Iris Recognition After Death

Mateusz Trokielewicz, Student Member, IEEE, Adam Czajka, Senior Member, IEEE, and Piotr Maciejewicz

Abstract—This paper presents a comprehensive study of post-mortem human iris recognition carried out for 1,200 near-infrared and 1,787 visible-light samples collected from 37 deceased individuals kept in the mortuary conditions. We used four independent iris recognition methods (three commercial and one academic) to analyze genuine and impostor comparison scores and check the dynamics of iris quality decay over a period of up to 814 hours after death. This study shows that post-mortem iris recognition may be close-to-perfect approximately 5 to 7 hours after death and occasionally is still viable even 21 days after death. These conclusions contradict the statements found in past literature that the iris is unusable as biometrics shortly after death, and show that the dynamics of post-mortem changes to the iris that are important for biometric identification are more moderate than previously hypothesized. The paper contains a thorough medical commentary that helps to understand which post-mortem metamorphoses of the eye may impact the performance of automatic iris recognition. An important finding is that false-match probability is higher when live iris images are compared with post-mortem samples than when only live samples are used in comparisons. This paper conforms to reproducible research and the database used in this study is made publicly available to facilitate research on post-mortem iris recognition. To our knowledge, this paper offers the most comprehensive evaluation of post-mortem iris recognition and the largest database of post-mortem iris images.

Index Terms—Iris recognition, post-mortem biometrics, forensics

I. INTRODUCTION

IDENTIFICATION of deceased individuals through their biometric traits has long been used for forensic purposes, exploiting characteristics such as fingerprints, DNA, or dental records to recognize victims of accidents, or natural disasters and crimes [1], [2]. Post-mortem iris recognition, however, has not received considerable attention, despite excellent performance of this method when applied to live eyes. Studying this unfamiliar area has at least two important goals:

a) To aid forensics: Can iris biometrics be a fast and accurate complement or alternative method to the existing approaches to post-mortem identification? If the answer is affirmative, it could be useful in cases when other methods cannot be applied, such as for victims of accidents with severed fingers or disfigured faces.

b) To improve security: Can dead iris be effectively used in presentation attack? Understanding the dynamics and reasons for post-mortem iris performance degradation allows to provide more precise answer to this question, and may help in development of countermeasures against forgeries with cadaver eyes.

To come up with as many answers as possible, this paper presents a comprehensive feasibility study of post-mortem iris recognition involving iris images acquired from 5 hours to almost 34 days after death in near-infrared (NIR) and visible-light (VIS). It is centered around the following six questions:

1) Is automatic iris recognition possible after death?
2) What are the dynamics of deterioration in iris recognition performance?
3) What type of images are the most favorable for post-mortem iris recognition?
4) What are the main reasons for errors when comparing post-mortem iris samples?
5) Which factors influence post-mortem iris recognition performance?
6) What are the false-match risks when post-mortem samples are compared against databases of live iris images?

To answer the above questions we acquired 1,200 NIR and 1,787 VIS images from 37 cadavers during multiple sessions organized from 5 to 814 hours after death. The bodies were kept in controlled mortuary conditions and stable temperature of 6° Celsius (42.8° Fahrenheit). Four independent iris recognition methods were used to show that automatic iris recognition stays occasionally viable even 21 days after death, and is close to perfect approximately 5 to 7 hours post-mortem. This allows to reject prior hypotheses that the iris cannot be used as biometrics after death [3], [4], [5], [6]. In this paper we also show that using the red channel of VIS post-mortem iris images can be considered as a good alternative to NIR samples. Images consisting of only red channel will be later referred to as ‘R images’ in the paper. We also show that the performance of cross-wavelength post-mortem iris matching (NIR vs R) is significantly worse than same-wavelength (NIR vs NIR and R vs R) matching. We analyze possible reasons for false match and false non-match instances, and by manual correction of the segmentation for the whole dataset we assess the impact of erroneous segmentation on the post-mortem iris recognition performance. The paper provides medical commentary on these post-mortem metamorphoses observed in the eye that degrade the recognition reliability the most. We discuss briefly relation of gender, age, and cause of death with post-mortem iris recognition. To our knowledge, this paper comprises the most extensive and comprehensive study regarding post-mortem iris recognition, and offers the largest dataset of iris images collected from deceased subjects.

The paper is organized as follows. Section II provides an overview of publications and claims related to post-mortem iris recognition. Section III describes the database used in...
In our previous work we showed that despite popular claims, the iris can still be used as a biometric identifier for post-mortem images, as well as corneal opacity, which fades out of the pupillary and limbic boundaries found in the eye while it is still a part of the cadaver. Ross observed that irises typically became useless from the recognition viewpoint only a few days after exposition to outdoor conditions, and if the bodies are kept outside for 14 days the correct verification decreases to only 0.6%. The real-life chance of recognizing an iris is estimated by the authors to be even less than 0.1%. The most recent paper in this field by Sauerwein et al. showed that irises stay readable for up to 34 days after death, when cadavers were kept in outdoor conditions during winter. The readability was assessed by human experts acquiring the samples and no iris recognition algorithms were used in this study, however it suggests that winter conditions increase the chances to see an iris even in a cadaver kept outside for a longer time.

