The Gap between Visualization Research and Visualization Software (VisGap) (2022)
C. Gillmann, M. Krone, G. Reina, T. Wischgoll (Guest Editors)

Physical Traces and Digital Stories: Exploring the Connections Between Forensics and Visualization

Victor Schetinger[1] and Saminu Salisu[1]

[1]TU Wien, Austria

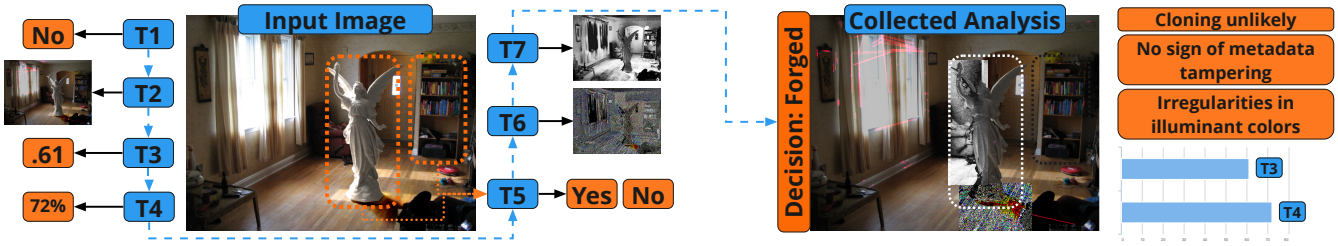

Figure 1: *Example of a possible digital image analysis in an idealized data visualization environment. A forensics expert applies a series of techniques (T1-7) in the image in a continuous process of investigation trying to decide if it has been forged. The output of techniques varies wildly in form, quality and scope, so that the analyst must carefully chose their steps, and how to piece all evidence together. This is essentially a spatial data visualization, oriented primarily by the image pixel space, but also by the physical geometry of the depicted scene. The collected analysis should summarize all relevant facts about the image and provide a coherent story to communicate the analyst's decision.*

Abstract

Forensics, or forensic science, deals with the analysis of evidence for investigation. It is a a wide and strongly interdisciplinary field that needs the coupling of research, practice, and communication to be useful. New techniques have to be constantly developed and applied in the field to solve social conflicts. Recent work suggests, however, that there are many gaps in this coupling, and we argue that there are lessons to be learned from them. Among the difficulties faced by forensics are the management of its interdisciplinarity and over-specialization, and the effective adoption of new research, which are also pressing for the visualization community. In this paper, we bring a gentle introduction to the challenges of forensics with a focus on its digital forms and explore connections to visualization. We believe these connections can be leveraged to further the development of both fields, and particularly that visualization and interaction are critical for the forensics process.

CCS Concepts

• ***Human-centered computing*** → ***Visualization;*** • ***Applied computing*** → *Law, social and behavioral sciences;* • ***Security and privacy*** → *Human and societal aspects of security and privacy;*

1. Introduction

We can start our exposition by stating the most obvious similarity between forensics and visualization: both deal with the analysis of evidence (data), and the exploration of hypothesis. However, for historical and practical reasons they developed independently. Forensics has been tied since antiquity to its use in a wide range of civil and criminal circumstances [Wat10, p. 27]. Before the development of the scientific method the effectivity of forensics was questionable, and to some degree the inventiveness of detective fiction helped shape its modern form [Tho03, p. 5]. All things considered, forensics is still just a component of the legal process, which involves many professions around the world, and ultimately the society in which it is inserted. What keeps everything together is language and communication [Cou16, p. 26]. Therefore, as with visualization, the final goal of forensics is to communicate.

In this paper, we discuss different forms of forensics and contextualize its methodologies and process in relation to visualization. To aid the reader and avoid terminological debates that would be out of scope, we will refer to these in a simplified way. Conventional forensics (CF) stands for the universe of methods based on

DOI: 10.2312/visgap.20221058

crime scene investigation and physical evidence, such as ballistics, fingerprint and DNA analysis, autopsies, and so on. We contrast this with digital forensics (DF), which deals with digital artifacts, of which central to our discussion is digital image forensics (DIF). A few other terms are used for exposition purposes, such as modern forensics, but they are locally contextualized and we hope their meaning is clear within the scope of the explanation.

The different facets of forensics all have their own research communities, methods, and can differ on their basic theoretical foundations, depending on the field they draw most on: medicine, psychology, material science, signal processing, and so on. In fact, this fragmentation is considered a serious issue to the application of forensics [RTWR*15, PKD19, BMG*20]. It is impossible to cover all of them with depth or fairness in this paper. We try to make a general overview with a focus on DIF for four reasons: (1) it is the context the authors have most experience in; (2) the photographic process makes DIF one of the forms of DF that is more grounded in the physical world, which is interesting for our discussion; (3) digital images are an ubiquitous, expressive medium, and (4) the analysis of a digital image can be easily paralleled to an open-ended exploration task (Fig. 1).

We begin by introducing traces (Sec. 2), the basic object of forensics investigation, and connecting them to autographic visualizations [Off20] (Sec. 3) to contextualize conventional and digital forensics in relation to visualization. Then, we discuss the role of narratives within forensics and the need to connect with audiences in both fields (Sec. 4). In Section 5, we present a general framework for DF. Digital traces and digital image forensics are presented as sources of challenges (Secs. 4.1 and 5.1), but also opportunities for visualization research. Finally, a fictitious analysis scenario is presented (Figure 1) to explore DIF tasks within an idealized visualization environment. At the end, we recapitulate the main lessons learned.

