**Recenzja rozprawy doktorskiej**

14 lutego 2022

**Kandydat:**

mgr inż. Mohammadreza Azimi

Wydział Elektroniki i Technik Informacyjnych

Politechnika Warszawska

ul. Nowowiejska 15/19

00-665 Warszawa

**Tytuł rozprawy:**

Investigation into the Reliability of Contactless Biometric Systems

**Recenzent:**

dr hab. inż. Adam Czajka

Department of Computer Science and Engineering

University of Notre Dame

384 Fitzpatrick Hall of Engineering

Notre Dame, IN 46556, USA

aczajka@nd.edu

Niniejsza recenzja została sporządzona wg zaleceń przekazanych przez Politechnikę Warszawską. W szczególności ocenia ona spełnienie przez rozprawę doktorską warunków określonych w art. 187 ust. 1-2 Ustawy z dnia 20 lipca 2018 r. Wg mojej wiedzy kandydat nie posługuje się językiem polskim w stopniu pozwalającym na swobodną dyskusję rozprawy, więc recenzja sporządzona została w języku angielskim.

**What scientific/research issue is considered in this dissertation (the aim and thesis of the dissertation) and was it formulated by the author in a comprehensible way?** (PL: Jakie zagadnienie naukowe/badawcze jest rozpatrywane w pracy (cel i teza rozprawy) i czy zostało ono dostatecznie jasno sformułowane przez autora?)

The main goal of this thesis is to investigate the impact of selected social factors on biometric recognition reliability. The main thesis states that “the reliability of biometrics can be significantly affected by social factors, but that this can be mitigated by proper data analysis techniques.” To support this claim, the author selected three biometric modalities (iris, speaker, and face recognition), formulated, and then verified four supporting theses: (S1) iris recognition is less effective for subjects suffering from type II diabetes; (S2) time of day when the voice is recorded affects text-independent speaker recognition systems; and (S3) face makeup impacts face recognition reliability. The main and supporting theses are formulated in a comprehensible way.

An immediate remark is that not all factors deteriorating biometric performance, especially those having their roots in biology, “can be mitigated by proper data analysis techniques.” I would rather agree that such mitigation is possible for selected situations and only to some degree. And this is what this thesis actually verifies.

Also, I am not sure if diseases should be considered as “social factors”. Instead, social factors, such as education, income, working and life conditions, diet habits, etc. may facilitate evoking certain diseases, which in consequence may impact biometric recognition. Thus, this thesis rather investigates the impact of selected nonideal, or adverse conditions (having various sources, not necessarily being social factors), in which biometric systems are placed, on recognition reliability.

**Does the dissertation include appropriate analyzes of past work, including world literature, state of the art, and industrial applications?** (PL: Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł, w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle?)

The list of cited literature has 188 entries, and the selection of sources in a satisfactory way represents the state of the art. The author could perhaps look at the most recent works related to correlations between iris recognition and gender. He says that *very little work has been focused on using iris texture patterns to determine ethnicity or gender* (p. 21), but there is a series of papers from the University of Notre Dame on possible use of iris texture and ocular regions for gender prediction [1,2]. These newer works would probably help in avoiding incorrect statements, such as *the iris can be used to indicate gender* (p. 61). This is incorrect if the “iris” means the non-occluded iris texture within the iris annulus, and the reason why it may be incorrect was explained in [3]. Namely, the iris texture has little to no correlation with genotypic features, including gender. What was found is that properties of ocular regions, mainly due presence of cosmetics used by women, may indicate a gender of the subject whose iris was photographed.

The author did not make connections between his work and applications in industry.

**Has the author solved the research issues defined in the dissertation, did he use an appropriate method and are the assumptions justified?** (PL: Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?)

The author designed and carried out a series of experiments to verify the formulated hypotheses. The experiments related to face recognition in presence of cosmetics and under different face expressions are correct.

I have identified the following weaknesses in the research methods related to the impact of diabetes on iris recognition and time of day on speaker recognition, and I would be happy to discuss them with the author during his defense:

1. Investigation whether the diabetes impacts iris recognition is based on comparison of iris matching results in two groups, which are heterogenous in terms of age of subjects. Table 3-1 suggests that the average age in the diabetic group is significantly higher than the average age in the healthy group. Iris recognition may perform worse for older subjects due to various factors not related to diabetes (e.g., *arcus senilis*, a grayish arc near the limbic iris boundary decreasing the contrast between the iris and the sclera, is one of known factors). Is it possible that the reason for the observed lower accuracy within the “diabetes” group is caused by age and not by the diabetes? If so, why did the author not make an attempt to reduce correlations between accidental factors (not related to diabetes) and the class-specific iris recognition accuracy?
2. (Related with (1)) The author was not able to find a method to detect diabetes from an iris image (p. 67). At the same time the author claims that diabetes has in impact on iris recognition. Aren’t these two statements contradictory? If it is not possible to detect diabetes from an eye image, could the observed differences for diabetic and healthy subjects be caused by data imbalanced in terms of the subjects’ age?
3. Having data collected only in a single morning session and in a single evening session does not allow to disentangle the effects originating from between-session measurements and daytime-related variability in human voice. Fig. 6-6 demonstrates between-session vs within-session score distributions, which can’t be used to conclude whether morning-vs-morning comparisons are different from morning-vs-evening comparisons. This experiment can only demonstrate whether morning-vs-morning comparison scores are different from evening-vs-evening comparison scores, and Fig. 6-6 suggests that they are not significantly different.