III. DATABASE OF POST-MORTEM IRIS IMAGES

A. Data Collection

A crucial part of this study was to create a new database of iris images, which would represent eye regions of recently deceased persons. We had a rare opportunity to collect iris scans from hospital mortuary subjects. The following section briefly characterizes the acquisition methodology and timeline of acquisition sessions.

1) Equipment: Two different sensors were used for image acquisition: a commercial iris sensor operating in NIR light IrisShield M2120U, and a consumer-grade color camera Olympus TG-3. Color images were collected simultaneously with NIR ones and each subject and each acquisition session are represented by at least one image of each type. The IrisShield sensor is equipped with a near-infrared illuminant, whose irradiance falls into the 710-870 nm band, with a peak at 810 nm.

2) Environmental Conditions: All acquisition sessions were conducted in the hospital mortuary. The temperature in the mortuary room was approximately 6° Celsius (42.8° Fahrenheit). Other conditions, such as air pressure and humidity were unknown, yet stable. The environmental conditions, in which the cadavers were kept prior to entering the cold storage are unknown.

3) Acquisition Timeframe: From 1 to 13 acquisition sessions could be organized for a given subject in this study. In each session at least one NIR and one VIS image were
acquired. Due to ethical concerns, no ante-mortem samples could be collected, hence the first session for each subject was always organized as soon after death as possible, typically 5 to 7 hours. The following sessions were organized based on the availability of deceased persons, who were subject to medical or police investigations, and were retained in the mortuary during varying time slots. The overview of acquisition sessions for all subjects is shown in Fig. 1. For three subjects, namely 20, 33, and 34, only a single acquisition session was possible.

![Graph showing hours post-mortem for each acquisition session](image)

**Fig. 1:** Hours post-mortem for each acquisition session plotted independently for each deceased subject.

4) Within-Session Acquisition Protocol: When collecting images within a single acquisition session, all samples can be considered separate presentations as recommended by the ISO/IEC 19795-2, i.e., after taking a photograph, the camera was moved away from the subject and then positioned for the next acquisition.

B. Quality of the Samples

For each image collected for this database, we have implemented the calculation of two quality metrics suggested in the ISO/IEC standard on iris image quality [19], namely: grayscale utilization and sharpness, as these are the only metrics from the current standard that do not require ground truth segmentation information, but rather require a raw iris image. Formulae for these calculations, statistical results, and quality scores for selected images are included in the Supplementary Materials.

C. Statistics

The dataset is a significant extension over the corpus used in our previous studies, comprising 1,200 NIR images, accompanied by 1,787 color images. These images represent eye regions of 37 different subjects (73 different irises, since only one eye was imaged for one cadaver). Age of the deceased ranged from 19 to 75 years old. 5 subjects were female and 32 were male. Causes of death included heart failure (18 subjects), car crash (7), suicide by hanging (7), murder (1), poisoning (2), and head trauma (2). The eye colors were blue-gray/light green (29 cadavers), light brown/hazel (5) and dark brown (3). Detailed description for each subject can be found in the metadata accompanying the released database.

D. Database Access

To conform with the reproducibility guidelines and facilitate research in this area, the database used in this study is made available along with the paper. A copy can be requested at: [http://zbum.im.pw.edu.pl/EN → Research → Databases → Warsaw-BioBase-Post-Mortem-Iris v2.0](http://zbum.im.pw.edu.pl/EN → Research → Databases → Warsaw-BioBase-Post-Mortem-Iris v2.0).

IV. MEDICAL BACKGROUND

A. Post-Mortem Changes To The Eye

1) General overview: Initially, the post-mortem decomposition of human organs may not be visible to the naked eye, since these processes start at the cellular level and then slowly progress to the macroscopic level. Early changes include algor mortis (body cooling), rigor mortis (desiccation with stiffening of the body), pallor mortis (paleness) and livor mortis (lividity), while late ones comprise of progressing decomposition caused by autolysis and putrefaction. Autolysis is a cellular self-destruction process caused by hydrolytic enzymes that were originally contained within cells. Putrefaction is a degradation of tissue caused by microorganism (e.g., bacterial) activity, and is visible macroscopically as discoloration or bloating of the skin.

