Abstract—We put forward a framework to address a problem created by the rapidly spreading use of imaging devices and related to involuntarily or unintentionally photographed individuals: their pictures can accumulate additional meta information via face recognition systems and can be manually tagged via social networks and publishing platforms. With this framework a user can express his/her picture privacy policy in a machine readable format and (to some extent) automatically enforce it. An easily understandable flag system is used to define restrictions on picture usage and linkability. This policy is encoded in an unobtrusive way into wardrobe patterns and accessory designs with almost no impact on apparel appearance or social interaction.

Index Terms—privacy invasion, involuntary photographs, unintentional photographs, DRM

I. INTRODUCTION

Imaging devices have infiltrated every corner of modern life. They are omnipresent in multiple forms such as photographic cameras, security cameras, and mobile phones. Products like Google Glass [1] have introduced wearable computing devices to the public, potentially enabling the recording of everything at anytime (c.f. Omniveillance [2]). Not only do these devices digitally document the life of the user, they also capture other individuals nearby.

People can feel uncomfortable about losing control over their pictures, and there are serious privacy implications [3] due to the massive publication of private pictures and other information along with them. Face recognition is built into many picture publishing systems such as Picasa and iPhoto and into social networks such as Facebook. Pictures enriched with personal (meta) information (Figure 1) can eventually end up in search machine indices, hence providing searchability by name, face similarity, date, and/or geographic location.

Picture rights, privacy, and publishing related problems are not a novelty of the Internet age, but they are amplified by it. Pictures in print media and their privacy implications are also regularly the subject of legal proceedings.

II. MOTIVATION

Many countries define rights regarding a person’s own image. However, they are not easy for a person to enforce. The image of a person might have been unintentionally captured by a photographer without the person noticing that his/her picture was being taken, the person may simply not know the photographer, or the person may not know when and where his/her picture was published and in which context. This lack of knowledge can hinder the person from exercising his/her legal rights. Moreover, the person has no way to inform potential or actual picture takers of their self-chosen restrictions on how their image shall be handled.

Likewise, a conscientious photographer might not have the chance to ask all the people whose image he/she captured for their consent to use their images. In any case, the person’s right to control how his/her image is used is lost due to a gap in the communication and control path from the person to
the photographer and/or publisher of the photo. Additionally, different countries regulate this right differently: some tie it to the act of publishing the picture while others tie it to the act of taking the picture.

A possible solution to the communication gap problem is to create a central database of privacy policies. Such a database could either use face recognition features or a unique code embedded in each person’s clothing as a lookup key. However, this method facilitates identification and creates a linkability capability that might contravene a person’s picture privacy policy. Furthermore, it creates a central database that could be a single point of failure and that could be misused for surveillance purposes.

III. CONTRIBUTION

Our proposed Personal Picture Policy Framework (P3F) eliminates the gap in communication from the photographed person to the photographer and/or publisher of the person’s picture. It incorporates a simple flag-based system that covers the most important restrictions a person might impose on his/her own picture. It is similar to the Creative Commons [4] system used for copyright restrictions on creative works (e.g., by photographers for their pictures).

A modular visual coding system is used to convey the policy information across the communication gap described above. The policy is embedded in the visual information of the photograph (e.g., as part of the clothing), making it an inseparable part of the picture so that it is highly likely to survive along the publishing path (Figure 1). Under favorable conditions, this information is hidden in such a way that it is unnoticed by the human eye. Hence, we call it Privacy Policy Hiding.

Our proposed formal logic is used to combine multiple policies found in one picture and to determine how to handle potential usage changes when the picture is passed from one entity to another for which other parts of the policy can be relevant (e.g., a picture and its meta information on a social networking site being indexed by a third-party search engine).

Our proposed automated system can be built into publishing software, social networks, and search engines so that they handle pictures appropriately on the basis of the relevant policy (blur out faces of people who do not want their images to be published, discard specified meta information, etc.).

IV. RELATED WORK

The World Wide Web Consortium (W3C) created a Platform for Privacy Preferences (P3P) Specification [5] to enable automated processing as well as human readable display of web site privacy practices. Since these policies can be quite complex, several authors created simplified human readable iconic representations [6]–[8], with Rundle [9] being one of the first. However, as Parsons in his “Privacy Commons” [10] points out, most icon sets lack the visual and semantic clarity and simplicity needed for the broad public. Most attempts have tried to cover too much detail to be understandable to users without an in-depth introduction. To be practical, such a system must focus on the main properties, even if it does not cover all special cases.