2. Traces and Locard's Principle

At the foundations of forensics is the famous Locard's principle [Loc20], which was summarized as *every contact leaves a trace*. Consider, for instance a ballistic forensics scenario within CF. The process of firing a gun leaves many physical traces: after pulling the trigger the cartridge is compressed, gunpowder sprays, etchings are imprinted on the bullet as it leaves the barrel, and when it hits something with high velocity, it either penetrates or ricochets, while being warped by the impact. Each instantiated event of a gun firing is unique in the way both the objects and the environment are transformed, and serve as physical media for a particular fingerprint. While many of these traces might be imperceptible, with the right tools and techniques an analyst can partially reconstruct these fingerprints, generating data about the event.

Taking a picture with a digital camera, similarly, is a process of imprinting traces. The content of the photography is given by the interaction of light with the environment, which then goes through the camera lens, passes the aperture, and hits the sensor, generating a gradient of voltages that is translated to a signal and digitally stored. Every element in this process contributes to the final result, so that the image file also represents a unique fingerprint for that event. Even two pictures taken a fraction of a second apart, in such a way that no human could visually tell the difference between them, will contain very different pixel information. Now, how are those traces used to answer questions, both on the ballistics and digital image forensics case? They provide grounding for narratives, in the same way empirical and experimental data can be used to ground scientific theories, and in all cases visualizing information plays an important role in the interpretation of phenomena.

3. Forensics as a Visualization Process

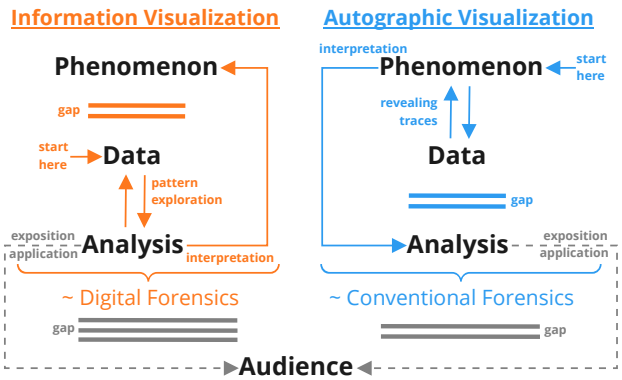

Figure 2: *Approximation of forensics as autographic and information visualization processes based on the framework by Offenhuber [Off20]. This characterization focuses on the interplay between Phenomena, Data, and Analysis. We include the Audience in the process through either the exposition or application of forensics to highlight an additional gap.*

The work of Offenhuber can be used to contextualize both conventional and digital forensics within a visualization framework. In "Data by Proxy - Material Traces as Autographic Visualizations" [Off20], he characterizes two forms of visualization according to the gap between phenomenon, data, and analysis (Figure 2). Information visualization, it is argued, deals with the manipulation of symbolic information, which constitutes a serious decoupling from the original phenomena to be analyzed, even if it generated the data. Autographic visualizations, in contrast, are materialized traces, indexes to the phenomena that can be used in themselves for the analysis (Figure 3). Autographic examples would include a mercury-in-glass thermometer as a visualization of room temperature, or a cut-section of a tree as a visualization of its age. In the former, the variable is encoded in the volume dilation of fluid inside the container, while in the later, in the tree rings that can be counted.

Through Locard's principle, one could say that CF deals with autographic visualizations, while DF can be paralleled to information visualization. We believe this inference to be interesting for our contextualization, but it is exaggerated for exposition purposes, and should not be taken too literally. For one, the process of visualizing traces in forensics is more linked to pattern matching to establish relationships between entities [Mar17], rather than measuring variables. Furthermore, modern laboratory equipment is digital and will provide data as output, blurring the comparison a bit.

Figure 3: *Example of autographic visualization. Through the attrition caused by people trying to insert the pin at the desired weight, traces are generated. One can estimate a distribution of the more common weights used in this machine (in orange) without resorting to statistics and symbolic manipulation.*

When Locard originally formulated his principle, over a hundred years ago, the definition of a trace had a much more precise scope, limited by the techniques and scientific understanding of the time about what was physically possible. DNA evidence, for instance, was not part of the criminal ontology, and digital traces are an even more recent addition. However, his principle still hold if we address the nature of traces. Traces are signs, more specifically indexes, referencing some aspect of their circumstances of imprinting (the phenomenon). We propose an extension of Locard's principle that is useful for understanding modern forensics, and relates to Offenhuber's model: *actions generate traces, and traces generate patterns, which can then be causally linked to events*.

3.1. Gap Matching

The important thing to highlight is the distinction between mainly material or symbolic indexes being manipulated for analysis, and how this distance from the phenomena could affect interpretation. This was the main goal of Offenhuber when proposing these two characterizations of visualization, and he further offered a combined, more seamless model. Interestingly, this is very much in line with critiques of forensics science that argue that over-specialization and decoupling of processes created dangerous gaps, and a more holistic approach should be pursued [RTWR*15, BMG*20].

In the digital forensics scenario these gaps can also be observed in different forms. In a DF investigation of a user's online behavior through activity logs both the phenomenon (e.g. user accessing a website) and the analysis can be considered within the digital realm. In DIF, however, a scene or event (physical phenomenon) is turned into data (digital representation of image) through photography, and all analysis happens on the symbolic domain, over pixels and aggregated data (Figure 1). Therefore, in DIF there is a larger gap between phenomena and analysis, and this results in serious challenges for interpretation, which are further discussed in Section 5.1.

