. *Leonardo da Vinci: A Closer Look. Exploring the Beauty and Complexity of Leonardo's Drawings Through a Study of his Materials and Methods*. London: Royal Collection Trust.

Ellis, M. H., and C. R. Johnson, Jr.. 2019. "Computational Connoisseurship: Enhanced Examination Using Automated Image Analysis." *Visual Resources Special Issue on Digital Art History* 35 (1-2): 125–140.

Farago, C. 1996. *Leonardo da Vinci, Codex Leicester: A Masterpiece of Science*. New York: The American Museum of Natural History.

Favaro, A. 1918. *Per la Storia del Codice di Leonardo da Vinci nella Biblioteca di Lord Leicester*. Florence: R. Deputazione Toscana di Storia Patria.

Fondazione Fedrigoni Fabriani. 2019. *Corpus Chartarum Fabriano (CCF)*. Accessed October 16, 2019. https://ccf.fondazionefedrigoni.it/.

France, F. G., M. Castle, D. de Simone, M. Hill, C. Bolser, S. Albro, and J. Bertonaschi. 2013. "Watermark Capture and Processing with Contemporary Desktop Applications." *Washington, DC: American Institute for Conservation Book and Paper Group Annual* 32 (41): 41–42.

Gonzalez, R. C., and R. E. Woods. 2016. *Digital Image Processing*. 4th ed. New Jersey: Pearson.

Harris, N. 2017. *Paper and Watermarks as Bibliographic Evidence*. Lyon: Institut d'histoire du hivre, 128–132. Accessed October 16, 2019. http://ihl.enssib.fr/paper-and-watermarks-as-bibliographical-evidence.

Heumiller, K., S. Choi, J. Stenger, and C. Graham. 2016. "Post Processing of Reflectance Transform Imaging for Isolation of Surface Impressions," *Archiving 2016: Final Program and Proceedings*. Springfield, VA: Society for Imaging Science and Technology, 15–20.

Hunter, D. 1930. *Papermaking Through Eighteen Centuries*. New York: William Edwin Rudge.

Johnson, C. R., Jr., W. A. Sethares, M. H. Ellis, and S. Haqqi. 2015. "Hunting for Paper Moldmates Among Rembrandt's Prints." *IEEE Signal Processing Magazine* 32 (4): 28–37.

Laurenza, D. 2019. *Leonardo da Vinci's Codex Leicester: A New Edition. Volume I: The Codex*, edited by Martin Kemp. London: Oxford University Press.

Loeber, E. G. 1982. *Paper Mould and Mouldmaker*, edited by L. Hills, and B. van Ginneke-van de Kastelle. Amsterdam: The Paper Publications Society.

Pedretti, C. 1962. *Chronology of Leonardo da Vinci's Architectural Studies After 1500*. Geneva: E. Droz.

Pedretti, C. 1964. *Leonardo da Vinci on Painting: a Lost Book (Libro A) Reassembled From the Codex Vaticanus Urbinas 1270 and From the Codex Leicester. With a Chronology of Leonardo's Treatise on Painting*. Berkeley: The University of California Press.

Pedretti, C. 1970, 1977. The Literary Works of Leonardo da Vinci, edited by J.P. Richter, 3rd ed. London and New York, 1970; with Commentary by C. Pedretti, Berkeley: University of California Press.

Pedretti, C. 1978. *The Codex Atlanticus of Leonardo da Vinci, a Catalogue of its Newly Restored Sheets*. New York: Johnson Reprint Corp.

Pedretti, C., and J. Roberts. 1981. *Leonardo da Vinci: the Codex Hammer: Formerly the Codex Leicester*. Los Angeles: Armand Hammer Foundation.

Rudin, L. L., S. Osher, and E. Fatemi. 1992. "Nonlinear Total Variation Based Noise Removal Algorithms." *Physica D: Nonlinear Phenomena* 60 (1-4): 259–268.

Ruiz, P., O. Dill, G. Raju, O. Cossairt, M. Walton, and A. K. Katsaggelos. 2019. "Visible Transmission Imaging of Watermarks by Suppression of Occluding Text or Drawings." *Digital Applications in Archaeology and Cultural Heritage*. Accessed October 22, 2019. https://doi.org/10.1016/j.daach.2019.e00121.

Sabetsarvestani, Z., B. Sober, C. Higgitt, I. Daubechies, and M. R. D. Rodrigues. 2019. "Artificial Intelligence for Art Investigation: Meeting the Challenge of Separating X-ray Images of the Ghent Altarpiece." *Science Advances* 5 (8). doi:10.1126/sciadv.aaw7416.

Styan, G. P. H. 1973. "Hadamard Products and Multivariate Statistical Analysis." *Linear Algebra and its Applications* 6: 217–240. doi:10.1016/0024-3795(73)90023-2. Accessed October 20, 2019.