Several systems have been proposed for the World Wide Web for minimizing the personal digital footprint. Besides the Do Not Track (DNT) header for web browsers currently being standardized by the W3C [11], there are a number of cookie, tracking, and advertisement blocker extensions for most browsers.

Various methods have been proposed for self-defending an individual’s privacy against face recognition. As face mummification is not socially or legally accepted everywhere, several methods attempt to defeat the face-finding algorithms used for pictures. The most common algorithms use Haar-like feature classifiers [12] for computationally lightweight face detection before further processing using more resource-consuming algorithms (e.g., bio-metric identification).

To inhibit the feature response of these algorithms, Harvey uses hair styles and make-up [13], while Yamada et al. [14], [15] uses infrared light sources in a pair of goggles that are visible to most camera sensors but invisible to the human eye. The hair-style and make-up approach is very time consuming in preparation and visually very dominant. It therefore hinders everyday social interaction and can provoke unwanted reactions. The goggles approach is much less intrusive but requires a constant power supply and infrared LEDs that can keep up with the ambient light. Another approach [16] is to bombard the camera with enough infrared light to create a back-lit condition that darkens the rest of the image. This is only feasible at short distances indoors as it is hard to compete against the much stronger daylight.

All three approaches are a form of digital mummification.
V. STRUCTURE AND ENVIRONMENTAL CONSIDERATIONS

Our proposed P3F sits on neuralgic nodes in the distribution path (Figure 2) where it decodes the privacy policies and automatically applies them as described below.

The framework has many constraints to fulfill to be feasible for the user as well as the publishing site or search engine operator. Users need a simple system that is easy to understand and use and that does not hinder them in any of their choices (like their wardrobe style or social interaction). Operators (once they agree or are forced to comply) want a computationally lightweight solution. Technical difficulties arise in finding a coding scheme that can convey a privacy policy whether it is photographed as a close-up or wide shot. Furthermore, the coding scheme has to fulfill the aesthetic needs of the user as well as constraints due to work or social etiquette.

A. Picture Privacy Policy

A usable policy needs to focus on key aspects to be easily understood and ultimately gain user acceptance. Our framework consists of three simple person-related restrictions (Table I) and two picture-wide restrictions (Table II).

1) Personal Flags: Flags are attached to a specific person in a picture and are applied individually. This means that it might be possible to find a picture by using the meta data of one person in the picture but not that of another person in the picture. That is, each person in a picture can set his/her privacy settings individually and gain informational self-determination.

The Do not Search flag specifies that the user does not want to be found through an internal or external search engine using a person-specific keyword. This includes the person’s real name, user name, birth date, and any other indexable data. Furthermore, it includes other images (e.g., “find similar faces,” “find other pictures of the same user”) or joined data (e.g. “other customers who bought this product,” “friend of the person”). In the case of Facebook, the user accepts being identified (“tagged”) in a photo but does not want this photo to show up if someone searches on his/her name or visits his/her timeline. However, the picture still can be included in an index based on the geolocation or timestamp information (e.g., Flicker’s "map this picture" feature).

The Do not Identify flag specifies that the user does not want to be identified in a picture. This includes automatic face identification as well as manual name tagging by other users. If this information should become available by other means despite this specification, it is not to be included in a search index.

The Do not Publish flag specifies that the user does not want to have any pictures of him or her published. If the person is not the main subject (e.g., his/her image was unintentionally captured) his or her face should be blurred, pixelated, or covered to make identification impossible. The publisher (e.g., newspaper editor, blog writer, or uploading social network user) can also crop the picture to exclude the person in question. A modern publishing system can blur faces automatically in accordance with P3F policy. (For justified exceptions, see Section V-B.)

2) Picture Flags: Two additional picture-related flags complete the privacy policy (see Table II). The No Geolocation flag specifies that geographic location should not be added, displayed, or indexed for this picture, and the No Timestamp flag specifies that a timestamp should not be processed for this picture.