2) The cornea: The most prominent metamorphoses observed in the eyes after death, and possibly the most troubling for iris recognition, are the changes to the cornea. A live cornea is a clear, transparent, dome-shaped structure in front of the eyeball. It is responsible for about 2/3 of the total eye optical power because of its curvature and the resulting refractive index. The cornea must remain transparent to refract light properly, but also to allow good quality iris image capturing. Its transparency is maintained by a controlled hydration with the tear film, produced by lacrimal glands and distributed by eyelids. As secretion stops, anoxia, dehydration and acidosis lead to progressing autolysis of the cells. Corneal thickness decreases immediately after death and increases thereafter. This results in opacification that increases with time. Upon death the cornea slowly becomes hazy. The change in corneal opacity is believed to be secondary to the change in hydration and architectural destruction of the collagen fiber network, functional alteration of corneal endothelium, deregulation of proteoglycan hydration and ion concentration in corneal stroma. It was confirmed that temperature has significant influence on protein degradation. Another effect associated with these mechanisms is the wrinkling of the corneal surface, manifesting itself with difficulties to obtain a good visibility of the underlying iris pattern. The progression of these effects is influenced by multiple factors, such as closure of the eyelids,
environment humidity, temperature, and air movement. It is also dependent on the age and general medical condition of a deceased person. Due to reduced intraocular pressure, we can also notice central depression of the globe, flaccidity of the eyeball, and loss in its firmness [17].

3) The iris: There are no evident changes to the iris surface observed after death. After demise, pupils are usually mid-dilated (a.k.a. ‘cadaveric position’), and in some cases they can be slightly dilated, because of the relaxation of the iris muscles and later they can become slightly constricted with the onset of rigor mortis of the constrictor muscles. In other cases, we may observe initial myosis within the first few hours after death with strong variations between individual cases. If rigor mortis affects ciliary muscles of two irises unequally, pupils in both eyes may have different apertures. Sometimes, if different segments of the same iris are unequally affected then the pupil may be irregularly oval or have an eccentric position. Shape and size of the pupils can also depend on the medical history of the subject, including treatment with drugs and eye surgeries.

4) Muscles of the iris: Death was once defined as the cessation of heartbeat and breathing, but with the development of cardiopulmonary resuscitation these can be restarted in some cases. Thus, we now typically rely upon the concept of brain death to define whether a person is clinically dead. Supravitality – sensitivity to excitation – relates to survival rates of tissue after complete irreversible ischemia (restriction of blood flow). The supravital reaction is the response of muscles to stimulation in the early period after death, as some cells do not die immediately after the brain death. Within the first couple of hours we can notice a decreasing pupillary reaction for pupillomotoric drugs, for example to pilocarpine and atropine [18]. Iris tissue response to myotic (pupil-constricting) and mydriatics (pupil-dilating) agents can be observed.

5) Other aspects: There are some other changes that the eye undergoes after death. The loss of intraocular lens transparency with time is due to the metabolic processes that take place between the lens and the aqueous and vitreous humor, and the aggregation of crystalline proteins in the fibers of the lens nucleus. It has been hypothesized that lens proteins aggregate to large particles that scatter light, causing lens opacity [17]. After death we may observe a black spot in the sclera, referred to as ‘tache noire’, caused by desiccation of the sclera with open eyelids, usually symmetrical corresponding to the position of the eyelids. Also, the vitreous humor – gelatinous substance contained in the posterior chamber of the eye, keeping the retina in place and maintaining the spherical shape of the eyeball – tends to liquefy, and later to dry, starting the process of eyeball collapse [19].

B. Visual Inspection of Post-Mortem Changes

We have taken the effort to carefully examine the samples throughout the time period since death for all subjects, and confront the observed changes with medical knowledge. This yielded a qualitative evaluation of post-mortem changes to the iris reported in this Section. Having both NIR and VIS images is crucial for such assessment, as these two types of illumination often reveal different appearance of the iris when changes to the cornea and the anterior chamber are present. This is shown in Fig. 3, where visible-light samples are compared against near-infrared samples for the same eye. Such differences are also reported on in the works of Aslam et al. [20] and Trokielewicz et al. [21] related to the disease influence on iris recognition performance. Both studies show that the NIR illumination typically used in iris recognition cameras is capable of alleviating corneal opacification effects to some extent.

A summary of example post-mortem changes that appear in the eye is presented in Fig. 2 together with a timeframe for a selected subject. It must be noted, however, that the dynamics of these changes are heavily subject-dependent and can happen with different rapidity, intensity and prevalence on the appearance of iris tissue.

First, a corneal opacification progresses with time since death, and it becomes visible after a few days post-mortem (e.g., 95 hours, or 4 days, after death, as depicted in Fig. 2). Second, a wrinkling of the corneal surface is expected to appear (e.g., 359 hours, or 15 days, as shown in Fig. 2). At this point, a strong influence on the automatic image segmentation procedures can be anticipated, as the iris tissue becomes less visible and additional patterns and light reflections emerge. Third, a loss of intraocular pressure in the eyeball due to post-mortem biochemical changes can be observed (e.g., 574 hours, or 24 days, as illustrated Fig. 2), causing the eye to slowly collapse into the eye socket. At this point in time, iris recognition methods are expected to seldom work, as the iris pattern is severely obstructed and thus challenging for iris image segmentation. Finally, after about a month, the eyeball was observed to dry out completely, leaving no traces of a healthy iris structure.

