BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF INFORMATION TECHNOLOGY
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

DEPARTMENT OF INTELLIGENT SYSTEMS
ÚSTAV INTELIGENTNÍCH SYSTÉMŮ

NEW TECHNOLOGIES FOR BIOMETRIC RECOGNITION BASED ON HAND CHARACTERISTICS

NOVÉ TECHNOLOGIE PRO BIOMETRICKÉ ROZPOZNÁNÍ NA ZÁKLADĚ CHARAKTERISTIK RUKY

PH.D. THESIS
DISERTAČNÍ PRÁCE

AUTHOR **Ing. MICHAL DVOŘÁK**
AUTOR PRÁCE

SUPERVISOR **Prof. Ing., Dipl.-Ing. MARTIN DRAHANSKÝ, Ph.D.**
ŠKOLITEL

BRNO 2022

Abstract

Biometrics allows us to identify or verify people based on their physical characteristics. Although the recognition itself may be very fast, extracting the chosen biometry features in the real world is a non-trivial task. To perform a recognition on a large number of people, for example at an airport, we need to perform the scanning and recognition in the shortest possible time, ideally without the user having to stand still. In this dissertation thesis, I focus on the development of a touchless and therefore hygienic biometric system, capable of performing identification, feature extraction and subsequent recognition on the fly using hand characteristics, and other systems and algorithms needed to solve this complex problem.

This thesis presents the current state of research and development in the field of biometric identification based on selected hand characteristics, with a focus on identification using hand geometry and fingerprints. Further it also discusses the technological principles used for the collection of the characteristics.

According to the dissertation objectives, formulated on the basis of the research, the practical part of this thesis deals with the development of biometric systems for use to identify large numbers of people in flight, for which most current biometric systems based on hand characteristics are not suitable. The thesis then addresses related issues, in particular presentation attack detection and biometric multimodality.

Abstrakt

Biometrie nám umožňuje identifikovat či verifikovat osoby na základě jejich fyzických vlastností. Přestože identifikace samotná je velmi rychlá, extrakce zvolené biometrie v reálném světě je netriviální problém. Pro snímání většího množství lidí, například na letišti, potřebujeme dosáhnout snímání a rozpoznání v co nejkratším čase, ideálně za chůze uživatele. Tato disertační práce se zabývá vývojem bezdotykového a tedy hygienického biometrického systému, který je schopen provést extrakci identifikačních charakteristik ruky a následně rozpoznat uživatele za běhu. Spolu s tím práce obsahuje další systémy a algoritmy potřebné pro řešení tohoto komplexního problému.

Práce prezentuje současný stav výzkumu a vývoje v oblasti biometrické identifikace na základě vybraných charakteristik, se zaměřením na identifikaci pomocí geometrie ruky a otisků prstů. Dále se pak věnuje technologickým principům využitých pro sběr těchto biometrických charakteristik.

Dle disertačních cílů, formulovaných na základě rešerše, se praktická část této práce zabývá vývojem biometrických systémů pro použití k identifikaci velkého množství lidí za běhu, pro které většina současných biometrických systémů založených na charakteristikách ruky nejsou vhodné. Dále pak se práce zabývá souvisejícími problémy, zejména pak detekcí prezentačního útoku a biometrickou multimodalitou.

Keywords

Biometrics, hand geometry, 3D hand geometry, biometric identification, biometric verification, line scan camera, dynamic acquisition, on the fly acquisition, stereo vision, watchlist technology

Klíčová slova

Biometrie, geometrie ruky, 3D geometrie ruky, biometrická identifikace, biometrická verifikace, řádková kamera, dynamické snímání, snímání za běhu, stereo vize, seznam sledovaných položek

Reference

DVOŘÁK, Michal. *New technologies for biometric recognition based on hand characteristics*. Brno, 2022. Ph.D. thesis. Brno University of Technology, Faculty of Information Technology. Supervisor Prof. Ing., Dipl.-Ing. Martin Drahanský, Ph.D.

Rozšířený abstrakt

Biometrie je vědní obor využívající fyzických či behaviorálních znaků osoby k určení její identity. Použití biometrických údajů pro rozpoznání uživatele se již stalo zcela běžnou součástí jak spotřební elektroniky, zejména mobilních zařízení, tak i bezpečnostních přístupových systémů nebo systémů pro sledování docházky. S rostoucím výpočetním výkonem vestavěných zařízení a jejich klesající cenou se otevírají možnosti využití biometrie v aplikacích, které do nedávna nebyly technologicky realizovatelné či finančně atraktivní.

Hlavní část této práce se zabývá řešením problému propustnosti běžných biometrických skenerů a jejich nasazení v atypických aplikacích. Požadavek na vložení ruky uživatele do skeneru a vyčkání, dokud snímání není ukončeno, představuje hlavní překážku pro rychlou identifikaci a určuje, kolik uživatelů lze pomocí dané technologie rozpoznat za jednotku času. Jako součást komplexního řešení se práce taktéž věnuje přidruženým tématům, zejména pak biometrickému rozpoznání na základě 2D a 3D geometrie ruky, rozpoznání na základě otisků prstů, biometrické multimodalitě a prezentačnímu útoku na biometrické systémy.

Práce je rozdělena do tří hlavních částí:

- V první kapitole je představena rešerše existujících akademických a komerčních řešení včetně rozboru použitých technologií, představení relevantních částí biometrie a fyziky návrhu optických systémů.
- V druhé kapitole je představen návrh a vývoj nových biometrických systémů a technologií, které řešící jeden z biometrických problém v současné době existujícíchaspekty problému dané zadáním.
- Ve třetí kapitole následuje prezentace ovládacích a biometrických algoritmů vyvinutých pro zařízení, která byla vyvinuta v rámci druhé kapitoly. Tato kapitola taktéž popisuje získávání dat, zpracování dat a vyhodnocení dat získaných ze systému představených v rámci kapitoly dva a dále popisuje výsledky ze souvisejících výzkumných činností.

V průběhu práce byly rozvedeny dva hlavní technologické koncepty, a to snímací technologie s nastavitelným zorným polem pro snímání heterogenních biometrických údajů a technologie pro bezdotykové snímání biometrických údajů za běhu.

Technologie pro snímání heterogenních biometrických údajů umožňuje vizuální zpracování různě velkých částí těla, bez nutnosti měnit fyzickou konfiguraci zařízení, při zachování rozlišení na jednotku plochy danou použitým senzorem. Toto zařízení změnou zorného pole umožňuje snímání prstu, ruky nebo celé paže při zachování plošného rozlišení.

Požadavky na takové zařízení jsou dány, dle rešerše v oblasti biometrie, minimálním plošným rozlišením, což v tomto případě bylo 350 bodů na palec (DPI) a zorným polem o šířce alespoň 400 mm a délce 180 mm s tím, že délka je operátorem nastavitelný parametr. Pro realizaci proměnného zorného pole byla využita řádková kamera připevněná na ose s lineárním pohonem. Na základě výše uvedených požadavků a zvolených rozměrech snímacího zařízení byly vypočítány ideální parametry optických prvků. Vypočítané ideální parametry vedly k výběru kamery AL6144-16gm Basler, objektivu AF Nikkor 50mm f/1.8D, lineární osy NL 208 TR a elektronických komponentů realizujících pohyb kamery po lineární ose. Zvolené komponenty byly využity pro konstrukci laboratorního prototypu. V rámci vývoje tohoto zařízení byly navrženy a realizovány systémy pro pohyb osy, ovládání kamery, osvětlení a napájení, včetně implementace jejich řídících algoritmů.

Laboratorní prototyp byl použit pro vytvoření databáze, která obsahuje 56 snímků dlaně a 89 hřbetů ruky o rozlišení 6,144 × 5,120 px, které představuje zorné pole o velikosti 417.1 × 195.3 mm. Snímky dlaně byly použity pro demonstraci použitelnosti prototypu pro účely biometrické identifikace. Z výřezů obrazů, které obsahují konce prstů a s použitím

algoritmů na zvýraznění papilárních linií bylo extrahováno 230 otisků prstů. Tyto otisky byly kvalitativně ohodnoceny pomocí nástroje VeriFinger (verze 10) a FiQiVi. Medián kvality VeriFinger byl určen jako 32. Výrobce VeriFinger Neurotechnology doporučuje kvalitu pro verifikaci alespoň 30, což vzhledem k obecnosti návrhu svědčí o vysoké věrnosti snímaného obrazu, která je více než dostatečná pro použití při biometrické identifikaci využívající geometrii ruky. Tyto výsledky včetně demonstrace využití v biomedicíně byly prezentovány a publikovány ve sborníku konference BIODEVICES.

Koncept skládání obrazových dat pro vytvoření komplexní informace byl taktéž zkoumán při realizaci 3D skeneru založeném na principu stereo vidění. Zařízení primárně vyvíjené pro projekt VRASSEO Ministerstva vnitra bylo použito pro ověření možnosti využití stereo vidění jako způsobu získání 3D biometrických příznaků. Zařízení skládající se z pohyblivého vozíku, který přesouvá definovaným způsobem kameru po zakulacené kolejnici, umožnilo získání obrazových dat snímaného objektu z arbitrárně velké množiny úhlů. Přestože rekonstrukce 3D objektu ze získaných dat byla ověřena a pro účely projektu byla dostatečná, výsledná hrubost představovala takovou překážku pro biometrickou aplikaci, že byl tento směr výzkumu přerušen ve prospěch slibnějšího směru, a to snímání za běhu.