In Figure 2 we re-frame Offenhuber's model to point another important gap: between analysis and audience. The audience is as crucial for forensics as it is for visualization, and it must be able to understand how each level unfolds: from action to trace, from trace to pattern, from pattern to event, and from event to story. A common criticism of forensics can be characterized as a gap with the audience, either through failure of application or exposition [RTWR*15, BMG*20, Tho19, SF08]. This can be caused by many factors such as the complexity involved in explaining results, the lack of access to resources or new technologies, poor education, and untrustworthy techniques. In the case of DF, the gap is even wider: digital traces are more distant from real phenomena, therefore harder to articulate, and due to dynamic nature of technology still allude proper regulation [Hor19]. In the next section, we explore the different roles within the scope of audience for forensics, which in its wider form can be considered society as a whole, and the narrative processes that tie them together.

4. Forensics as Storytelling

The general goal of forensics is to form causal chains between agents, objects, and events, that will resolve a conflict [Tho19]. If a bullet found on a victim can be linked to a gun, the registered owner of the gun implicitly becomes a character in this story. They might not be the one who pulled the trigger, but they are legally bounded to be responsible for that object, and are therefore accountable for it. If it was lost or stolen, for instance, nevertheless this person will be involved in the process and be a link in the causal chain. Furthermore, not all objects must be explicitly present. Even if the bullet was not found, or if the crime weapon is missing, other evidence may be brought in to support a narrative, such as a receipt for ammunition of a specific caliber, traces of gunpowder, or an empty holster. Therefore, forensics is bounded by the social processes that require such narratives, their rules, conventions, and language [Cou16].

It is important to realize that the legal process, and therefore the construction of legal narratives vary wildly around the world. Not only the legislation itself, but what is considered socially acceptable, culture, and the economic infrastructure all weigh on the extent to which forensics will be used and justified [Tho03, SF08]. In Brazil, for example, there are legal precedents for using psychographed letters (as in, spiritually channeled by a medium) in court [Pit17]. The idea of case law, which draws on past decisions to resolve future conflicts, further creates a collective social reality by reinforcement.

The role of the forensics expert, then, is regulating the trust in this storytelling process by grounding evidence in reality, through knowledge, skills and techniques [BMG*20]. The authority of the expert is a crucial aspect in this social network, and is greatly backed by the ontological success of its discipline [Wat10, ch. 3, 6]. The advances in medicine in the last couple centuries made healthcare ubiquitous, and medical science one the pillars of conventional

forensics, notably through the role of the coroner. A doctor is a respected specialist that can communicate a story to a judge or jury, and, even if they don't fully understand the details, their familiarity as patients awards epistemological weight to a doctor's words.

4.1. Digital Stories

What about, then, of a digital forensics expert? While a medical story can be somewhat related to the experience of visiting ones general practitioner, a legal narrative involving data objects can have a surreal tone [Pol09], and the handling of its elements be challenging to law agents [Bel19, Gog10]. In April 2018, when Mark Zuckerberg made a testimony before congress as part of the Cambridge Analytica trials, it was clear that the congressmen did not have a grasp on the technology involved in Facebook, and he took advantage of this in his defense by obfuscating communication. The congressmen were very experienced in policy, and probably make use of a variety of digital technologies in their daily lives, but lacked the epistemological foundations to understand its inner workings and articulate it in their discourse. If they had access to experts in DF, they might have been able to better express their concerns and evaluate Facebook's accountability, but since technical, intangible entities are at play, there might be no legal precedent to directly draw conclusions from. The EU general data protection act (GDPR) [VVdB17] can be seen as an effort to regulate the digital world, but legal systems are still struggling to adapt [TID19,ZB20].

If a forensics expert travelled back in time to the 1800s with the equipment to perform DNA analysis and tried to use it to solve crimes, they would not be very effective. The understanding of what DNA is and how it constitutes evidence was not available, and therefore the expert's arguments would not make credible stories for their peers (time travelling notwithstanding). Digital stories, or rather narratives about digital objects are still not well ontologically grounded. They are not derived from material traces and interactions, which adhere to the laws of physics, chemistry, or biology. Rather, they are technical articulations. Were the algorithms for compression, or the standards for text encoding been developed differently, many of our tools and theories would fail [IFP21, Hor19]. However, we can trust that if our time travelling expert brought working equipment and followed protocol, they would be able to draw veritable conclusions from the DNA analysis, even if useless in a court of that time. In the end, the usefulness of forensics is bounded by the ability of society in articulating it, of using it in narratives [Sch18, p. 136].

In this regard, we can find another connection to visualization, represented in Figure 3 by the arrows between the Analysis and the Audience. The ability of the public to understand a visualization will limit its impact (exposition) or usefulness (application), in the case of a visualization system. Both visualization literacy and domain knowledge might be required to properly understand a graph or interact with an application. The visualization literature, however, is much more concerned with the user and in solving possible knowledge gaps, for instance through guidance and onboarding [SCW*22], while the forensics literature is focused on developing forensics techniques [BMG*20, SIPO17, Ver20, FFdCJS20].

5. Digital Forensics and Visualization

To be able to properly deal with digital stories within an investigation framework, one can draw on a few of the available methodologies such as the digital forensics phases (Figure 4) by the National Institute of Standards and Technology (NIST). It includes data collection, data preservation, data processing and data visualisation [AMM11, Mar14], and as such could be considered a visualization pipeline by itself.