3) Flag Precedence: Flags are encoded in symbols, markings similar to 2D barcodes, patterns similar to 1D barcodes, or facilitating other visual techniques (e.g. watermarking) onto one’s wardrobe (e.g., shirt or jacket) or onto accessories (e.g., hat, cap, or button) as described in Section VII-A. Multiple visual encoding schemes are needed to blend unobtrusively into the desired wardrobe style, as this is often predetermined by external factors. Some of these codes might be so subtle that are unnoticed by the human eye - such as a 1D barcode in a stripe pattern on a tie or t-shirt.

However, it is infeasible for a person to carry around a set of shirts and change them in accordance with each occasion during the day. The person may therefore have different policies attached to different articles of clothing, with a defined precedence. The precedence order follows the ease with which a specific article is changeable in public. For example, one does not normally change a pair of trousers in public, so the policy attached to one’s trousers has the lowest precedence. Since ties, scarves, and jackets are easier to change and since a cap or button can be changed on the fly, they have the highest precedence.

A person can display the No restriction policy flag encoded on an article of clothing with the highest precedence.
to cancel out his/her fallback or default policy encoded on another article of clothing.

If multiple policies are displayed on different articles of clothing with the same precedence level, the individual restrictions are added up. More specifically, policy $P$ consists of an $n$-tuple of restrictions or flags $\langle R_1, R_2, \ldots \rangle$. The $n$-th policy of precedence level $p$ for an individual $i$ is denoted as $P_{i,p,n}$. The effective policy $EP$ is the $n$-tuple of all the strongest individual restrictions at the highest precedence level $p_{max}$, found for $i$. For person-related restrictions, let $\langle \text{G}<\text{F}^{\circ}<\text{X}<\text{P}<\text{P} \rangle$, whereas $\langle \text{G}<\text{F}^{\circ} \rangle$ and $\langle \text{X}<\text{P} \rangle$.

$$EP_i = \left\{ \max(P_{i,p_{max},1\ldots n}[R_1]), \max(P_{i,p_{max},1\ldots n}[R_2]), \ldots \right\}$$

B. Photographer’s and Publisher’s Stakes

P3F was designed as an opt-out procedure to publishing systems for use by individuals wanting to restrict how their picture is used. Although this is not the most privacy supportive design, a system that does not include the current reality as the default will have a hard time gaining broad acceptance.

P3F does not strictly prevent publishing of pictures in contravention of the user’s policy. Manual exceptions are allowable for two main reasons.

1) Side Agreement: The photographed person gave his/her consent for publishing a specific picture to the photographer before or after the photo was shot.

2) Justified Exceptions: The are generally several legal exceptions to the rights a person has regarding his/her picture. They include exceptions on the use of pictures by the media of people of public interest and on the use of police booking photographs.

Theses exceptions might be implemented in a publishing system (e.g. social network or newspaper system) as an additional manual work step to override a restriction. In such instances, the publisher (e.g. a social network user, blog poster, or newspaper editor) must confirm that he/she has the permission of the photographed person to publish the picture or that an exception applies.

As the original P3F policy remains part of the picture, another publisher that (re)uses the picture will also have to contact the photographed person for permission to publish or state the exception that applies.

This mechanism also accounts for possible mistakes and false positives that may occur during processing.

VI. ENFORCEMENT

Since P3F is an opt-out system, the impact on the status quo for the industry is minimized. Nevertheless, reservations could remain and thus reduce the incentive for wardrobe producers to offer such coded clothing. However, two examples demonstrate how similar systems were successfully adopted in the past.

After a public outcry shortly after the introduction of Google Street View, the service started to blur faces and license plates [17]. In Germany, Google additionally agreed to provide an opt-in feature after the Minister of Justice of Rhineland-Palatinate, the data protection supervisor for Schleswig-Holstein, and Germany’s Federal Consumer Protection Minister threatened the company with legal action. Since 2009, German home owners can blur the image of their home [18].

Another example is the integration of a banknote detection algorithm in popular software (e.g., Photoshop and PaintShop Pro), several printers, several scanners, and most color copying machines [19]. In 2004, the Central Bank Counterfeit Deterrence Group [20] (founded by the G10) published a Counterfeit Deterrence System software module for detecting banknotes that has subsequently found its way into many products although it is only available as a closed source module and there is no legal obligation for companies to include it.

VII. TECHNICAL ARCHITECTURE

The overall architecture of P3F is displayed in Figure 3. Clothing and accessory manufacturers can use the P3F encoder to create a visual marking or pattern that matches or blends into any wardrobe style. Individuals can use it as well for some accessories (e.g., a button) and some articles of clothing (e.g., homemade articles).