Technologie pro snímání biometrických údajů za běhu představuje hlavní část této práce. Požadavky na tuto technologii lze shrnou následovně:

- Systém musí být schopný snímat obrazy dlaně v plošném rozlišení alespoň 500 DPI,
- Systém musí být schopný snímat obrazy dlaně při pohybu dlaně o rychlosti alespoň 500 mm/s,
- Systém musí být schopný snímat obrazy s hloubkou ostrosti alespoň 35 mm.

Dodržení těchto požadavků bylo určeno jako instrumentální pro biometrickou aplikaci, ať již se záměrem použít otisky prstů, geometrii ruky či jejich kombinaci.

Vzhledem k tomu, že v době psaní této práce byl tento koncept v akademii neexistující a v průmyslu se objevovala pouze jedna podobná aplikace, bylo nutné zařízení navrhnout od fyzikálních principů, a to s ohledem na parametry snímacího elementu, objektivu, fyzické vzdálenosti mezi snímacím a snímaným elementem a osvětlení. Na základě teoretického návrhu bylo vytvořeno zařízení pro zkoušku konceptu. Tento koncept byl prezentován průmyslovému partnerovi TBS a ve spolupráci s ním byl vytvořen laboratorní prototyp a konečně i prototyp komerční. Tento poslední prototyp byl dobře přijat na bezpečnostním veletrhu INTERSEC v Dubaji.

Finální verze systému umožňuje identifikaci spolupracujícího uživatele za běhu, kde uživatel při chůzi nastaví ruku vhodně vůči zařízení a nechá jí projít snímací bránou. Tento průchod je zaznamenán a aktivuje zařízení. Kamera acA5472-17um Basler s 1" objektivem o ohniskové vzdálenosti 35 mm následně snímá obraz ruky, která je osvětlena pomocí k tomuto účelu navrženým světelným panelem. Pro napájení a spínání jednotlivých komponentů byla navržena a vyrobena vlastní deska kombinující obvody pro tyto dvě funkcionality.

Na verzi pro zkoušku konceptu byla demonstrována schopnost snímat a extrahovat otisky prstů v kvalitě přijatelné pro biometrické rozpoznání za letu, a to na databázi vzniklé za pomocí 82 dobrovolníků, z nichž každý sepnul zařízení alespoň 10 krát. Výsledných 908 obrázků dlaně zahrnuje 4,525 snímků konečků prstů, které obsahují informace potřebné pro extrakci otisku prstu. Tyto obrázky byly zpracovány pomocí algoritmů na zvýraznění otisků prstů. Průměrná kvalita těchto otisků prstů po jejich zpracování byla dle Verifinger SDK verze 9.0 na hodnotu 68.0 a dále pak na hodnotu 1.85 dle NFIQ, což představuje kvalitu dostatečnou pro identifikaci i verifikaci.

Na laboratorním prototypu byla testována schopnost zařízení rozpoznat uživatele za běhu na základě geometrie ruky. Výsledky byly prezentovány v článku odborného peri-

odika IET Biometrics. Míra úspěchu rozpoznání byla demonstrována jako srovnatelná se současným stavem techniky, toto řešení navíc na rozdíl od současného stavu umožňovalo identifikaci bezdotykově a za běhu. Sběr dat na tomto zařízení byl uskutečněn s pomocí 99 dobrovolníků, kteří byli instruováni, jak zařízení použít. S jejich pomocí bylo získáno 573 obrázků dlaně. Databáze byla anotována za účelem rozpoznání na základě geometrie ruky. Klasifikátor náhodných stromů (Random forest classifier) byl použit pro realizaci výše uvedeného účelu a došel k míře úspěšnosti 0.960 se standardní odchylkou 0.015, což na poli rozpoznání za běhu představuje špičku a při srovnání s dotykovými technologiemi se jedná o kompetitivní výsledek.

V neposlední řadě byla poslední verze zařízení prezentována na veletrhu INTERSEC a taktéž sloužila jako vzor pro podání a následné udělení mezinárodního patentu DE102020200569A1: Einrichtung zur biometrischen Identifikation mit Hilfe von Fingerabdrücken und/oder Handcharakteristiken und Verfahren zur biometrischen Identifikation mit Hilfe dieser Charakteristiken (Zařízení pro biometrickou identifikaci pomocí otisků prstů a/nebo charakteristik ruky a způsob biometrické identifikace pomocí těchto charakteristik).

Schopnost snímat různé biometrické příznaky souběžně dává možnost využít principů multimodální biometrie. Touto problematikou se práce s ohledem na souvislost s hlavním tématem zabývá ve spojitosti s aplikací při seznamu sledovaných položek (Watchlist), kde multimodální biometrie je prezentovaná jako způsob snížení ceny způsobené chybovostí systému. Výsledky tohoto výzkumu byly publikovány v článku odborného periodika IET Biometrics: Biometric-Enabled Watchlists Technology.

Poslední související výzkumný aspekt, kterým se práce zabývá, je otázka detekce prezentačního útoku (presentation attack detection). Využití 3D příznaků pro systémy rozpoznání geometrie ruky je zkoumáno jako možný bezpečnostní prvek a je prezentováno ve zvané kapitole v publikaci 3D Hand Geometry Recognition, kterou vydalo akademické nakladatelství Springer. Zároveň byl zkoumán princip detekce založený na multispektrální analýze. Pro ověření tohoto principu byl vyvinut funkční prototyp, který byl následně přihlášen jako užitný vzor.

Tato práce si vytyčila poměrně ambiciózní cíl a to nabídnout změnu paradigmatu snímání biometrických příznaků. Zcela neskromně je třeba říct, že po výzkumné stránce byl tento cíl naplněn.

Technologie pro snímání heterogenních příznaků za pomocí kamerového systému s nastavitelným zorným polem, který využívá řádkovou kameru, byla včetně relevantních algoritmů realizovávána a výsledky publikovány ve sborníku konference BIODEVICES. Ačkoliv byl koncept ověřen a využitelnost pro biometrickou aplikaci potvrzena, ve vývoji této technologii nebylo pokračováno, ve prospěch systému pro snímání za běhu.

Jako řešení problému nízké prostupnosti biometrických systémů tato práce navrhla, vyvinula a otestovala biometrický systém, který je schopný snímat obrazy ruky bezdotykově za běhu, v kvalitě dostatečné pro biometrickou identifikaci na základě biometrických charakteristik ruky.

Zařízení a princip rozpoznání na základě otisků prstů za běhu bylo ve spolupráci s průmyslovým partnerem Touchless Biometric System, patentováno mezinárodním patentem a funkční vzorek byl prezentován na bezpečnostním veletrhu INTERSEC. Zařízení se setkalo s dobrým přijetím a verze 3D Fly, kterou průmyslový partner dovyvinul je v současnosti nabízena na jeho stránkách.

Princip rozpoznání na základě geometrie ruky za běhu, včetně popisu systému a algoritmů, byl publikován v impaktovaném akademickém periodiku IET Biometrics. Článek představuje jak návrh optického systému dle nadefinovaných obecných požadavků na za-

řízení, tak i popis sběru dat a implementaci algoritmu pro rozpoznání, využívající klasifikátor náhodných stromů. Míra úspěšnosti tohoto systému je pak bezkonkurečně nejlepší ve své kategorii, neboť se jedná o jediný systém rozpoznávající na základě geometrie ruky za běhu. Nicméně i ve srovnání s klasickými identifikačními systémy si i přes přidanou komplexitu počíná lépe než mnohé jiné, pomalé dotykové systémy.

V průběhu práce na systému pro biometrickou identifikaci za běhu byly zkoumány i další související technologie a biometrické principy. Nejvýznamnější výsledek přinesla spolupráce na publikaci zabývající se biometrickou multimodalitou a jejím přínosem v technologiích využívajících seznam sledovaných položek (Watchlist), který byl publikován jako článek v impaktovaném odborném periodiku IET Biometrics.

Jako řešení problému detekce prezentačního útoku bylo představeno zařízení využívající multispektrální analýzu pro detekci nástroje prezentačního útoku, jehož princip byl registrován jako průmyslový vzor. Problematika a prevence detekce prezentačního útoku společně s biometrickou identifikací založenou na geometrii ruky byla diskutována v rámci zvaných kapitol v knize 3D Hand Geometry Recognition vydané akademickým nakladatelstvím Springer.

Úspěšným vývojem a prezentací zařízení pro biometrickou identifikaci za běhu, společně s dalšími výzkumnými aktivitami presentovanými v této práci, považují zadání za splněné. Publikace v impaktovaných odborných periodikách a konferenčních sbornících, autorství a spoluautorství na zvaných kapitolách a původcovství mezinárodního patentu považuji za uznání kvalit akademickými a odbornými autoritami.

New technologies for biometric recognition based on hand characteristics

Declaration

I hereby declare that this Ph.D. thesis was prepared as an original work by the author under the supervision of Prof. Drahanský. I have listed all the literary sources, publications and other sources, which were used during the preparation of this thesis.

.....
Michal Dvořák
April 29, 2022

Acknowledgements

First and foremost I would like to express gratitude to my supervisor, Prof. Drahanský for his guidance, support and and especially patience he extended me during my Ph.D. study. His breadth of knowledge has helped me immensely when trying to find the research direction in the field of biometry, and his supervision helped me focus my efforts. I would also like to thank Prof. Waleed H. Abdulla and Prof. Svetlana Yanushkevich who have welcomed me to their universities in New Zealand and Canada and allowed me to collaborate with them, it was an experience which had profound impact on me and I am trully grateful for it. Last but not least, I would like to thank my my parents, and fiance (possibly wife depending on the date of the defence). Without their support and inquires about the state of the thesis, it might have never get done.