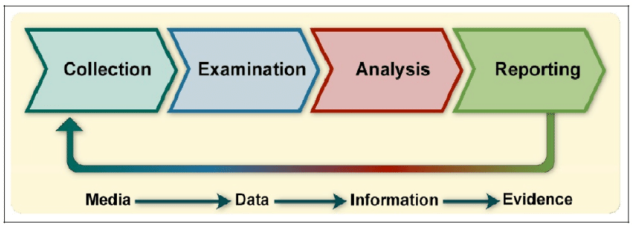

Figure 4: *Diagram of different the phases in a digital forensics investigation processes based on the NIST framework [AMM11]. The phases includes data collection, examination, analysis and reporting. In each phase, it is considered that there is a development such that media is held at collection, and only at reporting it can be considered evidence.*

Collection, the first step, and the whole management of forensics data can be extremely challenging [TMC17b, QLL*14], due to the large amount of data needed and its heterogeneity. This is much more serious in CF than in DF, where evidence is physical and needs to be physically stored, taken care, and accounted for [RTWR*15]. Hybrid approaches can also exist, for instance by having pictures of evidence, which is another reason why DIF is so important.

The final step in the digital forensics NIST investigation step (reporting), requires presentation of evidence collected, preserved and processed from the investigation. The evidence must assume a digital form and can involve some form of visualisation in the case of quantitative data with accompanying media files. This final stage of the digital forensics process highlights the importance of a good visualisation to communicate with the audience.

The datasets involved in digital forensics processes emanate from digital devices such as mobile phones, computers and smart systems, which naturally generate large amounts of information [Hal17]. When it comes to DF, almost everything that can be digitally tracked can be used for analysis in some form, from phone calls [CFF13] to network traffic [CT20]. The work by [Hal17] explored the use of visualisation techniques to reduce the amount of time and effort in analysis, while increasing investigative efficiency and accuracy. Tassoni et al. [TMC17a] further explored the issues practitioners face in an ever changing mobile centric and portable electronic device society, and provide an in-depth summary of visualization techniques that are used in DF.

TimeSets is another example of an effective visualisation technique built purposely to aid intelligence investigation [SXW*16]. TimeSets is a timeline visualisation showing sequence of events

with sets relation. As part of a study carried out with domain experts in [SXW*16], observation on effective visualisation technique included use of color to indicate severity of a situation, grouping information by trust level, and information display position with top placed information trusted more than bottom information.

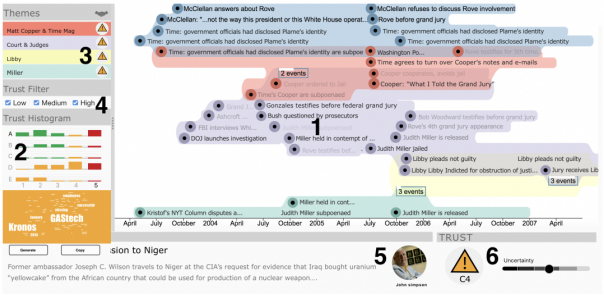

Figure 5: *Dashboard for TimeSets with uncertainty intelligence visualisation [SXW*16]. Events are organized according to severity and case.*

5.1. Digital Image Forensics

DIF, as the name suggests, deals with the analysis of digital images. It is a particularly challenging field because it is strongly tied to both digital artifacts and human factors. Besides (or, in spite of sometimes) being traces for criminal narratives, digital images are ubiquitous in our communication. DIF can be seen as a special part of DF that still lacks proper contextualization within a larger investigation framework, and techniques that are more considerate of the audience (Figure 3): both analysts that need to use them (application), which are the audience for new research, and the general public that needs to consume the output of analysis (exposition).

The *raison d'être of* DIF is also solving conflicts of trust between human peers: who took this photo? Is this image fake? Is the person in this picture John Doe? However, the types of questions that can be made about a digital image and the types of answer that can be provided vary wildly, and cannot stand the same ontological scrutiny as CF. For instance, it is impossible to say that an image is true in any sense, only that there is no evidence it has been tampered from capture to storage in the camera [SIPO17, Far18]. To further the problem, digital photographs of evidence have been increasingly used for analysis due to the practicality this offers to the investigation pipeline [EKN20], increasing the gap with the phenomena and turning many CF scenarios partially into DIF.

Medium and Message There is a strong dichotomy between medium and message in DIF. The digital format, compression, place of storage, and even the physical elements used in its capture can be considered part of its medium, while the message or content is the intangible visual information instantiated in it. Modern camera pipelines have many components and steps that transform the light information from the scene: from the lens, color filter array, sensor architecture, to the camera software that automatically tries to enhance picture quality. On the one hand, this is great for forensics, as per Locard's principle each component leaves its individual trace, and therefore the final image is causally linked to every one of them, but on the other it makes it impossible to determine what transformations should be allowed in an image to preserve its "truthfulness" to an event [SIPO17].

In practice, the majority of DIF techniques focus on the medium, and have a similar application than the ballistic forensics case: matching traces to the components that generate them. Very subtle details, such as the chromatic aberration of the lens [JF06], or the noise of the sensor [SIPO17] can be used to link a picture to a camera, or to identify a region of the image that has been altered from its expected trace pattern. However, these techniques can be weak against attacks such as compression or filtering [IFP21].