Contrary to initial predictions, we did not come across any sample that would be affected by tache noire. Also, the severe corneal opacification was visible in original VIS samples only, while NIR and R images worked in favor of exposing post-mortem iris texture better than original VIS samples, as depicted in Fig. 2.

V. EXPERIMENTAL STUDY

A. Iris Recognition Methods

For a comprehensive analysis of how iris recognition can perform when used with post-mortem samples, we have employed four independent iris recognition methods. Three of them are commercially available products, and one is an open source solution.

1) VeriEye: This commercial product is offered by Neurotechnology in the form of the Software Development Kit (SDK) [22]. The manufacturer does not divulge algorithm details, apart from the claim that off-axis iris localization is employed with the use of active shape modeling. The algorithm has been evaluated by NIST in their ICE 2005 [23] and IREX [24] projects. VeriEye is the only algorithm in this study which returns a similarity score between two iris images, rather than a difference score. Hence, the higher the score, the
Fig. 2: Example measurements for a single subject throughout the period of time post-mortem. Symptoms that are expected to appear as time since death elapses are denoted above the top row. OSIRIS segmentation results are also shown for selected images, presenting the degradation of iris image segmentation quality with passing time. Red-colored portions of the image are denoted by the algorithm as not representing the iris and therefore do not participate in comparison.

Fig. 3: From left to right: visible-light (VIS) (top row), near-infrared (NIR) (middle row), and red channel (R) (bottom row) images obtained in the first three sessions: 5, 16, and 27 hours after death. Progressing corneal opacity can be spotted in visible-light images, but NIR imaging seems to be more insensitive to corneal haze.

better the match between samples. Scores between 0 and 40 denote different-eye (impostor) pairs, while scores above 40 are expected for same-eye (genuine) pairs.

2) IriCore: Similarly to the VeriEye method, the IriCore matcher is offered commercially as the SDK [25]. IriTech Inc. does not disclose any details on the underlying algorithm. The software is claimed by the manufacturer to conform with the ISO/IEC 19794-6 standard [26] and has been shortlisted by NIST in 2005 [23] as one of the best iris recognition solutions. IriCore returns dissimilarity score from 0.0 to 2.0, with same-eye (genuine) scores expected to fall in between 0 and 1.1, and different-eye (impostor) scores between 1.1 and 2.0.

3) MIRLIN: (Monro Iris Recognition Library) is a third method offered on the market in the form of an SDK by FotoNation Ltd (formerly Smart Sensors Ltd) [27]. The underlying algorithm employs discrete cosine transform (DCT) calculated for overlapping iris image patches to deliver binary iris features [28]. Similarly to Daugman’s original method, the resulting iris codes are compared using exclusive or (XOR) operation and normalized by a number of valid bits (corresponding to iris portions that are not occluded), yielding a fractional Hamming distance. Comparing two images of the same eye should result in a score close to zero, while the distance between images of two different irises is expected to oscillate around 0.5.

4) OSIRIS: Open Source for IRIS, an academic solution developed within the BioSecure EU project [29], has been open-sourced by its authors. Its principles follow the original works of John Daugman, with iris image segmentation and subsequent normalization to dimensionless polar coordinate system. A binary iris code is calculated using phase quantization of the Gabor filtering outcomes. Similarly to the MIRLIN method, the fractional Hamming distance between the codes is used as a comparison score. Values close to zero are expected for two same-eye images, while scores oscillating around 0.5 should be produced when two different-eye images are compared. However, due to compensation of the eyeball rotation, different-eye score distributions will more likely be skewed toward 0.4 – 0.45 range. We introduced a modification of the original OSIRIS method to include score normalization as proposed by Daugman [30]:
\[ HD_{\text{norm}} = 0.5 - (0.5 - HD_{\text{raw}}) \sqrt{\frac{n}{N}} \]

This transforms the samples of scores obtained when comparing different eyes into samples drawn from the same binomial distribution, as opposed to drawing sample scores from different binomial distributions with \( \sigma \) dependent on the number of bits \( n \) that were available for comparison (commonly unmasked bits). \( N \) is the 'typical number of bits compared (unmasked) between two different irises,' being said to equal 911 or 960, depending on the data. The \( N \) parameter is estimated for a particular database of iris images. Since post-mortem iris samples are different from images of live iris images, and the OSIRIS does not necessarily use identical Gabor wavelets as used in [29], we estimated iris images, and the OSIRIS does not necessarily use identical post-mortem iris samples are different from images of live iris images. Since all of the iris images involved in this study, including both NIR and visible-light samples, as depicted in Fig. 4.