Contents

Introduction

In the world today, there is a demand for entry mechanism that fulfills both the requirement for ease of access, as well as the security properties necessary to prevent an unauthorized access. Whether it is used as a system protecting a location or a software, the ideal solution would be the one, where the authorized user must do nothing, but approach said system, while unauthorized attacker, will be unable to access it.

Biometrics is a field of study, that uses physical or behavioral features of a person in order to determine his or her identity. Already many biometrics are being used and are becoming a very common feature today both in consumer electronics, especially mobile devices as well as security access systems or attendance tracking systems. Most used biometrics are fingerprints, face geometry and eye-based systems commonly based on retina and iris recognition. But as no biometrics is usable or practical in every situation, new approaches to recognition based on other biometrics are being developed, alternatively, existing systems, that have not, for example due to technological issues, experience the same rate of adoption, are being improved. One of these approaches is a utilization of hand geometry and associated features present on the palmar side of the scanned hand, such as palmprint and vein patters.

As with any other biometrics, the hand-geometry based biometrics arises from an assumption, that the shape of one's hand is unique for every individual. Or, that probability of two humans having the same hand shape given the defined database size is low enough, to be considered negligible. At the same time, the biometry needs to be immutable in time, that within the constraints of acquisition process, the model of scanned hand will remain identical across multiple acquisitions.

In both the commercial and academic world, the devices used for the biometric scanning are mostly unchanging, especially when it comes to underlying technology used. This has led to very robust and fast methods of biometry acquisition, but only for the originally developed scenarios. As such some avenues of research have not been explored. Some biometrics have been temporarily abandoned such as hand geometry identification, which until recently, saw a very limited use. This was mostly due to computing limitations as well as hardware costs, the existing implementations and research concentrated solely on basic 2D features of human hand at a low resolution. This approach, while very cost effective, was too constrictive to allow for wider spread.

With the increased processing power of embedded devices and increased capabilities of acquisition devices however, new avenues of research are opening, especially in the area of 3D hand geometry identification, and multimodal biometrics, that combines features from multiple biometrics, such as hand geometry, palmprint and fingerprints.

The second avenue of research addresses the throughput limitation of conventional biometric scanners. The majority of currently used biometric technologies requires a physical contact with the device, as is most notable in the fingerprint identification-based systems or

at the very least requires the user to stop moving for the duration of the scanning process. This requirement presents an upper limit to how many users can be processed using the mentioned technology. Not to mention the hygienic concerns, that should be considered if such technology has to be used in the public area. To solve both of these problems a new scanning technology needs to be developed, that allows the biometric scanning process to take place on the fly, does not require to be standstill during acquisition and forfeits the requirement of physical contact between the user and the biometric system. This requirement especially in the view of the recent COVID pandemics, seems to be a very useful and worth investigating, as it would pave a way to a much more hygienic recognition without relying on the morally controversial face tracking and recognition.

This work strives to explore and contribute to the above-mentioned areas by investigating possible improvements or creation of algorithms and or devices, that allow for on the fly biometric data collection as well as exploring the viability of various hand-based biometrics for the purposes of biometric identification. This work shall research and discuss the hand related biometrics, with focus on 2D and 3D hand geometry based biometrics, fingerprints and multimodal biometrics.

This thesis is separated into three main parts- in the first chapter, biometry as a field of study shall be introduced, describing the taxonomy, principles and evaluation metrics for understanding the later parts of the thesis. Brief introduction into optical system design will be presented as it will be referenced in the practical part of this thesis. Technological principles used in scanners for collection of biometric features will be presented in this part, as well as research of current state of the art in scanning technologies both in academia and commercial applications.

The third chapter of this thesis will feature the description of the development of novel scanning technologies with focus on touchless and on the fly scanning of biometric features. This part will present the design of the optical and electrical properties of the system and its relevance for biometric purposes as well as methodology and scenarios designed and utilized for the data collection using said system.

The fourth chapter will describe the data collection processes and methods for processing the data acquired. Dealing mainly with the image enhancement methods, used to alleviate the noise and distortions in the images, introduced by the on the fly data collection. Algorithms dealing with the feature extraction methods for the hand geometry and fingerprint identification both original and using state of the art will be shown. This part will also include the evaluation of the individual identification methods, thus presenting a solid metrics for this new biometric system. This chapter will also discuss the research and the results of the tangential areas, related to the hand biometry, as explored during the main research. Focusing on the exploration of multimodal biometry, as well as the presentation attack instrument development, using various materials and presentation attack detection utilizing multi-spectral analysis, and 3D geometry analysis. Furthermore, in this chapter, we will discuss the possible future development in the area of the hand biometry, based on the research in this thesis.

The claims, devices and results in the above-mentioned areas will refer to the respective publications, submitted and accepted by various academic journals and conferences, as well as the utility models, patents and testimonials of the industrial partners, where relevant.

Dissertation goals

1. **Development of biometry enabled optical scanning devices**
   This goal aims to investigate current methods of data acquisition utilized by existing biometry enabled systems which use the features present on the hand. Based on this research shortcomings shall be identified and avenues of potential improvement and novelty pursued.
   New systems that tackle one or more of these shortcomings shall be presented and discussed. Design process of the individual subsystems shall be presented and justified in regards to the theoretical principles and the intended goal. These devices shall be developed in a manner that facilitates the goal number two of this dissertation thesis.
   Novelty and scientific contribution of the developed systems shall be demonstrated using relevant patent applications and/or receiving endorsements from an industrial partner.
2. **The applications and evaluation of novel biometry enabled devices**
   Secondary goal is to demonstrate the utility of the systems developed in as part of the first goal in the manner compliant with the topic of this thesis.
   Software and firmware developed to facilitate biometric application of the discussed and the development presented.
   Database collection where relevant shall be discussed as well as the initial conditions, the properties of the databases and the intended application.
   Additional applications such as multimodal biometry and its advantages in biometric recognition shall be discussed if found relevant.
   Novelty and scientific contribution of the developed systems shall be demonstrated by publishing the results in academic journals, conferences and/or receiving endorsements from industrial partner.

Chapter 1

Theory

In this chapter, we will cover the necessary theoretical background, with the expected focus on the biometry, especially hand-based biometry such as fingerprint, palmprint and hand geometry identification and verification. Various evaluation metrics used in this field and the current state of the art of the individual biometrics will be presented here. Due to the practical aspects of this thesis, part of this chapter will be dedicated to the sensorics and optics theory, focusing on the subset necessary for the optical system design, as the further chapters will assume this knowledge.

Biometry is a field of study, that deals with identification of a user based on his or her unique active or passive physical properties. Where the active physical property is characteristics defined by the user's expression and is both physical and behavioral, such as gait or signature characteristics, whereas the passive is solely physical and utilizes biometrics such as fingerprint, iris, retina, face, DNA and multitude of other less common approaches to recognition. While there are historical reasons and findings that people have been using biometric principles to identify themselves or their creation since antiquity, and perhaps even earlier. The biometry utilizing scientific method can be traced to the late 19^{th} or early 20^{th} century, where anthropometry, i.e. the use of bodily measurements to identify a person, and dactyloscopy which utilizes fingerprints, have been first introduced as a secure approach to identification. Both methods are readily used to this day.

As will be elaborated on in the upcoming parts, in its most basic form, biometric system consists of a sensor, for data collection, feature extractor module, that locates and extracts the unique features of the chosen biometrics, matching module that can compare the set of features and determine the likelihood of the sets belonging to the same user or in case of verification systems, determine the most likely user identity. The system should also include a registration module, that allows enrollment of the new set of user features also known as biometric templates, and a storage module to save these templates.

1.1 Hand geometry recognition

Historically, identification and self identification of humans, based on their hand shape, is one of the first examples of applied biometry. This is due to low demands on the acquisition and processing system. Unlike fingerprint based systems, where the resolution has to be high enough to capture the papillary lines, or even the sweat glands on the papillary ridges, for hand geometry acquisition, these demands are much lower.

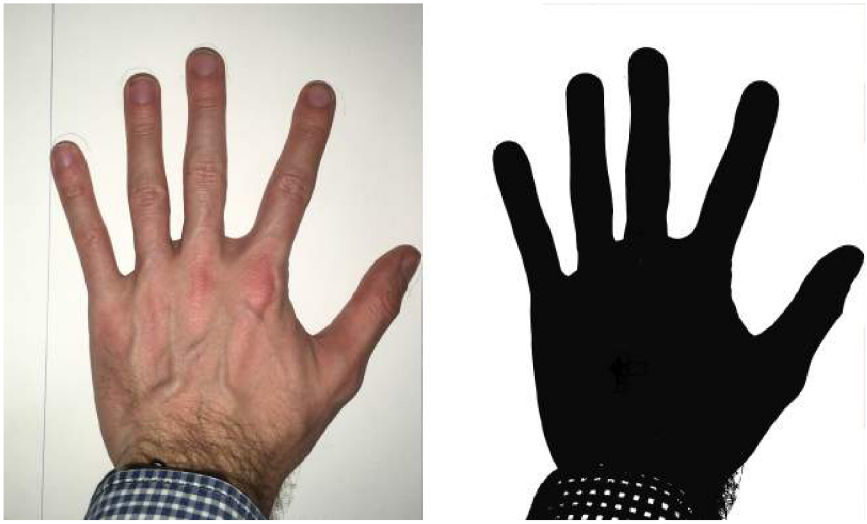

Figure 1.1: 2D hand geometry acquisition example, source image (left), extracted shape (right).