Tasks Some techniques in the DIF arsenal are automated, and process images without any need for human input, while other techniques require interaction by the analyst selecting parameters, areas in the image, or reacting on iterative outputs. A proper analysis should include the outcome of many techniques, since the information each provides is very limited in scope and confidence. This can be compared to having the opinion of multiple physicians before accepting a diagnosis. However, combining the output of multiple techniques is an extremely complex problem [FARTB13], and still one of the open challenges of DIF.

Very broadly speaking, there are two main tasks in DIF: provenance identification and tampering detection. The first deals with linking images, devices, and people, which is practically similar to CF, and focuses on the medium. Tampering detection tries to identify if an image has been modified (maliciously or not), and will consider both medium and content. An analyst working on either case generally will have a toolkit with an array of preferred or sanctioned techniques, and will use them to appraise an image. Their actual praxis is determined more by experience and local factors than any state-of-the-art. A sad reality is that the majority of actual DIF work deals with investigating sexual abuse of children [WKRR22, FBAMM18], and there is very little research that helps one know how to go about analyzing an image in general. The literature is lacking on use cases, methodologies, and guidelines, and published techniques provide no support for prospective users [SIPO17, Ver20].

6. Fictional Scenario: The Missing Angel

Let us consider a fictitious image forensics analysis to understand how it could work in practice. Imagine the following case: a client of an insurance company is claiming he was robbed and many objects were stolen from his house, including a very expensive angel sculpture he had worked on. As part of the proof, he sent a picture he had taken a few days before of the angel in his living room. However, this picture was a clever forgery made by inserting a synthetic 3D model within a picture [KSH*14] with the intent of defrauding the insurance company for extra money. Suspecting something was wrong, the company hired an expert to analyze the images provided. How would that analysis go?

Generally speaking, a DIF analyst will have a toolbox of preferred techniques, many of which will be black boxes. There are some commercial solutions available, but to the best of our knowledge not a clear standard that is adopted by experts. In the recent

years, due to the popularity of deep learning and the rise of deep fakes, many forensics techniques have been proposed with a "deep" approach [Ver20], however, they are limited in their interaction and visualization capabilities. In this fictional scenario, let us imagine our analyst has a forensics tool developed by a visualization researcher, that allows him to plug-in techniques, execute them and compose their outputs (Figure 1). Since techniques generally vary greatly in output and confidence [SIPO17], the analyst will execute the minimum amount of techniques that cover the broader possibilities to avoid false positives.

6.1. Analysis Techniques

The first obvious step is examining the image metadata (T1), as many editing software leave watermarks that can be easily caught. If nothing is found, a second good options is looking for cloning traces, which involve copying and pasting image patches and are very common in forgery. There are automatic techniques that can correlate similar regions in the image (T2), but they might be meaningless. Similarly to a data exploration process, the analyst might then not know what to do, and decide to run automated black-box techniques that crunch statistical features and give some score estimation that the image has been tampered (T3, T4). These techniques are very common and can be effective, but they provide results that are hard to interpret [Ver20], contextualize [SIPO17], and combine [FARTB13]. Some of them might require selecting a region of interest (T5), which helps the analyst in isolating interesting elements and dealing with non-explanatory outputs. By testing objects in the room, the analyst realized there was something wrong with the angel.

The most revealing techniques are those that allow the human in the loop to combine hunches and localized insights without prescribing judgement, such as those that output maps of features in image or pixel space (T5, T6). This can be understood in terms of the relationship between Phenomenon, Data, and Analysis (Figure 3). Image space still preserves a close relationship to the phenomenon in the sense that pixels have a mapping to the physical world, and information about the scene can be obtained simply through gazing at the photo. This is not the case with T1, T3, T4, and to some extent T5, where the outputs can only be interpreted by referring to the semantics of the algorithm (sometimes black boxes), and have no spatial reference. This makes it harder for the analyst to apply their results during the analysis (application), and to use them in argumentation (exposition). Therefore, they are less efficient in bridging the gap between the Analysis and the Audience.

Image-space techniques might estimate a local compression errors for pixels in the whole image, or the local correlation between neighboring pixels [SIPO17]. Recalling Locard's principle, even small changes in an image might leave traces. Adding or removing objects cause discontinuity because no two scenes have the exact same lighting conditions, no two cameras are the same, and while two files can be compressed in the same way, making an exact match is very burdensome. These details might be imperceptible to the human eye [SOdSC17], but such techniques act as magnifying glasses of difference. Through analyzing the outputs of T5 and T6, the analyst realized inconsistencies in illuminant colors and shading, and it was clear that the image had been forged.

6.2. Interfacing

Having a proper analysis environment is crucial for both executing the analysis, and for the subsequent reporting of results. If we switch the image data in the previous example for map data, with the outputs from T2, T7, and T6 being different map layers such as temperature, vegetation, height, etc. the task of detecting forgery can be abstracted in a geovisualization context as, for instance, finding an appropriate area for building. It is essentially a matching of constraints over regions of interest. However, DIF is rarely treated as a visualization problem by the community, and therefore it lacks the proper treatment of its effective tasks.

In Figure 1 the reader can imagine a system that provides the analyst with the option of selecting techniques to execute from an extensive pool, and then layering and organizing the results so that, at the end, a composed picture aggregates all evidence (right image). In the case of the missing angel this would be invaluable, as the analyst must defend their position that the image has been forged. The accused client might cast doubt against the analysis, as is the nature of the legal process, and even hire their own analyst. The more external parties get involved in a dispute, the more crucial is the potential for synthesis and exposition that a forensics tool might offer. Once again, this is not unlike visualization, where collected insights might impress domain experts but fail to convince organizational stakeholders or be effectively implemented in processes.