Valero, C. 2018. *Enhancing Watermark Images: A Photoshop Method*. Washington, DC: MacMillan Education Center, Smithsonian American Art Museum. Washington Conservation Intern Talks, April 5, 2018. http://www.conservation-wiki.com/wiki/BPG_Annual_Meeting_Tips_Sessions#Enhancing_Watermark_Images:_A_Photoshop_Method_.28Claire_Valero.29 Accessed Oct. 20, 2019.

Van Staalduinen, M., J. C. A. van der Lubbe, F. Laurentius, G.Dietz, and T. Laurentius. 2006. "Comparing X-ray and Backlight Imaging for Paper Structure Visualization." *Proceedings of the International Conference on Electronic Imaging and Visual Arts*, 108–113.

Veltman, K., and K. D. Keele. 1986. *Studies on Leonardo da Vinci I: Linear Perspective and the Visual Dimensions of Science and art*. Munich: Deutscher Kunstverlag.

Zonghi, A., and A. Zonghi. 1953. *Zonghi's Watermarks*. Edited by A.F. Gasparinetti. Vol. 3. Hilversum: Paper Publications Society.

Appendix 1. Solving the Optimization Problem

Let $\mathbf{T}$ represent a transmitted light image, and let $\mathbf{A}$ and $\mathbf{B}$ be images of the corresponding surfaces so that all three are aligned and have the same dimensions. The process of removing the surface features from $\mathbf{T}$ can be stated as the problem of finding parameters x, y, and z such that

$$f(x, y, z) = 1/2||\mathbf{T} - x\mathbf{A} - y\mathbf{B} - z\mathbf{C}|| \tag{1}$$

is minimized, where $||.||$ represents the square of the Frobenius norm and where $\mathbf{C}$ is a matrix of all ones used to automatically adjust the grey-scale intensity of the resulting image. Thus the goal is to locate

$$(x^*, y^*, z^*) = \min f(x, y, z)$$

The cost function $f(x, y, z)$ can be conveniently rewritten in terms of the Hadamard product o (the element-by-element product as in (Styan 1973) as

$$f(x, y, z) = 1/2 \cdot \mathbf{1}_n^T((\mathbf{T} - x\mathbf{A} - y\mathbf{B} - z\mathbf{C}) \text{ o } (\mathbf{T} - x\mathbf{A} - y\mathbf{B} - z\mathbf{C}))\mathbf{1}_m$$

where $\mathbf{1}_n$ is an n-dimensional vector of all ones, and so $\mathbf{1}_n^{\mathrm{T}}\mathbf{N}\mathbf{1}_m$ is equal to the sum of all the elements of $\mathbf{N}$. This can be rewritten again using the fact that the sum of all elements in the product $\mathbf{N}$o$\mathbf{M}$ is equal to the trace $Tr\{\mathbf{N}\mathbf{M}^{\mathrm{T}}\}$, so that

$$f(x, y, z) = 1/2Tr\{(\mathbf{T} - x\mathbf{A} - y\mathbf{B} - z\mathbf{C})(\mathbf{T} - x\mathbf{A} - y\mathbf{B} - z\mathbf{C})^{\mathrm{T}}\} \tag{2}$$

The minimization can be accomplished in several ways. The gradient of the cost function $f(x, y, z)$ with respect to the parameter x is

$$\begin{aligned}\frac{df(x,y,z)}{dx} &= Tr\{\mathbf{A}(x\mathbf{A} + y\mathbf{B} + z\mathbf{C} - \mathbf{T})^{\mathrm{T}}\} \\ &= \mathbf{1}_{\mathrm{n}}^{\mathrm{T}}\,(\mathbf{A}\mathrm{o}(x\mathbf{A} + y\mathbf{B} + z\mathbf{C} - \mathbf{T}))\mathbf{1}_{\mathrm{m}} \\ &= \Sigma_{i,j}a_{i,j}(xa_{i,j} + yb_{i,j} + zc_{i,j} - m_{i,j})\end{aligned} \tag{3}$$

and the derivatives with respect to y and z are similarly

$$\begin{aligned}\frac{df(x,y,z)}{dy} &= Tr\{\mathbf{B}(x\mathbf{A} + y\mathbf{B} + z\mathbf{C} - \mathbf{T})^{\mathrm{T}}\} \\ \frac{df(x,y,z)}{dz} &= Tr\{\mathbf{C}(x\mathbf{A} + y\mathbf{B} + z\mathbf{C} - \mathbf{T})^{\mathrm{T}}\}.\end{aligned} \tag{4}$$

Combining (2) with (3) and (4) allows direct implementation of a gradient descent strategy