At this date, the hand geometry based systems, while existent, are not at the forefront of the development, due to limitations that will be examined in this chapter. However, in order to build on the existing research and development, the current state needs to be examined and analyzed, both from the academic point of view, and commercially available solutions, to establish available avenues of research that are also sought after by the security industry.

In this chapter, we will examine current academic state of the 2D hand based biomterics as well as the current state of research in the area of 3D hand based biometerics. Both from the acquisition side of the process, as well as subsequent processing and evaluation part of identification. After that, we will examine the commercially available devices, that utilize these principles, to better understand the requirements placed on these devices and use it to postulate the appropriate course of development.

Experimental results show that images acquired at resolution of only 100 dots per inch (DPI) can be processed with satisfactory performance [36]. On Fig. 1.1 we can see the result of segmentation process, from which the features would be extracted. Understandably the computing and acquisition requirements are lower, then when compared to fingerprint biometrics.

1.1.1 State of the art

Hand geometry verification has been explored for many years. Ross [83] presented a prototype for acquiring a hand image at 640x480 pixels using a charge-coupled device (CCD) camera. This device utilises five pegs that guide the hand into an appropriate position before the acquisition. The user remains stationary in front of the device and places the hand at the extreme position the pegs allow; then a photo is taken. Sixteen measurements of absolute length measurements were collected from each of the collected photos. Euclidean distances were used for the classification.

The approach to this form of recognition has not changed significantly since then. Putra [80] and Levicky [98] used different classification algorithms. Yoruk [110] and Vergara

Table 1.1: Performance comparison of related recognition systems [35]*.

Author	Feature source	Classification	Features	Acquisition method	Pegs?	Database	SR [%]
Angadi [5]	Eedge connected graph	SVM classifier	30	Touch based	No	1,440	97.95
Adan [2]	Hand contour	Normalized Similarity Measure	27	Touch based	No	470	96.6
Singh [89]	Hand contour	SVM classifier	24	Touch based	No	480	96.53
Boreki [7]	Hand contour	Euclidian distance classifier	58	Touch based		360	96.2
Yoruk [110]	Global hand shape	Hausdorff distance and ICA	2048 / 200	Touch based	No	458	95.78
Firas [3]	Hand contour	BPNN with MSE value	66	Touch based	No	100	93
Putra [80]	Hand contour	DTW distance	8 codes	Touch based	Yes	900	84
Levicky [98]	Hand contour	GMM	21	Touch based	Yes	408	80.18
Vergara [99]	DWT hand image	Nearest Neighbor	31	Touch based	No	120	70.2
Svoboda [94] (2D only)	Hand contour	Mahalanobis distance	41	Contactless	No	128	EER= 5.36
Kanhangad [52] (2D only)	Hand contour	Euclidean distance	—	Contactless	No	3,540	EER= 6.3
Raghavendra [81]	Face image	Gabor-SRC	—	Contactless/ OTF	No	54 users	91
Kuzu [57]	Finger Vein images	Densenet-201	—	Contactless/ OTF	No	900	97.97
Proposed system	**Hand contour**	**RFC**	**263**	**Contactless/ OTF**		**564**	**96**

[99] did not use the hand contour as a feature source, yet they used the global hand shape and discrete wavelet transformed hand image, respectively. However, the systems are very similar in design and operation. The hand should be placed on a flat surface during acquisition, and a digital camera placed above or below is used to capture the hand image. The most significant design difference is that, unlike [83], [80] and [98], newer works no longer use the guiding pegs. The aforementioned [110], [99] no longer use them, as well as, for example [5][2][89][7] and [3].

All these systems are pegless, meaning the user may position his or her hand freely in the scanning area. However, the system does require that user places his hand on top of a surface, meaning that these systems are not truly contactless.

These and other hand geometry based systems and their performances can be seen in the Tbl. 1.1. This table also includes performance of system developed as a part of this thesis and will be explored in future chapters [35]*.

In the most systems, as will be further explored in a subsequent part of this chapter, only the hand silhouette is extracted from the acquired image and all the measurements are then performed on this binary mask. The advantage is two fold, due to lack of demand for skin feature-level details, the acquisition aperture can be scaled down, making the final solution more attractive from economic standpoint. The resultant low resolution and binary nature of the segmented image also leads to decreased complexity of identification algorithms, making the final process either less time consuming, or placing lower demands on the processing units, making it a suitable solution for embedded systems.

1.1.2 Recognition methods

Unlike fingerprint-based biometrics, where the majority of the systems are based on the minutiae analysis, the hand geometry does not have an established “best” approach, aside from the fundamental idea of physical hand dimensions being used. These approaches need to consider that there are multiple joints inside one hand, which allows for different position of individual fingers and their phalanxes on every image capture, this combined with different pressure, placement or even physical state of the hand will make each reading unique and chosen method needs to take all this under consideration. In general, we can split the today’s methods in two categories. Methods based on direct measurements of various hand parts, and hand shape alignment comparison.

* The asterix (*) indicates that the citation is from a publication which was authored or co-authored by the author of the thesis.

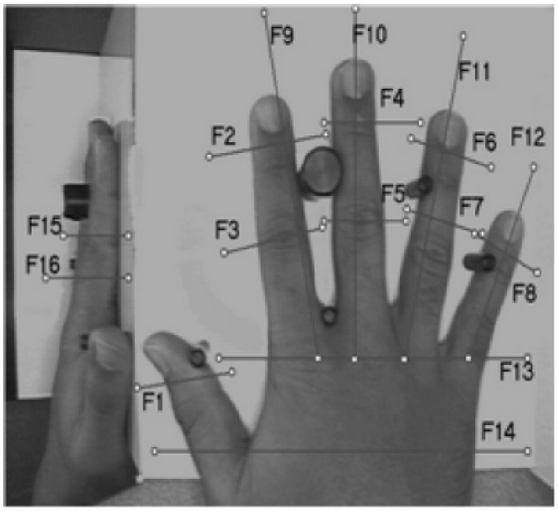

Figure 1.2: Feature extraction from two cameras system [83].

Methods based on direct measurement

This method, perhaps the oldest one, compiles a vector of dimensions of individual fingers and other parts of hand gained directly from the acquired image of a person's hand. This method has been first presented by Ross [83] in his prototype. The example on Fig. 1.2, shows configuration where 16 features are being inferred from a downward facing CCD sensor. To extract the F15 and F16 features a mirror is used. To the best knowledge of author, all commercially available devices today use only the top view camera without the mirror, giving up the side view related features.

The extraction of features can be then done either by thresholding the image and gaining the dimensions straight from the binary image, alternatively, array of pixel values is taken from an unprocessed image and the features are determined individually. The second approach may prove advantageous, especially in case of localized skin discolorations, which can be incorrectly processed, if global or even local transformations are used. It is worth noting, that if utilizing 2D imaging, increase of the number of features above certain threshold adds very little additional useful information for purposes of recognition. During enrollment, one or multiple measurements are performed and saved as a model vector, which we then use during identification/verification phase, where the model vector is matched against the newly acquired vector. Difference can be calculated using any metrics, such as Sum of Absolute Difference Eq. (1.1), weighted Sum of Absolute Difference Eq. (1.2), Euclidean distance Eq. (1.3) or weighted Euclidean distance Eq. (1.4) [24]. The measured features $x_i = [x_1, x_2, .., x_z]$ are then compared to template features $y_i = [y_1, y_2, .., y_z]$ with σ_i being the weight of a given feature. While the usage of weights is not required, given the differences in absolute sizes of these feature, and difference in mobility of individual segments, it has proven to be beneficial.

$$\sum_{i=1}^{z} |x_i - y_i| \tag{1.1}$$

$$\sum_{i=1}^{z} \frac{|x_i - y_i|}{\sigma_i} \tag{1.2}$$

$$\sqrt{\sum_{i=1}^{z} (x_i - y_i)^2} \tag{1.3}$$

$$\sqrt{\sum_{i=1}^{z} \frac{(x_i - y_i)^2}{\sigma_i}} \tag{1.4}$$

Methods based on shape matching

Unlike the direct measurement approach, this method utilizes whole outline of the hand. The template is an outline itself or a normalized set of information that may be used for reconstruction. The following steps describe the general approach to this method:

1. Preprocessing
2. Outline extraction. If the device is well designed, only thresholding is needed for object separation and then apply morphological operations to gain an outline of the hand.
3. Extraction of individual fingers. To compensate for variable finger orientation, the outlines of individual fingers need to be extracted. The tips and valleys of the fingers can be searched for and the range of pixels between them is considered the finger. The orientation normalization can be performed here to prepare for the next step.
4. Perform alignment operations with the template. To do this, any algorithm that measures distance of two metric subspaces can be used, for example Modified Hausdorff Distance [26] where the orientation and position of one outline is gradually changed to achieve closest alignment with the template.
5. Calculate the distance of the pair. MAE (Mean Alignment Error) can be calculated, as an average distance difference between corresponding points of the outlines.
6. If MEA is less then a threshold, it is declared the measured hand is of the same person as the one saved in template.

Fig. 1.3 shows the principle of the comparison, outline is extracted from two hands, one during the enrollment process and other during the identification process. The template of the first is stored. During the identification we prepare a maximum possible overlap of individual fingers and calculate the MEA. Based on the resulting score, we decide whether the hand belongs to the same user. On Fig. 1.3 (right) we can see that while the thickness of fingers displays a high degree of similarity, the length is different. Our threshold value should determine that these two hands are not of the same individual. Both of the described methods serve the purpose they are designed for, both however are plagued by limitations we will look at now.