7. Lessons Learned

In our exposition we have covered a large amount of subjects. First, we to tried establish isomorphisms between forensics and visualization (Figure 2 by linking the concept of trace to its use in autographic visualizations, as proposed by Offenhuber [Off20]. Then, we explored the fundamental role of narratives, which both motivate and guide the practice of forensics, by discussing its social and technical aspects. Finally, we zoomed in DF and particularly DIF to analyze in depth some of the challenges that might be present and provide a practical example.

Here we try to summarize a few of the most important lessons. There are interesting connections that arise from the similarities between visualization and forensics, for both deal with data, analysis, and human factors, and these provide opportunities for visualization to collaborate with forensics. However, there are also some structural issues that we can identify as societal challenges, affecting both fields.

7.1. Opportunities

DF As highlighted by [Hal17], Visualisation techniques already play an important role on modern day digital forensics processes which likely involves analysing large dataset from smart devices such as mobile phones, computers and even smart watches. Effective visualisation techniques can save investigators and information consumers time and effort by narrowing down areas on interest from big data samples using methods such as timeline visualisation

and geographical map visualisation to limit the scope of investigation to a specific time and place.

DIF The forensics analysis of a digital image can be paralleled to a data exploration task, where an expert is looking for evidence (insights) about the image, but many tasks can be open ended, since it is impossible to exhaustively test all techniques and possibilities contained within the data (the digital image). Guidance and onboarding could be deployed to aid analysts in using techniques and testing hypothesis [Sch18, ch. 4]. Surprisingly though, discussion on the role of visualization, interaction, and supporting tasks is almost completely non-existent in the DIF literature [FFdCJS20,SIPO17,Far20], which focuses on individual techniques. An ideal DIF tool might look like a visualization dashboard where different insights and outputs of techniques would be combined and super-imposed on the image during the analysis process, collaborating to sense-making and the subsequent exposition of results to the audience.

7.2. Challenges

Organizational The most pervasive and hard to solve problems both in forensics and visualization come from coordination between organizations, which hinders collaboration between individuals. Simply put, the most severe criticism of forensics [RTWR*15, RRC18, SF08] is rooted at compartmentalization. A researcher and a police investigator are under different pressures, and their organizational settings afford them different opportunities, priorities, and resources. Therefore, even with the best goodwill by all individuals involved, cooperation implies competition [KFS08]. When this is compounded by all the different agents and organizations meshed together, systematic issues emerge. In visualization this is clearly illustrated by the difficulty in collaborating with domain experts, which has been long recognized a a challenge in the field [SMM12]. Even after successful research and development of visualization (or forensics) technology, the key to its adoption is often the engagement of the right stakeholders.

Interdisciplinarity the theoretical foundations of the different disciplines involved in forensics and its societal applications are not always compatible, and this leaves dangerous open gaps [Sch18]. This can be observed not only in the contrast between "science" and "law", but also between CF and DF, because empiric investigations of physical and digital phenomena should not be treated equally. Within visualization, this subject is often tackled by the digital humanities [Jän16, SRF*19, HEAB*17], where it is argued that the convenience and efficiency of digitally processing data brushes over subtle but crucial human factors. This is a serious issue in applications of law, where automatic techniques and machine learning are employed to offset the high cost of forensics, generating dangerous black boxes with power overs peoples' lives [DF18,Hor19]. The development of better inter- and trans-disciplinary epistemological theories is still paramount for both visualization and forensics, and their relationship with the audience.

8. Conclusion

In this paper, we have introduced the field of forensics, with a focus on its digital form. We tried to explore the subject not by treating it as an application domain for visualization, but perhaps as a parallel, different form of visualization tied to the legal system that developed on its own. Forensics is essential to society with its role of bringing scientific and technical developments to aid the resolution of conflicts. Through its challenges, one can learn important lessons about bringing research to practice and improve understanding on the gap between visualization research and visualization software.

Acknowledgements

This work was partially supported by the Austrian Science Fund (FWF) as part of the project KnoVA (#P31419-N31), and HumaneAI (EU Horizon #761758).

References

[AMM11] AZEMOVIĆ J., MUŠIĆ D., MOSTAR: Efficient model for detection data and data scheme tempering with purpose of valid forensic analysis. 4

[Bel19] BELSHAW S. H.: Next generation of evidence collecting: The need for digital forensics in criminal justice education. *Journal of Cybersecurity Education, Research and Practice 2019*, 1 (2019), 3. 4

[BMG*20] BAECHLER S., MORELATO M., GITTELSON S., WALSH S., MARGOT P., ROUX C., RIBAUX O.: Breaking the barriers between intelligence, investigation and evaluation: A continuous approach to define the contribution and scope of forensic science. *Forensic Science International 309* (Apr. 2020), 110213. URL: https://linkinghub.elsevier.com/retrieve/pii/S037907382030075X, doi:10.1016/j.forsciint.2020.110213. 2, 3, 4

[CFF13] CATANESE S., FERRARA E., FIUMARA G.: Forensic analysis of phone call networks. *Social Network Analysis and Mining 3*, 1 (2013), 15–33. 4

[Cou16] COULTHARD M.: *An Introduction to Forensic Linguistics*. Routledge, July 2016. URL: https://doi.org/10.4324/9781315630311, doi:10.4324/9781315630311. 1, 3

[CT20] CLARK D., TURNBULL B.: Interactive 3d visualization of network traffic in time for forensic analysis. In *VISIGRAPP (3: IVAPP)* (2020), pp. 177–184. 4