Additional question relates to the iris recognition methodology:

1. According to the equation (3-2) on. p. 37, the selected iris recognition method does not incorporate occlusion masks into calculating the Hamming distance. Occluded parts of the iris should not contribute to the matching score. Is it only a mistake in the thesis writing and the author did incorporate the mask into calculating the comparison score in his programs? A minor remark is that USIT was proposed by Rathgeb et al. (<https://www.wavelab.at/sources/USIT>), not Monro et al. (p. 39).

**What are novel elements of the dissertation, what are the author's independent and original accomplishments, what is the position of the dissertation in relation to the state of the art and the level of technology represented in the world literature?** (PL: Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy i poziomu techniki reprezentowanych przez literaturę światową?)

The accomplishments presented in Chapters 3, 4, 5 and 7 have been previously published as peer-reviewed papers, coauthored by Mr. Azimi. Upon my request, Mr. Azimi delivered a written statement on his individual contributions to these papers. According to the statement, Mr. Azimi was a leading author responsible for all aspects of the research, except for new data collections. Such contributions are sufficient and justify having these works included in the dissertation. The most important and original element of the thesis, triggering interest among other biometric researchers, relates to the impact of diabetes on iris recognition. Once the doubts related to experimental design are clarified, it may serve as valuable contribution to the community. Especially availability of the database makes this research reproducible.

**Has the author demonstrated capabilities to present the obtained result in a correct and convincing way (brevity, clarity, editorial correctness)?** (PL: Czy autor wykazał umiejętność poprawnego i przekonywującego przedstawienia uzyskanych przez siebie wyników (zwięzłość, jasność, poprawność redakcyjna rozprawy)?)

The thesis is written in a clear way. There are several statements that would require a better justification or more precise formulations:

-- *Using a biometric identifier is more reliable than traditional methods such as passwords or identity cards* (p. 1). This may be not true if put into a concrete authentication context. For instance, a weak face recognition system (with high FAR and lacking liveness detection) may be less reliable and more harmful than password-based authentication.

-- *Retinal detachment can permanently change the trait of the iris* (p. 2). Retina and iris are two different parts of the eye. Appropriate citation should be provided to support the above claim.

-- I would consider sensors (their “aging” and technological changes) as one of the factors of degradation of biometric performance over time (p. 2).

-- *Pupil dilation cannot directly influence the unique iris features themselves* (p. 45). Although the meaning of “influence” is unknown here, pupil dilation does have an impact on iris recognition accuracy, mainly due to a nonlinear nature of iris deformation, which is not compensated by linear methods, such as Daugman’s “rubber sheet” model.

Chapter 2.1.2.2 seems to be a bit chaotic: iris and speaker recognition systems are intertwined into a single narrative. I also noticed a few minor language errors: text repetitions “such as such as” (p. 7), typos “Daughman” instead of Daugman (p. 21), paragraph repetitions (e.g., approval of Ethics Committee of the Warsaw University of Technology, usage of the IriShield MK 2120U sensor, or USIT method), or unclear statements (e.g., what do “truncated Chebyshev series processing minimax properties” or “peaks of pixel paths” mean?). These, however, do not preclude from understanding the main research efforts and obtained results described in the thesis.

**What is the value of the dissertation for engineering and technical sciences?** (PL: Jaka jest przydatność rozprawy dla nauk inżynieryjno-technicznych?)

This thesis tackles an important problem of biometric reliability under adverse circumstances and makes an attempt to assess how selected biological and behavioral factors impact three popular biometric modalities. The thesis in its present form rises the awareness of the community to a potential relation between diabetes and iris recognition and offers first experiments to demonstrate causality, and thus has a positive value. I have two main concerns related to the experimental design, that should be addressed before the next publications coming out of this research: (a) class labels are correlated with average age in diabetes-related research, and (b) assessment of how time of day impacts the speaker recognition can’t be done with single acquisition sessions. Once these doubts are clarified, the value of the conducted research will be significantly higher.

In summary, **I classify this dissertation to category:**

**(c) meeting requirements** (PL: (c) spełniająca wymagania)

References

[1] S. Lagree and K. Bowyer, "Predicting ethnicity and gender from iris texture," IEEE HST, 2011

[2] A. Kuehlkamp and K. Bowyer, "Predicting Gender From Iris Texture May Be Harder Than It Seems," CVF/IEEE WACV, 2019

[3] A. Kuehlkamp *et al.*, “Gender-from-Iris or Gender-from-Mascara?”, CVF/IEEE WACV 2017