Classification algorithms

Most of the research for the hand geometry based recognition can be observed in the exploration of classifiers used as can be seen in the Tbl. 1.1, many approaches ranging from direct processing and comparison of the feature values, such as the use of Euclidean distance in works [53], [7], Mahalanobis distance in [93], statistical approach of Nearest neighbour classifier in [99], classical machine learning approaches such as support vector

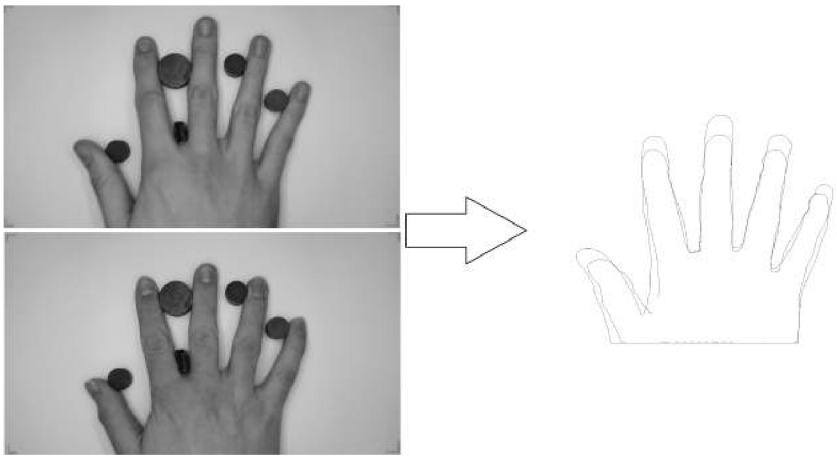

Figure 1.3: Hand image used as a source and other hand image (left), outline overlap after finger positioning (right).

machine (SVM) classifier in [5], [89] and Gaussian mixture model classifier used in [98]. Even the Back propagation neural networks have been used in [3].

All this is to say, that the hand geometry based classification, is a intensely researched topic with many publications serving as a testament to this fact. And due to the nature of the input data, many classification algorithms can still be explored.

To the extent of knowledge of author, there are no existing pegless and by extension contactless commercial devices for hand geometry recognition now available on the market. While it has been demonstrated that pure 2D approach can be used to remove the peg requirement for hand geometry recognition and the chapter presented one of such approaches. There is research into utilization of depth sensing devices, as a method of tackling this problem [49] as well as the need for the physical contact with the sensing device.

1.1.3 Commercially available implementations

To the best knowledge of author, there are only two commercially available systems being currently developed, both by SCHLAGE. Line intended for use as a security admittance mechanism HandPunch© and line for time logging mechanism HandKey©. At the time of writing this chapter, the newest model of the HandKey© series was HandKey II[1]. HandPunch GT-400[2] and HandPunch 4000 then belong to the HandPunch© line[3]. HandKey II, as can be seen on Fig. 1.4 is primarily designed as an access device, it is of sturdy construction and offers security features such as duress code. The base user memory of 512 users can be further expandable. Per available information, the hand geometry features are stored as a 9-byte large compressed template.

The HandPunch series is primarily designed as an attendance and time logging terminal. HandPunch 4 000 also utilizes 9-byte large template, with default database size of 530 users, which can be expanded to 3,498. The HandPunch GT line is also intended as an attendance terminal. The main difference is in the size of the saved template, which is increased to

[1]Available at: http://schlagebiometric.com/products/hk-2/
[2]Available at: http://schlagebiometric.com/products/gt-400/
[3]Available at: http://schlagebiometric.com/products/hp-4000/

Figure 1.4: Image of HandKey II hand geometry reader[1].

20-bytes from original 9. The device does not offer a database expansion, but the default size of 1,000 users is larger than the base model of 4000 line. On Fig. 1.5 the appearance of HandPunch device can be seen.

1.1.4 Limitations of currently employed 2D based biometrics

While the lower requirements on the resolution of the sensor can make the system more cost effective and research shows that the resolution required for hand geometry can be further decreased [36], it also is a cause of several shortcomings that need to be addressed. We can split them into those, that can be solved or mitigated by modifying or improving on existing concept, and those that originate from the 2D acquisition itself, which makes them unsolvable without changing the basic concept. The database size is a large issue for the hand geometry based technologies, as we can see from examination of commercially available solutions such as HandKey II, the focus on increasing viable database size is necessary for further spread of this technology. As has been discussed earlier in the chapter, increasing the number of features extracted from the 2D images, does not yield an adequate increase in the uniqueness. However, it has been demonstrated that, by including the 3D features of the hand, the entropy is increased, which in turn allows for larger user databases.

Just as the requirements for resolution are low, for these biometrics, so is the difficulty of developing a presentation attack instrument that can be used to bypass this technology. As was demonstrated in [11] a simple paper silhouette can be used as a presentation attack instrument that can be used to introduce a successful impostor see Fig. 1.6 for an example of triviality of presentation attack instrument that can be used for the presentation attack. The palmprint recognition faces the similar issue, especially if the simple optical approach is used. The multi-dimensional approach can assist by increasing the complexity of an appropriate presentation attack instrument. Necessary dimensions gathering from uncooperative target becomes a non-trivial task.

Figure 1.5: Image of HandPunch GT-400 hand geometry reader[2].

Figure 1.6: Paper silhouette presentation attack instrument used to bypass commercial device [11].

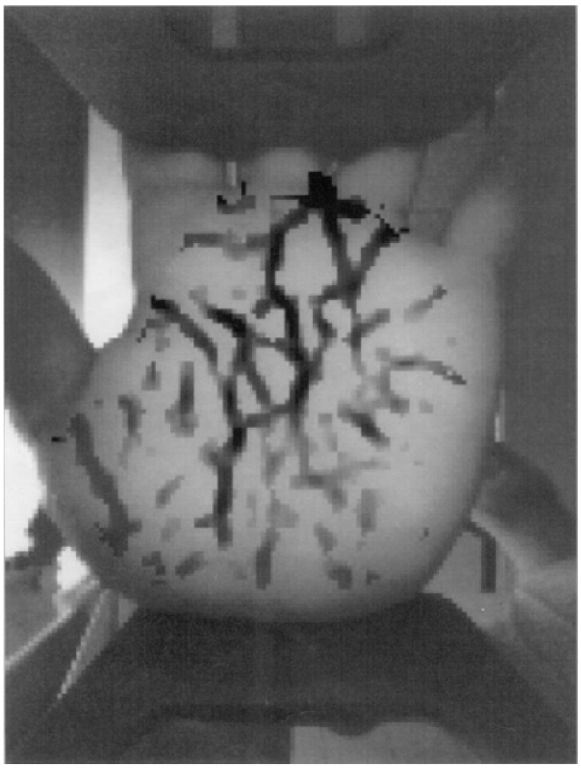

Figure 1.7: Surface vein detection using NIR.

To increase the security of these systems, two approaches are being utilized and further researched. The first approach is to expand the unimodal system into multimodal system, where instead of recognition based solely on one biometrics, we use one or more of others, alternatively we use these additional biometrics simply as means of presentation attack detection and focus solely on detecting their presence.

The presentation attack detection, aims at measuring parameters, we expect to be present in a human hand, but not in a presentation attack instrument. In case the system is developed as touch based, we commonly encounter checks based on electric properties, such as conductivity and capacity. Among other used we can encounter systems detecting heat, utilizing skin spectroscopy, odor analysis and high-resolution skin analysis.

For the system to remain touchless the vein detection can be utilized. The vein detection can either be performed as a surface detection, where by utilizing a near infrared light source (NIR) and performing an acquisition with a sensor without an IR filter, i.e. Fig. 1.7. We can detect veins on palm or dorsal part of hand as the optical absorption of NIR by hemoglobin is higher than that of other tissue and will appear darker on the resulting image.

Instead of capturing the reflection, system can instead capture the dispersed light passing through the hand as in Fig. 1.8. Due to the hand structure, systems using this approach can detect the passing light on finger area only, otherwise a very strong light source would have to be used. As with the reflection approach, the NIR (Near infrared) absorption by hemoglobin is used and the veins on the resultant image can be detected. Both these methods can be easily implemented into a biometric system, as the NIR can be used as the sole source of the artificial lighting, and both CCD and complementary metal-oxide-semiconductor (CMOS) sensors, have high sensitivity in the IR range of the electromagnetic

Figure 1.8: Surface vein detection using NIR.

spectrum. As stated earlier, the vein detection can be used solely as a presentation attack detection, or a source of additional unique features.

The other approach is to record the 3D as well as 2D features during enrollment and verification.

Both the transition from 2D to 3D and the multimodal approach can serve as a presentation attack detection, as the data required for construction of presentation attack instrument are both less accessible and make the model reconstruction more complex. To the knowledge of author, no study has demonstrated a 3D hand model or hand geometry with bloodstream model reconstruction from an uncooperative user with accuracy high enough to create a viable presentation attack instrument.

As will be further discussed in the chapter, the transition from 2D to 3D biometrics has been experimentally accomplished and supports the assumption that 3D features will lead to increased entropy of the model.