B. Database of Iris Images

1) Near-infrared and visible-light samples: During data acquisition, we had the opportunity to collect images using two types of cameras: one producing near-infrared images of VGA size (640 \times 480 pixels), and the second producing color photographs of high resolution. Both near-infrared and visible-light images were used in visual inspection presented in Sec. IV-B. Using only red channel of color iris samples acquired in visible-light has been found to offer high recognition accuracy even when being matched with near-infrared samples. Thus, in this study, two types of samples are used: a) original near-infrared and compliant to ISO/IEC 19794-6 and ISO/IEC 29794-6 standards, and b) red channel of visible-light images manually center-cropped to conform the VGA image type, as defined in ISO/IEC 19794-6. Hence, the resolution of all images is 640 \times 480 pixels. This cropping of visible-light sample additionally protects the identity of donors, as original high-resolution visible-light images contained significant portions of face region. An example pair of near-infrared sample and the cropped red-channel image of the same eye is shown in Fig. 4.

2) Manually-annotated ground-truth iris masks: The erroneous execution of the segmentation stage is usually a main cause of drops in iris recognition performance, when samples presented to the algorithms are of challenging nature. To test whether this is also the case for post-mortem data, we have taken the effort to prepare manually annotated iris masks for all of the iris images involved in this study, including both NIR and visible light samples, as depicted in Fig. 4.

Fig. 4: Example NIR and R images of the same eye (65 hours after death) and the corresponding manually annotated masks.

3) Dataset of live iris images: For the purpose of assessing false-match risks in scenarios when post-mortem samples are expected to be compared in an open-set scenario with the existing datasets of live iris images, we have collected a complementary dataset of iris images from living 74 subjects. To minimize the bias in the data, we have used the same sensor (IriShield MK 2120U) as used for the collection of post-mortem data. Two subject-disjoint subsets gathering data from 37 subjects each were created, containing 557 and 611 images, respectively.

C. Types of Analyses

Due to difficult data acquisition resulting in rather sparse and irregular image capture moments, the experiments conducted in this study are carried out three-fold: (i) a short-term analysis that is based on samples collected in the first acquisition session for each subject, (ii) a long-term analysis that employs the entire data representing a maximum period of 814 hours after death, and (iii) an open-set analysis to assess the false match probability when post-mortem samples are compared to live samples.

D. Failed Comparisons Statistics

Iris identification methods are often equipped with image quality control mechanisms that prevent comparing samples with unacceptable properties. That is, if at least one of two samples being compared presents a less-than-accepted quality, the comparison fails and no valid comparison score is returned. Iris image quality seems to be defined differently in the implementations used in this study, since the numbers of failed comparisons is non-uniform across matchers. It is also difficult to hypothesize on what the exact reasons of these failures are, since out of four recognition algorithms, only one of them, IriCore, is explicitly said by the manufacturer to be compliant with the ISO/IEC 29794-6 standard on iris image quality. As for the open-source OSIRIS matcher, neither ISO/IEC metrics nor any other methods for discarding bad samples are implemented. For this method, samples denoted as those that failed to produce a template did so because of errors returned by low-level OpenCV routines used in OSIRIS, which may be a result of low quality of the processed samples. For the remaining two matchers, VeriEye and MIRLIN, there is no information how the iris image quality is assessed in these commercial products.

Table [I] summarizes numbers of all comparisons in short-term and long-term analyses along with numbers of the failed comparisons. For instance, the most restrictive method, MIRLIN, fails to compute almost 44% of comparison scores when comparing NIR samples, and fails to compute over 40% of comparison scores when comparing R images. In turn, the IriCore software was able to compute all short-term comparison scores and failed to compute only 0.11% of long-term comparison scores for NIR images. Analyses presented in the remainder of this Section are based only on comparison scores that were correctly calculated. We should expect that methods controlling the iris image quality more restrictively should present a higher recognition accuracy (e.g., lower EER) than the methods implementing weaker quality discrimination.
TABLE I: Number of failed comparisons (genuine and impostor) and total number of comparisons calculated for all variants of analyses done in this study. Numbers are presented separately for each iris recognition method and suggest that iris image quality control mechanisms differ significantly among methods when post-mortem samples are used.

| Method      | NIR images | R images | NIR images | R images |
|-------------|------------|----------|------------|----------|
| Short-term analysis | 39,621     | 76,245   | 245,904    | 447,882  |
| Long-term analysis   | 0          | 0        | 282 (0.11%)| 0        |

E. Baseline Post-Mortem Performance: Short-Term Analysis

Fig. 5 illustrates Receiver Operating Characteristic (ROC) curves obtained in short-term analysis, when images were acquired within the first acquisition session (5-7 hours post-mortem). Iris recognition performs reasonably well on average in short-term post-mortem horizon, what conforms with earlier works on a smaller data sample [10], [11]. The best performing method, IriCore, achieved $EER = 4\%$ for NIR images, and $EER = 2\%$ for R images.