Once a picture of the user wearing something with the user’s policy embedded ends up in a publishing system, two other components are used to implement the policy. The P3F decoder searches for and decodes the embedded policy and attaches it to the appropriate person (or face). System integrators might decide to keep the original image and the extracted P3F information in their database and apply them at the point of publishing (Figure 3, “Publishing 1”). Others might decide to immediately run the picture through the P3F picture manipulator to remove the meta data and blur the appropriate faces (Figure 3, "Publishing 2").
A. Visual encoding

In its current form (Section V-A), an individual's policy is encodable in six bits. Thus, the encoding scheme does not have to offer a high data density, but it must meet certain other technical requirements.

1) Technical requirements:

a) **Illumination stability:** The code should be decodable under a wide range of lighting conditions. However, under conditions making face identification impossible, a decoding failure is tolerable.

b) **Blurriness tolerance:** Picture blurriness can arise from sub-optimal auto-focus mechanisms because the photographer actually focused on another object or person or moved the camera during exposure (a common problem with amateur photographers).

c) **Size and clipping invariance:** The code should be decodable from shots with different fields of view. Therefore, it should be so redundant that a partial capture in a close-up produces results as good as those in a wide shot. Furthermore, in a wide shot, a larger part of the code is recorded but with a reduced resolution compared to a close-up. Fine encoding that repeats multiple times is better for close-up shots while coarse encoding is better for wide shots. Ideally, a code unifies both traits.

d) **Distortion stability:** People do not always face the camera head-on, especially when they are being photographed unintentionally. Furthermore, the human body is not a flat board, and loose clothing tends to fold and wrinkle. Another faults may arise from lens distortion or improper washing or drying of the person's clothing.

e) **Noise robustness:** Another artifact introduced by cameras is noise, especially in low-light and low-contrast situations due to the automatic camera gain amplifying the sensors background noise.

f) **Computational weight:** The detection algorithm should be lightweight because operators of publishing systems will most likely demand one that conserves computational resources.

g) **Compression stability:** Digital photography greatly depends on picture compression algorithms. They commonly destroy details in pictures and introduce artifacts. These algorithms are often based on a psycho-visual model of human visual perception and are therefore not optimized for computer vision purposes. The most common compression method for photographs on the Internet is JPEG.

h) **Blind decoding ability:** The decoder should have the ability to decode the data without prior knowledge of the original pattern used to encode the data or the data that is being looked for (a common prerequisite for some watermarking techniques).

i) **Detection accuracy:** Detection accuracy should be high with a slight bias toward false positives since people typically feel more comfortable with more privacy than with less. False positives can still be overridden by the publisher if necessary.

2) Aesthetic Demands:

a) **Dress code:** Dress codes are often imposed by society, the employer, or another external entity. The coding scheme should thus produce markings and patterns that blend into the imposed dress code.

b) **Fashion:** People additionally often have their own fashion demands. The coding scheme should thus produce markings that blends into the individual's fashion style.

c) **Adaptive:** Clothing is sold in many different colors and shapes. The code should thus be versatile and work with many different colors and shapes.

d) **Unobtrusive:** The application of P3F should require only a slight adjustment in clothing style. The code should be subtle with low visual impact. It should be unrecognizable by other people, thus minimizing social complications.

B. One encoding to rule them all?

As this wide range of partially contradictory requirements and demands suggests, it is unlikely that there is one encoding and marking scheme that meets all of them and therefore is suitable in all situations. P3F is thus based on a modular design with different encoding schemes. As Figure 4 illustrates, these schemes ideally come in a split form to reduce the computational impact. A lightweight detector finds candidates for code occurrences that are subsequently fed into a potentially computational more demanding decoder.

The P3F decoder additionally searches for individuals by using common face detection techniques and assigns all found policies to these persons. Finally, an effective P3F policy is deduced for each person by using the precedence rules described in Section V-A3.

C. Commonly used and available visual encoding schemes

1) **1D Barcodes:** Linear barcodes are the obvious choice for all articles of clothing with a stripe pattern. They are computationally easy to detect using frequency analysis. Clipping stability can be achieved by continuous repetition. However, common linear barcodes typically lack size invariance and the ability to repeat them continually due to quite zones and distinctive begin and end markers.
2) 2D Barcodes: Some barcode schemes (e.g., Microsoft Tag [21]) offer customizability up to the point where the data is completely disused by a picture (Figure 5). However, most barcodes require an easily spottable synchronization marker or a quite zone around them. Both make it difficult to hide them visually in a repeatable pattern. See Table IV for a comparison.