1.2 Fingerprint recognition

Fingerprint recognition is perhaps the most widespread biometric technology currently on the market. We can trace back the field of dactyloscopy i.e. methods of fingerprint identification to 1892 where the methods and science behind the fingerprint identification is established by Francis Galton in his work Fingerprints. His approach begins to be utilized from 1900. Majority of the mobile phones, access systems even time logging systems, use the fingerprint to unlock, identify or to verify the user. However, since this technology has been in development for the last two centuries, there are many approaches to both the scanning and the matching both experiencing tremendous advances. The underlying principles have remained similar.

Fingerprint is essentially a set of friction ridges; these friction ridges are formed by elevated portions we call valleys, and depressed called ridges. It is these surface distortions that are being analyzed and acquisition systems aim to create an accurate scan of this

profile using its designated technology. Friction ridges may start, end, diverge and join during its length, we call these ridge characteristics minutiae. Type and position of these minutiae are one type of features, that can be used for the recognition.

So far, only the fingerprint minutiae have been mentioned, however other features can be extracted. Generally other extractable features have to do with friction ridges themselves, rather than solely minutiae. Alternatively, with scanner of sufficiently high resolution, the sweat pores on top of the individual friction ridges can be scanned. This allows for extraction of additional features that can be used for recognition.

In this section we will focus on the feature extractor module, and the matching module, as some of their specifics, are unique for biometry. We will explore sensor and feature extraction modules specifically in relation to the intended biometrics. The description of the matching module will be focusing on the statistical measurements and other classification metrics, used when describing the performance of the biometric system. In other words, in this section we will assume that the data is already available and we need only to process it. The description of the actual technologies employed to collect this data will be explored in subsequent section focused on acquisition as regardless of the underlying technology, output is generally a matrix of data, that describes a corresponding area of finger's skin in a manner consistent with the technology used. For optical sensors we would therefore expect a 2D image formed of light intensity values, which will vary depending on skin deformations. While for capacitive sensor a 2D matrix of capacity values, or binarized data determined from these values, depending on height differences of human skin caused by friction valleys and ridges. In this chapter, we will assume that the output data from all sensors will be an image, and image processing techniques can be applied. As this thesis focuses on the low-level processing rather than the high-level biometric recognition method. This chapter will elaborate in more detail into the methods of processing the raw image as well as the fingerprint enhancement and will describe the identification processes only in the sense of technological overview, since various libraries have been used to gauge the quality of the developed system, but were not in fact custom developed for the purpose of this thesis.

Regardless, in order to collect the data needed for the biometric recognition, the surface of the epidermis, or dermis needs to be scanned at sufficient resolution. This resolution typically ranges from 250-1,000 DPI, conventionally 500 DPI is a standard for fingerprint minutiae extraction and 1,000 DPI for sweat pores extraction. There are multiple technologies that are being used for this task, we will generally within this thesis focus on the two most widespread: capacitive and optical.

1.2.1 Image preprocessing and fingerprint enhancement

The main role of the image preprocessing during the fingerprint enhancement is to act as an effective filter that increases the contrast between the foreground ridges and the background and reducing noise. While the methods have improved since 1988, there are lots of similarities. The most cited and at the same time still very used approach was introduced by Hong [45], an approach which uses normalization, combination of orientation field a frequency map and Gabor filters to enhance the fingerprints. Similar in architecture are the works by Chikkerur [14] and Charkaborty [10]. Notable alternative approach is the use of convolutional neural network (CNN) as done, for example in the work by Li [62].

In theory chapter, we will focus mostly on the Hong's work, as in this thesis a modified version of Hong's approach is being used. The modified approach has been used used and published in a paper by the thesis author [35]*. The different approach had to be

Figure 1.9: Example of effect of enhancement on on the fly acquired finger images [35]*.

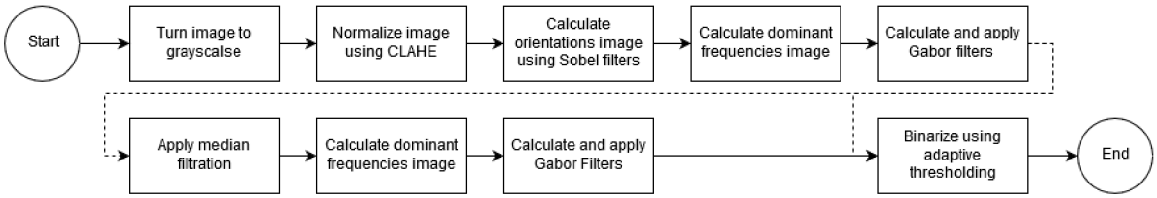

Figure 1.10: The block diagram of the enhancement process. Adapted from [35]*.

used for the normalization stage, as the on the fly approach does not offer an even global illumination, thus the global normalization does not offer the intended effect. As stated in the paper:

„However, we modified this method to suit our system to deal with the effect of direct light source and resultant shadows. To eliminate the illumination variance due to the depth of the object, we used local adaptive normalization, CLAHE (Contrast Limited Adaptive Histogram Equalization) [78], with block size of 8×8, while in [45] they used global normalization. Image orientation is used to compute the Gabor filters of size 7×7 with Gaussian envelope having a standard deviation of 4.“ [35]*

1.3 Other hand biometrics recognition

Due to its availability, hand has been subjected to intensive research. It is therefore understandable, that many other biometrics have been explored in relation to the human hand. Some of these have presented an interesting alternative to identification technologies. While it is fair to say that no other biometry has achieved such a prominence as fingerprint recognition or hand geometry recognition, it is still worthwhile to investigate what other biometrics are there present at or in the human hand.

In this sub section, we will focus on the systems related to the palm print recognition the thermal hand imaging recognition and the knuckle recognition as the biometrics, with the most research done. While the focus will be primarily on the technologies involved in the aforementioned systems, we will also look at the methods employed for the purposes of recognition.

Palmprint recognition

Palmprint recognition has a significant presence in the biometry research area. Especially due to its being instrumental in the forensic sciences. There are two general approaches to the palmprint recognition, using high-resolution images such as [88] and [27], and the low-resolution images such as [43] and [111]. The main difference between the two approaches being, that creases in the low resolution recognition are the primary source of data, as opposed to the high resolution data recognition, which utilizes the ridge characteristics and the creases are in a source of noise, that needs to be filtered. Survey going into more details has been done in [56]. For the purposes of this thesis, it is sufficient to know, that the high resolution palmprint recognition uses many of the same algorithms as the fingerprint recognition.

Due to the shape of the palm, most of the acquisition technologies now are optical and differ mainly in the way the hand is being scanned, with the contact based being the most common one such as presented in [107], where a hand is placed on the glass surface and a CCD sensor acquires the palm image through the surface. Alternatively a contactless approach has been also investigated, for example [21].

Thermal imaging recognition

To the knowledge of the author, thermal imaging hand recognition is a not very widespread recognition method. While it has been demonstrated, it is indeed viable for example in [100]. It is fair to say that it is simply an extension of the hand geometry-based recognition with different imaging approach. Where instead of using a visible spectrum a near infrared spectrum is used.

The area where the thermal based imaging has gained prominence, is in the hand tracking and gesture recognition. Especially now, due to the decrease in price of the thermal imaging cameras, this technology has been explored more than before. In the publication [114] and [54] we can see that the usage of the thermal based imaging makes the segmentation stage of the image processing much more trivial, but also that the subsequent image analysis utilizes the same methods as the classical visible spectrum based image processing.

What we can take from this area, is the fact, that the appropriate technological setup of the capture system, enables us to developed a less performance demanding system, then an unoptimized hardware setup would. In this case a different spectrum has been used for improved segmentation stage. As we will learn in the subsequent chapters of this thesis a similar approach has been utilized to develop a capturing setup in a way that gives us de facto preprocessed data (to a degree) and the subsequent processing is less performance intensive than it otherwise might have been.

To serve as a lead to the the next subsection, it is fair to point out that thermal imaging offers an additional benefit, that the visible light imaging does not. In this particular case, it gives us an opportunity to use the data as a basis for the presentation attack detection, in other words to ensure, that the device is not presented with a presentation attack instrument or a fake human hand analog, to try and bypass the system. As most of the artificial presentation attack instrument do not include a source of heat and would be therefore easily recognized by the thermal imaging camera.

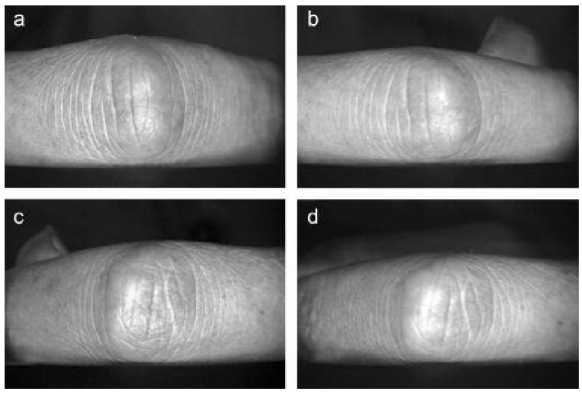

Figure 1.11: Example of four sample knuckle print as shown by Zhang, a,b from user one and c,d from user two [112].