Interestingly, having taken iris images in visible light with a high-resolution camera, and using their red channel, allows for a better performance of all of the involved iris recognition methods when compared to the performance achieved for VGA (i.e., low-resolution) NIR images. We can observe even six-fold gains in recognition accuracy for the VeriEye matcher, measured $EER$-wise ($2\%$ error with R images versus $13\%$ error for NIR images). One of the contributing factors for such a favorable performance of R images is likely a dominant proportion of lightly-colored eyes in the dataset, cf. Sec. III-C, for which iris recognition is known to work well even without a specialized NIR camera, provided that the images are of good quality [31]. For completeness, we also include an analysis of cross-spectral matching scenarios, where R images are matched against NIR images (dashed black lines in Fig. 5), however, the results of such matching are discouraging enough not to be investigated any further, and thus are not considered in the long-term analysis. Apparently, for post-mortem images, the differences in sample presentation under different illumination are large enough to make such a scenario unusable.

Note that all algorithms, except for the OSIRIS, did not reject any sample due to low quality (cf. ‘Short-term analysis’ columns in Tab. I), which makes these methods perfectly viable for samples collected a few hours after death. The worst method in short-term analysis, MIRLIN, presents equal error rate of $7\%$ for R images and $20\%$ for NIR images. However, this could possibly be compensated by more restrictive quality control since MIRLIN, as IriCore, did not reject any sample prior to matching. Even so, the worst result ($EER=20\%$) is still far better than $50\%$ expected for a random classifier, indicating a possibility to correctly recognize a large subset of cadaver irises.

Finally, when examining results obtained when matching manually segmented samples using the OSIRIS matcher (bottom-right plot in Fig. 5), we may draw a conclusion that most of the recognition errors can be associated with erroneous execution of the automatic segmentation stage, as this method, when processing manually segmented images, presents $EER = 1\%$ for R images and $2\%$ for NIR images. This shows that when post-mortem samples are carefully segmented to represent the correct iris region, iris recognition is a viable method for identification.

F. Biometric Capability Decay Study: Long-Term Analysis

Samples acquired in the second and subsequent sessions are sparsely distributed in time and across the subjects. Thus, Figs. 6 through 14 present genuine and impostor scores calculated by all the methods, between session 1 images (5-7 hours after death) and all samples acquired in the following sessions, together with close-up analysis of example false matches and false non-matches. With a difficulty to obtain ante-mortem and then post-mortem iris scans from the same individuals, and after visual inspection of samples (cf. Sec. IV-B), we assume that iris scans obtained shortly (i.e., a few hours) after death would not differ much from those obtained ante-mortem. Thus, images acquired in the first session serve as gallery samples in all subsequent analyses.

In this section we also visually investigate selected image pairs that generated false matches and false non-matches (assuming default acceptance thresholds). Each method, except for OSIRIS with Daugman’s score normalization, generated false matches and all matchers generated false non-matches. Thus, we selected the worst pair of images in each case, i.e., the most similar images of different eyes and the most distinct images of the same eye, both in terms of the comparison score. We were also able to read the segmentation results in two methods (MIRLIN and OSIRIS), which helped to find a reason behind a given error.

So, how long after death can the iris still offer enough features to generate a correct match?

For the OSIRIS method, when employing automatic image segmentation, we can expect correct matches for samples obtained up to 150 hours post-mortem for NIR images, and 263 hours for R images, Fig. 6. However, when the manual corrections to the segmentation stage are introduced, Fig. 8 the recognition horizon for NIR samples extends to 263 hours, while for R samples it declines to only 215 hours post-mortem, which may indicate that some of the correct matches
obtained for R images were a result of badly segmented irises, and not of a genuine match of iris codes. Plots of impostor comparisons shown in Figs. [6] and [8] suggest that there is no clear trend in comparison scores when time after death increases, also in the case when manual segmentation was applied.

Figure [7] presents the worst results obtained from OSIRIS prior to applying score normalization. Incorrect segmentation explains the reason for a false non-match obtained for the left pair of samples. The right pair and the segmentation results show a very interesting case of a false match. Even with incorrect segmentation, as observed in this case, we still have some ‘non-occluded’ image areas that are significantly different (part of the skin on the first image and part of the wrinkled iris/cornea on the second image). So why the false match is observed? The reason for that may be a very small number of bits being compared due to application of occlusion mask during calculation of the Hamming distance. Note that the occlusions found by the OSIRIS in this image pair are almost mutually exclusive: ‘non-occluded’ part is mostly located on the left part of the hypothetical iris on the first image, while ‘non-occluded’ part of the iris shown on the second image is located mostly on the right. This false match does not happen after applying score normalization that penalizes low numbers of mutually un-occluded bits in OSIRIS codes.

As for the remaining matchers, IriCore and MIRLIN were able to deliver correct matches for samples acquired even 503 hours post-mortem. VeriEye occasionally recognizes samples acquired up to 260 hours after demise, at the assumed acceptance thresholds. IriCore was again a method that rejected a very small number of comparisons (0.11% of NIR images and 0% of R images) when compared to other methods (cf. ‘Long-term analysis’ column in Tab. [1]). This means that in favorable conditions (IriCore software and IrisShield sensor come from the same manufacturer) iris recognition may still be possible almost 21 days after death. Unfortunately, as this method does not allow for generating segmentation results, we are not able to investigate whether these correct matches are indeed a result of genuinely matching the corresponding iris features, or of an incorrect segmentation causing the similar portions of the image to be matched.