[DF18] DRESSEL J., FARID H.: The accuracy, fairness, and limits of predicting recidivism. *Science Advances 4*, 1 (2018), eaao5580. URL: https://www.science.org/doi/abs/10.1126/sciadv.aao5580, arXiv:https://www.science.org/doi/pdf/10.1126/sciadv.aao5580, doi:10.1126/sciadv.aao5580. 7

[EKN20] EDIRISINGHE P., KITULWATTE I., NADEERA D.: Knowledge, attitude and practice regarding the use of digital photographs in the examination of the dead and living among doctors practicing forensic medicine in sri lanka. *Journal of Forensic and Legal Medicine 73* (2020), 101995. 5

[Far18] FARID H.: Digital forensics in a post-truth age. *Forensic science international 289* (2018), 268–269. 5

[Far20] FARID H.: Image Forensics. In *Computer Vision*. Springer International Publishing, Cham, 2020, pp. 1–10. URL: http://link.springer.com/10.1007/978-3-030-03243-2_877-1, doi:10.1007/978-3-030-03243-2_877-1. 7

[FARTB13] FONTANI M., ARGONES-RÚA E., TRONCOSO C., BARNI M.: The watchful forensic analyst: Multi-clue information fusion with background knowledge. In *2013 IEEE International Workshop on Information Forensics and Security (WIFS)* (2013), IEEE, pp. 120–125. 5, 6

[FBAMM18] FRANQUEIRA V. N., BRYCE J., AL MUTAWA N., MARRINGTON A.: Investigation of indecent images of children cases: Challenges and suggestions collected from the trenches. *Digital Investigation 24* (2018), 95–105. 5

[FFdCJS20] FERREIRA W. D., FERREIRA C. B. R., DA CRUZ JÚNIOR G., SOARES F.: A review of digital image forensics. *Computers & Electrical Engineering 85* (July 2020), 106685. URL: https://www.sciencedirect.com/science/article/pii/S0045790620305401, doi:10.1016/j.compeleceng.2020.106685. 4, 7

[Gog10] GOGOLIN G.: The digital crime tsunami. *Digital Investigation 7*, 1-2 (2010), 3–8. 4

[Hal17] HALES G.: Visualisation of device datasets to assist digital forensic investigation. doi:10.1109/CyberSA.2017.8073402. 4, 6

[HEAB*17] HINRICHS U., EL-ASSADY M., BRADELY A. J., FORLINI S., COLLINS C.: Risk the drift! stretching disciplinary boundaries through critical collaborations between the humanities and visualization. 7

[Hor19] HORSMAN G.: Tool testing and reliability issues in the field of digital forensics. *Digital Investigation 28* (Mar. 2019), 163–175. URL: https://www.sciencedirect.com/science/article/pii/S1742287618303062, doi:10.1016/j.diin.2019.01.009. 3, 4, 7

[IFP21] IULIANI M., FONTANI M., PIVA A.: A leak in prnu based source identification—questioning fingerprint uniqueness. *IEEE Access 9* (2021), 52455–52463. doi:10.1109/ACCESS.2021.3070478. 4, 5

[Jän16] JÄNICKE S.: Valuable research for visualization and digital humanities: A balancing act. In *Workshop on Visualization for the Digital Humanities, IEEE VIS* (2016). 7

[JF06] JOHNSON M. K., FARID H.: Exposing digital forgeries through chromatic aberration. In *Proceedings of the 8th workshop on Multimedia and security* (2006), pp. 48–55. 5

[KFS08] KEYTON J., FORD D. J., SMITH F. L.: A mesolevel communicative model of collaboration. *Communication theory 18*, 3 (2008), 376–406. 7

[KSH*14] KARSCH K., SUNKAVALLI K., HADAP S., CARR N., JIN H., FONTE R., SITTIG M., FORSYTH D.: Automatic scene inference for 3d object compositing. *ACM Transactions on Graphics (TOG) 33*, 3 (2014), 1–15. 5

[Loc20] LOCART E.: *L'Enquête Criminelle et les Méthodes Scientifiques*. Ernest Flammarion, Paris, 1920. 2

[Mar14] MARTURANA F.: *DEVICE CLASSIFICATION IN DIGITAL FORENSICS TRIAGE*. PhD thesis, 06 2014. doi:10.13140/2.1.2399.0725. 4

[Mar17] MARGOT P.: Traceology, the bedrock of forensic science and its associated semantics. In *The Routledge international handbook of forensic intelligence and criminology*. Routledge, 2017, pp. 30–39. 2

[Off20] OFFENHUBER D.: Data by Proxy — Material Traces as Autographic Visualizations. *IEEE Transactions on Visualization and Computer Graphics 26*, 1 (Jan. 2020), 98–108. doi:10.1109/TVCG.2019.2934788. 2, 6

[Pit17] PITTELLI M.: Psicografia como meio de prova judicial. *Vianna Sapiens 1* (09 2017). 3

[PKD19] PIETRO D. S., KAMMRATH B. W., DE FOREST P. R.: Is forensic science in danger of extinction? *Science & Justice 59*, 2 (2019), 199–202. URL: https://www.sciencedirect.com/science/article/pii/S1355030618302454, doi:https://doi.org/10.1016/j.scijus.2018.11.003. 2

[Pol09] POLLITT M.: Digital forensics as a surreal narrative. In *Advances in Digital Forensics V* (Berlin, Heidelberg, 2009), Peterson G., Shenoi S., (Eds.), Springer Berlin Heidelberg, pp. 3–15. 4