Knuckle recognition

Knuckle recognition or finger knuckle print recognition is an area of biometry, where as the name suggests the identification of a subject is achieved by observing and comparing the creases that form on the outer side of the finger knuckle. It's uniqueness was observed in [108] as a part of overall finger surface biometry investigation and has since been investigated on its own as an interesting alternative to more common fingerprint, face or other biometric recognition. The overall advantage comes from the fact that the creases are of bigger scale and therefore place lower demands on the camera used, in terms of resolution and acquisition time. Substantial amount of work has been done in [112] where use of Gabor filter was proposed, this was later expanded in [113] where local features extracted using Gabor filters were combined with Fourier transform coefficients serving as global features. A high recognition rate is achieved as a result of this combination. See Fig. 1.11 for an example of finger knuckle image used, for the knuckle print recognition. Perhaps due to decreased demand on the sensor system, finger knuckles biometry is on of the few biometries that has been used in an on the fly system. This has been demonstrated in [1] where a touchless non restrained system is presented, capable of performing biometric recognition.

1.4 Watchlist technology and Doddington's zoo

Biometric-enabled watchlist, is one of possible meant of performing a rapid risk assessment in an automated border control (ABC) infrastructure, also known as e-border [59]*. This technology is of particular interest to this paper, as the on the fly technology is a prime candidate for adoption in e-borders, due to its speed and therefore a possible high bandwidth. As elaborated on in the paper [59]*:

„We consider the watchlist as an intersection of advanced technical and scientific highways, taking into consideration of factors such as human rights. This conceptual projection includes the following dimensions:

1. **Inference engine:** Since the data about the persons of interest that is stored in the watchlist is often incomplete, uncertain,imperfect, fragmentary, and conflicting , a powerful inference engine is needed at various phases of gathering,evidence accumulation, control, fusion, recognition, tracking in the surveillance network, as well

as risk assessment and prediction. These inferences can be executed by Bayesian networks and their extensions such as dynamic Bayesian networks, as well as methods for dealing with conflicting information, and deep learning techniques based on neural networks. Technological components, including artificial intelligence and computational intelligence tools, should be developed, implemented, and deployed under fundamental social constraints and regulations.

2. **Social embedding:** The watchlist utilises mechanisms for embedding in social infrastructure such as exploring various databases, local security resources, big data analysis, and possibilities of surveillance networks for identification and tracking.
3. **Interview supporting technology:** Motivated by the fact that the traveller's cooperation is a crucial factor for improving the watchlist performance, the watchlist screening should be integrated into interview supporting machines – a well identified trend for deception detection as applied to e-border infrastructure.
4. **Countermeasures:** The watchlist can be vulnerable to attacks on its integrity. Some attack countermeasures include anti-spoofing techniques and anti-impersonation“ [59]*.

As for the taxonomic criteria of watchlist, following have been defined in [59]*:

- updating mechanisms;
- architecture (usually a distributed network);
- performance metrics (including social impact);
- social embedding (capacity of sources of information);
- watchlist inference (engine, capacity, and control); and
- data protection mechanism.

„It follows from the list above that any watchlist is a dynamical structure, mainly: architecture can be reconfigured as a distributed network of some watchlists, data storage can be changed or deleted via updating mechanisms, and performance varies and depends on the depth of social embedding, credibility of data, reliability of sources, and application constraints. In this paper, we use the term 'risk landscape' to emphasise the dynamic properties of the watchlist.

In e-borders, the problem 'Is this traveller on the watchlist?' is reformulated as 'What is the risk (or cost) of the traveller being wrongly matched/non-matched to the watchlist?'. Conceptually, this means that the watchlist check by a human operator is replaced by watchlist-based inference of risks [59]*“.

The risk calculation is composed of many parameters, some of which will be elaborated on later in the practical part of this thesis, however, one aspect that needs to be introduced in the theoretical part of this thesis, is a metrics that determines whether a person's biometric data is abnormal in a way that changes the likelihood of correct recognition. Whether we are talking about a susceptibility to false positive recognition or false negative recognition. This phenomena has been first described for speech recognition by Doddington et al. in [20], for other biometrics by Yager et al. [109] and another for various biometrics including hand geometry by Lopes et al. [66].

Doddington's Zoo

The category of vulnerabilities, the Doddington's zoo may be used to tackle is called *Impersonation phenomenon*. The phenomenon of impersonation (passive or zero-effort attack) has been experimentally detected in speech recognition by Doddington et al. and in other forms of biometric modalities. Unfortunately, as passive attacks are an inherent property

of watchlist technology, impersonation from other sources, such as social networks, can also influence the reliability of the watchlist [59]*.

Doddington identified in his experiments individuals who performed unusually during recognition in statistically significant number of samples. Based on these observations, Doddington designed categories, into which every user could be classified, based on the type of abnormality his recognition displayed.

Doddington risk categorisation is defined as the five types of classification in the recognition process:

- Category I ('sheep'), recognised normally;
- Category II ('goats'), hard to recognise as themselves;
- Category III ('wolves'), good at impersonating others;
- Category IV ('lambs'), easy to impersonate; and
- Special Category V ('worms'), hard to recognise as themselves while good at impersonating others.

An illustration of this classification can be seen in the Fig. 1.12. This phenomenon is especially important to take into consideration, when dealing with biometrics, whose features are not fully independent. In another words, while for example features in iris recognition are widely considered to be independent on each other. The face recognition, or knuckle recognition, doesn't have this property, i.e. person whose eye distance is above average, is likely to have an above average ear distance. It is this dependency, that increases likelihood of appearance of trend where in some individuals experience different overall recognition rate (both interpersonal and intrapersonal rate) then others.

As the hand geometry, according to the current literature, does exhibit these qualities, it is in the interest of the completeness of this thesis, to include this phenomena and investigate it.

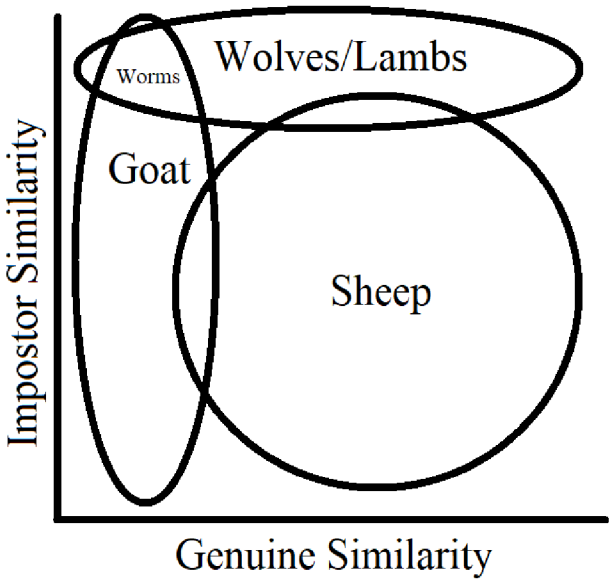

Figure 1.12: Doddington's risk categories and the type of their biometric abnormality.

1.5 Hand geometry acquisition methods

While in the previous chapters we have focused on the overall approach to hand geometry recognition, in this chapter we shall examine the acquisition technologies used for hand geometry based recognition themselves. The acquisition of raw data for the hand geometry recognition can be separated into two main categories: 2D acquisition where a simple projection of the hand is acquired, and 3D acquisition, where the depth information of the scene is collected as well.

The principle of 3D acquisition in biometrics has been successfully tested and researched. However, the price of hardware and high computational demands for processing this data limited the application to academic research. With the advance of low-cost 3D acquisition devices, and increasing computational power of even the mobile devices, the 3D acquisition has become a viable alternative to 2D based systems. For hand geometry, the expansion into 3D offers additional source and type of features that can be extracted. Whereas in 2D based systems, the features location was defined using only two coordinates, the extension into 3D allows for three coordinates description, with serves both as a presentation attack detection and source of increased entropy.

1.5.1 2D acquisition

The Bosphorus Hand Database[4] provided by the Bogazici University [110] used Olympus C-3020 digital camera[5] to collet their database.In [99] on the other hand, a commercial scanner has been used.

[4]Available at: „http://bosphorus.ee.boun.edu.tr/hand/Home.aspx“
[5]https://www.imaging-resource.com/PRODS/C3020/C32A.HTM

In the publications, can be seen, two distinct approaches have been utilized. In [110] the camera is placed above the assumed position of the hand and the dorsal side of the hand is then captured. In [99] the camera is placed below a transparent surface on top of which the hand is placed, this allows for acquisition of the palmar region of the hand. However in both cases, the palm is pressed or placed on a surface for support. This approach can be seen in most of the databases.

Another of the key parameters of the device is, whether the surface includes guiding pegs or not. As can be seen in Fig. 1.2, this setup immensely simplifies the subsequent image processing, as the placement of the hand and individual fingers is essentially fixed. The user is much less likely to commit his hand to a wrong position. This in turn allows to make various assumptions when we begin to process the images, i.e. we can assume where the center of the hand and individual digits are likely to be and are able to simplify the preprocessing stage. Similarly the design of the camera system can also benefit from these constraints, by for example using a camera with lower resolution, as the users' hands physically cannot be outside of expected area, or if they are, an invalid enrollment can be declared.

This is a significant design difference, and data from systems that lack these pegs such as shown in previous chapter by Fig. 1.13, require different and generally more complex processing prior to use for recognition. In other words, the actually hand needs to be localized and correctly identified especially in relation to its position, placement and orientation. As has been described in section 1.1, this puts further demands on the preprocessing stage. It is however fair to say, that the pin-less system is significantly more user friendly and an important milestone when getting to the touchless and on the fly systems.

1.5.2 3D acquisition - Time of flight

In [53] high accuracy 3D scanner is employed – in this research Minolta Vivid 910 has been used for compilation of a large database of right hand scans (3,540 items). Per documentation, the device offers an accuracy up to 0.22 mm in x-axis, 0.16 mm in y-axis and 0.10 mm in z-axis [6].