Figure [10] presents iris image pairs yielding false non-match (left pair) and false match (right pair) when the MIRLIN method is used. In this case we can also observe the segmentation results, which give a clear explanation of the observed errors in both cases. A false non-match is caused simply by comparing non-matching iris areas due to incorrect iris localization. A false-match is probably due to comparing sclera that is very similar in both samples, rather than actual iris texture.

Figure [12] presents the worst cases leading to false non-match (left pair) and false match (right pair) when the IriCore method is used. The reason for a false non-match is the collapse of the eyeball and severe cornea wrinkling observed in the second sample (acquired 622 hours post-mortem) of the left pair. The right pair of images, however, does not provide any obvious clue for the observed false match.

Figure [14] shows the worst pairs of images from the VeriEye method point of view. The possible cause of a false non-match
Fig. 6: Long-term analysis for all subjects for the OSIRIS matcher using automatic iris segmentation, for both NIR and R samples. Green dots represent correct behavior, i.e., a match for genuine pair, and a non-match for impostor pair, while red dots correspond to incorrect behavior, i.e., a false non-match for a genuine pair, and a false match for impostor pair between samples acquired in the first session (after 5-7 hours) and samples acquired in the following sessions.

Fig. 7: Example image pairs generating a false non-match and a false match for the OSIRIS matcher prior to score normalization. The comparison scores are shown in brackets. Left image of each pair is the gallery sample. We assume that OSIRIS returns a match when the comparison score is below 0.32. After score normalization, the false match shown above is rectified.

(left pair) is a compensation of lower intraocular pressure by manually pressing the eyeball (a finger of a personnel pressing the eyeball is visible in the second image). This made the cornea less wrinkled (when compared to the first image of the left pair), but simultaneously changed the visible texture, ending up with creation of different iris features for those samples. However, the right pair of images again does not provide a clear explanation for a false-match.

G. Assessment of False Match Risk When Comparing Post-mortem Samples Against Live-iris Samples

All the above experiments were performed for a closed set of deceased subjects. A question arises: what happens if post-mortem samples are compared in an open-set scenario with live irises of other subjects? To answer this, we used the IriShield MK 2120U sensor again to collect an additional database of iris images from 74 living persons, which was split into subject-disjoint live-test and live-reference sets, both comprising data from 37 subjects and approximately 600 images. The most accurate (in this study) IriCore method was used to generate and compare two impostor score distributions: (a) scores obtained when matching post-mortem iris images (from 37 cadavers, as introduced in this paper) against live-test iris images, and (b) scores obtained when matching the same live-test images with samples from the subject-disjoint live-reference set. The post-mortem set was created by randomly choosing images in a way that from each acquisition session half of the samples are selected. This ends up with balanced post-mortem and live iris sets.

Fig. 15 (left) presents the False Match Rate calculated for different acceptance thresholds and for two above scenarios (a) and (b). For the threshold equal to 1.1, as recommended
by the manufacturer, we observed small FMR < 0.1% in both scenarios. However, when the threshold is relaxed, the chances of getting a false match when live iris images are compared to post-mortem images are higher than in scenario when only live iris images are compared. At the same time, it is hard to see any clear dependency between false match rate and time after death when the iris was photographed, as depicted in Fig. 15 (right).

VI. CONCLUSIONS

This paper delivers the most comprehensive, known to us, analysis of post-mortem iris recognition. The concluding remarks are presented in the next paragraphs in a form of answers to the questions that stimulated this research and were posed in the introductory part. We are also aware of certain limitations of this study, and these are presented at the end of this Section.

A. Question 1: Is automatic iris recognition possible after death?

Post-mortem iris recognition is possible. In favorable conditions, when the sensor and the processing methods come from the same manufacturer, we observed EER=2% when comparing images acquired 5-7 hours post-mortem, collected in visible spectrum with a high-resolution camera, and then converted to grayscale using the R channel. The recognition accuracy can, however, be further improved by employing a human expert to manually annotate the correct iris region in the image, driving the EER down to as low as 1%, employing the OSIRIS matcher.

B. Question 2: What are the dynamics of deterioration in iris recognition performance?

Due to inevitable decomposition of human body, recognition accuracy becomes progressively worse. Despite a longer analysis horizon (814 hours) when compared to prior work in this field [7], [10], [11], [33], we did not find any iris image taken later than 503 hours after death that would result in a correct match in all four iris recognition methods used in this work. On the other hand, 503 hours (almost 21 days) is a horizon giving an ample amount of time to make post-mortem forensic analysis and suggests that the iris may deliver actual biometric features for a longer period than initially believed. One should also note that four methods used in this study presented highly heterogenous performance on the same set of post-mortem samples. For instance, IriCore failed to compute only 0.11% comparison scores, while MIRLIN failed to compute 43.51% comparison scores for NIR images. This may suggest that the algorithms implement significantly different mechanisms to control the quality of iris image pair used to compute the score.
C. Question 3: What type of images are the most favorable for post-mortem iris recognition?