[QLL*14] QI M., LIU Y., LU L., LIU J., LI M.: Big data management in digital forensics. In *2014 IEEE 17th International Conference on Computational Science and Engineering* (2014), pp. 238–243. doi:10.1109/CSE.2014.74. 4

[RRC18] ROUX C., RIBAUX O., CRISPINO F.: Forensic science 2020–the end of the crossroads? *Australian Journal of Forensic Sciences 50*, 6 (2018), 607–618. 7

[RTWR*15] ROUX C., TALBOT-WRIGHT B., ROBERTSON J., CRISPINO F., RIBAUX O.: The end of the (forensic science) world as we know it? The example of trace evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences 370*, 1674 (Aug. 2015), 20140260. URL: https://royalsocietypublishing.org/doi/10.1098/rstb.2014.0260, doi:10.1098/rstb.2014.0260. 2, 3, 4, 7

[Sch18] SCHETINGER V.: Beyond digital, imagens, and forensics: towards a regulation of trust in multimedia communication, 2018. 4, 7

[SCW*22] STOIBER C., CENEDA D., WAGNER M., SCHETINGER V., GSCHWANDTNER T., STREIT M., MIKSCH S., AIGNER W.: Perspectives of visualization onboarding and guidance in va. *Visual Informatics 6*, 1 (2022), 68–83. 4

[SF08] SAKS M. J., FAIGMAN D. L.: Failed Forensics: How Forensic Science Lost Its Way and How It Might Yet Find It. *Annual Review of Law and Social Science 4*, 1 (Dec. 2008), 149–171. URL: https://www.annualreviews.org/doi/10.1146/annurev.lawsocsci.4.110707.172303, doi:10.1146/annurev.lawsocsci.4.110707.172303. 3, 7

[SIPO17] SCHETINGER V., IULIANI M., PIVA A., OLIVEIRA M. M.: Image forgery detection confronts image composition. *Computers & Graphics 68* (2017), 152–163. 4, 5, 6, 7

[SMM12] SEDLMAIR M., MEYER M., MUNZNER T.: Design study methodology: Reflections from the trenches and the stacks. *IEEE transactions on visualization and computer graphics 18*, 12 (2012), 2431–2440. 7

[SOdSC17] SCHETINGER V., OLIVEIRA M. M., DA SILVA R., CARVALHO T. J.: Humans are easily fooled by digital images. *Computers & Graphics 68* (2017), 142–151. 6

[SRF*19] SCHETINGER V., RAMINGER K., FILIPOV V., SOURSOS N., ZAPKE S., MIKSCH S.: Bridging the gap between visual analytics and digital humanities: Beyond the data-users-tasks design triangle. In *Workshop on Visualization for the Digital Humanities, IEEE VIS* (2019), vol. 2019. 7

[SXW*16] SALISU S., XU K., WAGSTAFF A., BIGGS M., PHILLIPS G.: TimeSets for Uncertainty Visualisation. In *Computer Graphics and Visual Computing (CGVC)* (2016), Turkay C., Wan T. R., (Eds.), The Eurographics Association. doi:10.2312/cgvc.20161291. 4, 5

[Tho03] THOMAS R.: *Detective Fiction and the Rise of Forensic Science*. Cambridge Studies in Nineteenth-Century Literature and Culture. Cambridge University Press, 2003. URL: https://books.google.at/books?id=s6y9uaTPiBoC. 1, 3

[Tho19] THORNTON J. I.: Uses and abuses of forensic science. In *Science and Law: An Essential Alliance*. Routledge, 2019, pp. 79–90. 3

[TID19] TARAKANOV V. V., INSHAKOVA A. O., DOLINSKAYA V. V.: Information society, digital economy and law. In *Ubiquitous Computing and the Internet of Things: Prerequisites for the Development of ICT*. Springer, 2019, pp. 3–15. 4

[TMC17a] TASSONE C., MARTINI B., CHOO K.-K.: Forensic visualization: survey and future research directions. In *Contemporary digital forensic investigations of cloud and mobile applications*. Elsevier, 2017, pp. 163–184. 4

[TMC17b] TASSONE C. F., MARTINI B., CHOO K.-K. R.: Visualizing Digital Forensic Datasets: A Proof of Concept. *Journal of Forensic Sciences 62*, 5 (2017), 1197–1204. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1556-4029.13431. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/1556-4029.13431, doi:10.1111/1556-4029.13431. 4

[Ver20] VERDOLIVA L.: Media forensics and deepfakes: An overview.

IEEE Journal of Selected Topics in Signal Processing 14, 5 (2020), 910–932. doi:10.1109/JSTSP.2020.3002101. 4, 5, 6

[VVdB17] VOIGT P., VON DEM BUSSCHE A.: The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing 10*, 3152676 (2017), 10–5555. 4

[Wat10] WATSON K. D.: *Forensic Medicine in Western Society*. Routledge, Nov. 2010. URL: https://doi.org/10.4324/9780203840290, doi:10.4324/9780203840290. 1, 3

[WKRR22] WILSON-KOVACS D., RAPPERT B., REDFERN L.: Dirty work? policing online indecency in digital forensics. *The British Journal of Criminology 62*, 1 (2022), 106–123. 5

[ZB20] ZAEEM R. N., BARBER K. S.: The effect of the gdpr on privacy policies: Recent progress and future promise. *ACM Transactions on Management Information Systems (TMIS) 12*, 1 (2020), 1–20. 4