3) Augmented Reality Markers: These are similar to 2D barcodes but are highly distortion invariant and enable calculation of the relative distance and angle of the barcode surface to the camera. However, they are visually very prominent as they are optimized for real-time applications.

4) Symbols: Some symbols might be suitable for encoding. However, they are likely to be more intrusive and easily spotted by other people.

5) Watermarking techniques: Watermarking and stenography are often used in the context of digital rights management (DRM) and have been extensively researched over many years. However, their use for redigitized surfaces and especially for textiles is a new research area.

Watermarking can be roughly split [22] into techniques for detecting a known pattern (1-bit information) and those for reading arbitrary data. Furthermore, some algorithms (non-blind) need either the original image or original pattern whereas blind decoding algorithms can recover the watermark without an a priori knowledge. Most watermarking and steganographic techniques are designed for natural images and patterns and perform poorly for geometric shapes and drawn images.

Robustness against a print-scan attack [23] most closely resembles the normal use case in our application: a printed image or pattern with a watermark is redigitized with an optical sensor. Therefore, this use case is our main selection criteria.

Xin et al. [24] proposed watermarking for textiles by using the structure of the woven fabric. However, the watermark is only readable from very close distances.

Otori and Kuriyama [25] developed a very promising watermarking technique for natural textures that creates watermarks robust against analogue transportation paths (e.g., visual capture). Their implementation, however, is neither size and cropping invariant nor self synchronizing (i.e., it requires border lines).

Shirali-Shahreza et al. [23] put forward a very simple method for watermarking textiles based on collage steganography (i.e., object position in pictures). However, this method is non-blind and therefore not suitable for our application.

Zhu et al. [26] developed a watermarking system (originally for GIS applications) that is robust against resizing and cropping.

VIII. LIMITATIONS

The proposed framework does not try to mimic compulsory DRM systems. It is a best-effort system with an optional manual override. Furthermore, it does not allow the permissions to be changed after picture capture, unlike a central-database-based approach.

Although some visual marking systems are promising candidates as a foundation for further adaption, they all have their limits on strictly plain color cloths.

While the framework can handle multiple people in an image, partially hidden people could pose a challenge in matching policies to the individuals.

The proposed framework is aimed at the recognition of faces in pictures. While there are other forms of identification, such as identification of movement dynamics, face recognition is more commonly used and offers better distinction properties.

IX. FUTURE WORK AND EXTENSIONS

As none of the examined marking techniques fulfills all of the requirements we identified, there is a need for further research and development. Our next task is the implementation and evaluation of a prototype system based on our proposed framework.

Additionally, the development of a stable watermarking technique with a high data capacity would enable the use of a per-entity permission model based on private/public key

Fig. 5. Microsoft Tag offers great customizability for concealing the data pattern but still needs a distinctive feature (i.e., a bulky border) for synchronization [21]
cryptography. The individual would encode his/her publishing permission in the watermark and provide trusted publishers and/or friends with the decryption key. An automated P3F enforcer built into the social network or publishing system could then verify the permission by using a local permission-key database. To avoid introducing a new way of linkability, the watermarking algorithm should be parameterized with a part of the key. A publisher without the key would be unable to decode a deterministic bit pattern from the watermark.

X. Conclusion

The framework we have presented enables involuntarily or unintentionally photographed individuals to express their picture privacy policy in a machine readable format and (to some extent) automatically enforce it. An easily understandable flag system is used to restrict usage and linkability. This information is encoded in an unobtrusive way in wardrobe patterns and accessory designs. The encoding can be done using barcodes, watermarking, steganography, etc (Figure 6). While these forms of encoding are often used in context of DRM by corporations to protect their interests, they are used here for the benefit of individuals.

None of the techniques examined for building patterns around barcodes or the schemes examined for encoding information in a natural pattern meet the identified requirements and demands for a system that would fit into everyday life. This creates room for further research and development.

With regulatory help and/or enough public pressure, our proposed framework may one day be implemented in a system for widespread use.

ACKNOWLEDGMENT

This work was performed under the National Institute of Informatics international internship program.

REFERENCES