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} - \mu f(x_k, y_k, z_k) \begin{pmatrix} \frac{df(x,y,z)}{dx} \\ \frac{df(x,y,z)}{dy} \\ \frac{df(x,y,z)}{dz} \end{pmatrix} \tag{5}$$

where μ is a suitably small stepsize (typically, $\mu << \frac{1}{nm}$). Because $f(x,y,z)$ was chosen to minimize the square of the Frobenius norm, it is also possible to carry out the optimization in closed form. Observe that (3) can be rewritten as

$$\begin{aligned}\frac{df(x,y,z)}{dx} &= Tr\{(x\mathbf{A} + y\mathbf{B} + z\mathbf{C} - \mathbf{T})\mathbf{A}^{\mathrm{T}}\} \\ &= Tr\{x\mathbf{A}\mathbf{A}^{\mathrm{T}} + y\mathbf{B}\mathbf{A}^{\mathrm{T}} + z\mathbf{C}\mathbf{A}^{\mathrm{T}} - \mathbf{T}\mathbf{A}^{\mathrm{T}}\} \\ &= xTr\{\mathbf{A}\mathbf{A}^{\mathrm{T}}\} + yTr\{\mathbf{B}\mathbf{A}^{\mathrm{T}}\} + zTr\{\mathbf{C}\mathbf{A}^{\mathrm{T}}\} \\ &\quad - Tr\{\mathbf{T}\mathbf{A}^{T}\}\end{aligned} \tag{6}$$

and that (4) can be similarly rewritten

$$\begin{aligned}\frac{df(x,y,z)}{dy} &= xTr\{\mathbf{A}\mathbf{B}^{\mathrm{T}}\} + yTr\{\mathbf{B}\mathbf{B}^{\mathrm{T}}\} + zTr\{\mathbf{C}\mathbf{B}^{\mathrm{T}}\} - Tr\{\mathbf{T}\mathbf{B}^{\mathrm{T}}\} \\ \frac{df(x,y,z)}{dz} &= xTr\{\mathbf{A}\mathbf{C}^{\mathrm{T}}\} + yTr\{\mathbf{B}\mathbf{C}^{\mathrm{T}}\} + zTr\{\mathbf{C}\mathbf{C}^{\mathrm{T}}\} - Tr\{\mathbf{T}\mathbf{C}^{\mathrm{T}}\}\end{aligned} \tag{7}$$

The optimal values occur when (6) and (7) are simultaneously zero. Since all of the $Tr\{.\}$ terms in (6) and (7) are scalars, this 3 by 3 matrix problem can solved directly for the optimal values (x^*, y^*, z^*) such that

$$\begin{pmatrix} Tr\{\mathbf{A}\mathbf{A}^{\mathrm{T}}\} & Tr\{\mathbf{B}\mathbf{A}^{\mathrm{T}}\} & Tr\{\mathbf{C}\mathbf{A}^{\mathrm{T}}\} \\ Tr\{\mathbf{A}\mathbf{B}^{\mathrm{T}}\} & Tr\{\mathbf{B}\mathbf{B}^{\mathrm{T}}\} & Tr\{\mathbf{C}\mathbf{B}^{\mathrm{T}}\} \\ Tr\{\mathbf{A}\mathbf{C}^{\mathrm{T}}\} & Tr\{\mathbf{B}\mathbf{C}^{\mathrm{T}}\} & Tr\{\mathbf{C}\mathbf{C}^{\mathrm{T}}\} \end{pmatrix} \begin{pmatrix} x^* \\ y^* \\ z^* \end{pmatrix} = \begin{pmatrix} Tr\{\mathbf{T}\mathbf{A}^{\mathrm{T}}\} \\ Tr\{\mathbf{T}\mathbf{B}^{\mathrm{T}}\} \\ Tr\{\mathbf{T}\mathbf{C}^{\mathrm{T}}\} \end{pmatrix} \tag{8}$$

Observe that the optimal answer can be calculated directly as in (8), or (x^*, y^*, z^*) can be viewed as the solution to which the gradient algorithm (3)—(4) converges. If more data is available, for instance, if there are two transmitted light images

($\mathbf{T}_A$ and $\mathbf{T}_B$, photographed from both sides) the cost function

$$h(x,y,z,a_{A,}a_B) = 1/2||\alpha_A\mathbf{T}_A + \alpha_B\mathbf{T}_B - x\mathbf{A} - y\mathbf{B} - z\mathbf{C}||$$

may be more appropriate, and all five parameters (x, y, z, $\alpha_{A,}$ α_B) can be treated similarly. It may also be desirable in some cases to employ alternatives to $f(x,y,z)$ in (1). For example, minimizing the L_1–norm, using the spectral (eigenvalue) norm, etc. In these cases, the gradient strategy (3)—(4) may still be applied while there may exist no closed form solution analogous to (x^*, y^*, z^*) of (8).