In short-term analysis, when matching samples are obtained shortly after death, we have shown that employing high-resolution R images offers much higher recognition rates when compared to those obtained with NIR images (2% EER for R images versus 4% for NIR images, obtained for the IriCore matcher). Both types of images offer close-to-perfect performance when the image segmentation stage is executed manually, with small difference in favor of the R images (1% ERR versus 2% for NIR images, obtained for the OSIRIS matcher). This may prove important for forensic applications, which usually involve a human expert, who could then perform the necessary segmentation stage. In addition, we have shown that, probably due to significant differences in the appearance of iris features in post-mortem samples under different wavelengths, a cross-spectral scenario cannot be recommended, with ERRs not dropping below 15%, even with manual image segmentation in place.

In the long-term analysis, there is no single conclusion on whether NIR or R images are better for post-mortem iris recognition. For the OSIRIS method with manually corrected image segmentation, NIR images seem to offer better chance of getting a correct match. For MIRLIN, R images are generating far more false matches than NIR images, while for IriCore it is the opposite, with NIR images causing more false matches. However, NIR images also allow for correct matches during a longer period post-mortem, compared to R images. VeriEye, on the other hand, seems to work better in general, when R images are used, namely offering less false-matches and more correct matches. Therefore, the matching performance in regard to the image type is heavily matcher-dependent.

D. Question 4: What are the main reasons for errors when comparing post-mortem iris samples?

After visual inspection of the automatic segmentation results for falsely matched and non-matched samples, one of the most obvious reasons for failures is incorrect localization of iris texture due to post-mortem decomposition processes. However, most of these failed examples still show the iris...
E. Question 5: Which factors influence post-mortem iris recognition performance?

The data collected for the purpose of this study represents images acquired from subjects who passed away due to various reasons and of different age. Also, samples for both female and male subjects were collected. This gives a rare opportunity to examine whether cause of death, age and gender can give a priori insights on the expected performance of post-mortem iris recognition. Significantly worse comparison scores were observed for subjects who were either poisoned or murdered. Significant differences were observed neither between males and females, nor among groups of comparison scores sorted by

the age in the moment of death. However, due to a relatively small number of available genuine comparison scores, these results should be considered as qualitative assessment, rather than formal statistical analysis.

F. Question 6: What are the false-match risks when post-mortem samples are compared against databases of live iris images?

As shown in Sec. V-C, false-match probability may be higher when live iris images are compared with post-mortem samples than when only live samples are used in comparisons. This translates to a higher chance of observing a false match when post-mortem probe sample is compared to a gallery of live iris images, and it calls for post-mortem-specific iris matching strategies to address a possibly of higher false match probabilities in post-mortem iris recognition.

G. Limitations of this research

Although this study brings a few groundbreaking deliveries, and the largest database of post-mortem iris images, one should proceed with the conclusions cautiously and consider this work also as an important call for intensified research in post-mortem iris recognition. The most important limitation that we are aware of is a relatively small dataset (37 deceased subjects). This data were collected in a very difficult environment of the hospital mortuary and acquisition moments could
Fig. 13: Same as in Fig. 6, but for the VeriEye matcher.

(a) False non-match (0)  
(b) False match (46)

Fig. 14: Same as in Fig. 7, except for the VeriEye matcher. We assume that VeriEye returns a match when the comparison score is above 40.

Fig. 15: **Left:** FMR as a function of IriCore acceptance threshold for comparisons between live irises and live vs post-mortem irises. **Right:** FMR as a function of post-mortem acquisition time calculated for the IriCore acceptance threshold 1.2, when live irises are compared with post-mortem irises.

not interfere with various examinations, including criminal proceedings. Time spent in the mortuary was beyond our control, hence irregular acquisition sessions. The lack of ante-mortem samples limits the calculation of the reference templates to the earliest post-mortem images. However, visual inspection suggests that the first-session samples do not exhibit any visible deterioration when compared to living irises. Also, it would be valuable to repeat our analyses for samples collected under varying ambient conditions, as done by Bolme et al. [13]. However, we are not aware of any database of post-mortem samples different than offered with this paper, and collected under varying environmental conditions, or for more cadaver subjects that would be available to the researchers at the moment of preparation of this paper.

**ACKNOWLEDGMENT**

This study had an institutional review board clearance and the ethical principles of the Helsinki Declaration were carefully followed by the authors. All legal requirements necessary for making the dataset publicly available to researchers have been met. The authors thank Ms Katarzyna Roszczewska and Ms Ewelina Bartuzi for their help with preparation of the ground-truth occlusion masks. The authors also thank Dr. Kevin W. Bowyer who provided valuable comments on an early draft of this paper.