Research done in [52] performs a detection of four fingers on intensity map, this information is then used for extraction of corresponding data from the depth map. Each fingers is then split into chosen amount of segments along the length of the finger. At these points the features such as curvature and normals along the finger are calculated based on the profile of the cross-section. Fig. 1.14 shows a depth profile of a cross section, calculated curvature and normal features.

In the paper, the 2D and 3D data are experimentally matched and EER (Equal Error Rate) and AUC (Area Under ROC Curve) are calculated.

1.5.3 3D acquisition - Structured light

While any 3D mapping method may be used, in can be clearly observed that in current research, most utilized method is that of active triangulation. Whether the 3D laser digitizers are used for this purpose [52] and [53], which offer superior accuracy, or systems based on structured light projection [93] that are now being mass produced and are therefore affordable.

[6] Available at: http://www.upc.edu/sct/documents_equipament/d_288_id-715.pdf

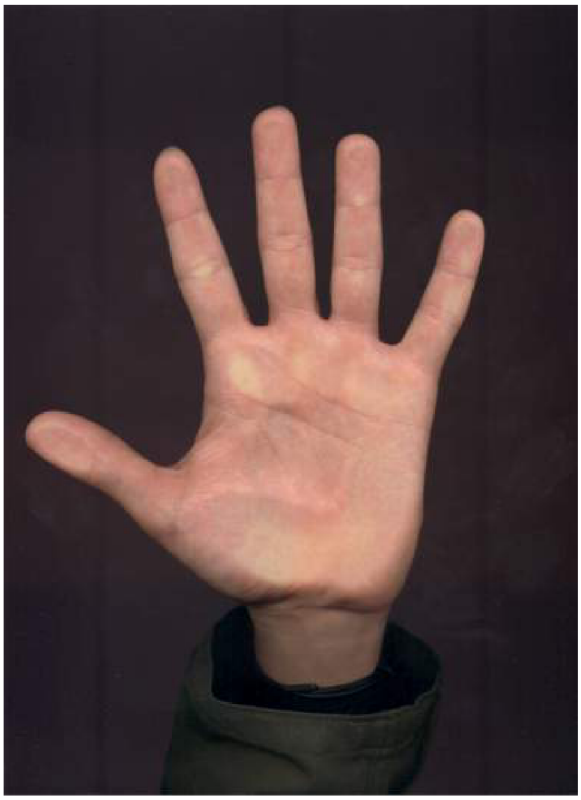

Figure 1.13: En example of pegless system [53]

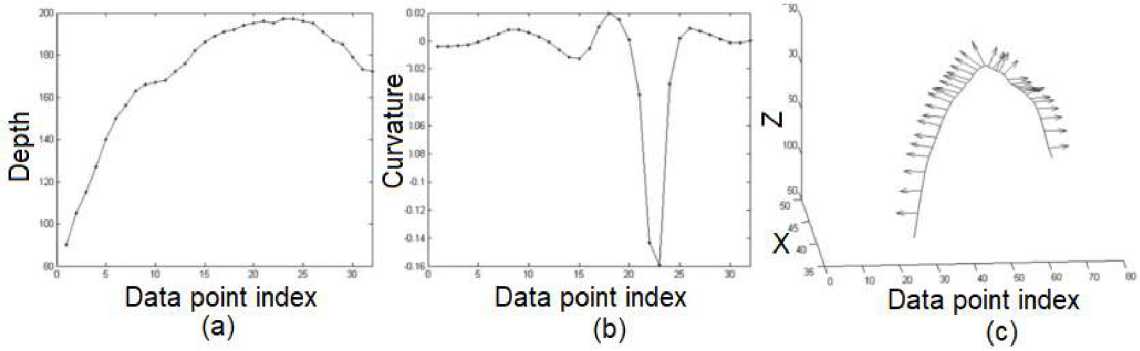

Figure 1.14: (a) Cross-sectional finger segment and (b) its computed curvature features; (c) Normal features computed for a finger segment. Adapted from [53].

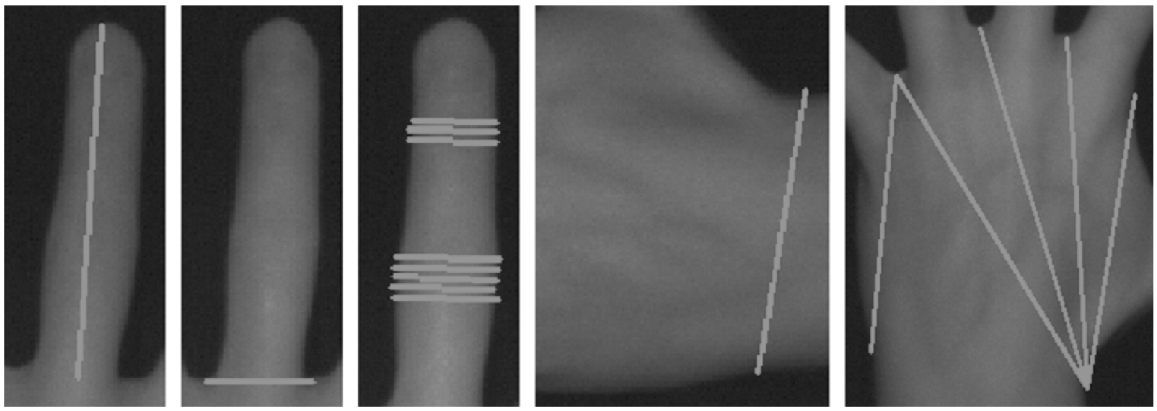

Figure 1.15: Extracted 2D features [93].

When a light pattern of an observable frequency is project onto a deformed surface such as human hand or palm, the pattern itself becomes apparently deformed. By detecting and measuring this deformation relative to the non-deformed pattern a 3D mesh can be reconstructed that serves as a model of observed object. This approach using active triangulation has been used in several publications, varying both in the type of light and the pattern used.

IR pattern projection

Utilizing IR camera and IR projector became viable thanks to the mass production of affordable 3D sensing technologies for purposes of entertainment, such as Microsoft Kinect and Intel RealSense. Both aforementioned devices include not only the depth sensor, but also standard RGB (Red, Green, Blue) imagining sensor, with software libraries, that map the coordinates of one onto the other.

In [93] 2D and 3D features of human hand are extracted using Intel RealSense camera. After normalizing the image during preprocessing, minutiae such as fingertips, finger valleys and wrist lines are located. Using these minutiae feature vectors containing 41 2D features and 137 3D features are constructed.

- Finger length (2D);
- Finger valley distance (2D);
- Finger width (2D);
- Wrist to valley distance (2D);
- Finger axis surface distance (3D);
- Finger width (3D).

Fig. 1.15 and Fig. 1.16 show the respective features on the depth images and intensity images.

The matching of the feature vectors is performed using Mahalanobis distance [68], where the weights for individual features are calculated using the large margin nearest neighbors method (LMNN) [103].

In the work by Svoboda [93], 2D and 3D feature vectors are calculated separately. This allowed the author to compare both the performance of the matching algorithm using 2D and 3D feature vectors separately, as well as the algorithm which combines the 2D and 3D data and uses the combined feature vector for the matching.

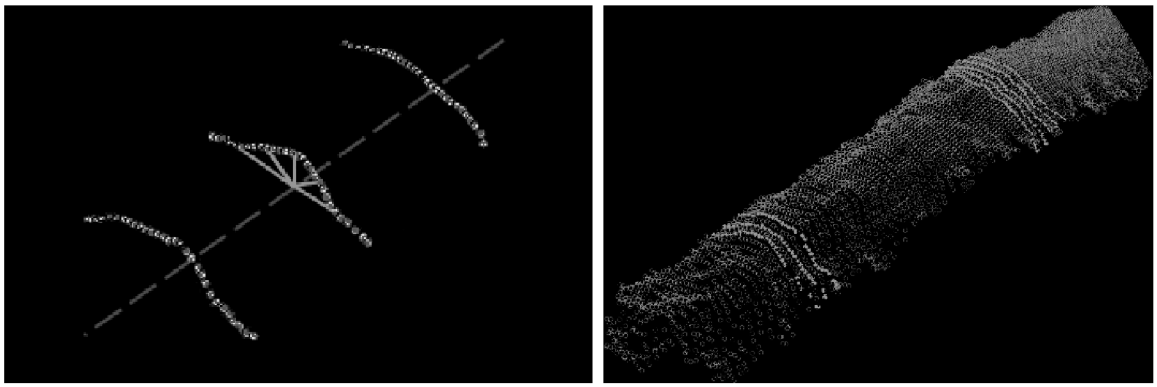

Figure 1.16: Extracted 3D features [93]

Unlike in the experiment with the high precision 3D scanner, the performance of only 3D features in recognition algorithm is inferior to the 2D hand based geometry biometric. However, the performance of combined features, is superior to a 2D only approach by a large margin, despite using the low-cost technology. It is also worth noting, that the performance is significantly higher, when one source of features is burdened by heavy noise, as one technology can be immune to source of noise, that the other is burdened by, i.e. glare in RGB sensor does not influence the depth mapping. Overall the combination of 2D and 3D feature vectors in the algorithm has demonstrated performance comparable to the state of the art approaches.

Lasers and diffraction grating

While the utilization of commercial low-cost cameras has demonstrated acceptable results, the fact that these devices are designed to operate at wider range of distances and greater depth of field means, that the principle can be optimized for the specific application. Since we also have the information about the range of colors we can expect from human hand, a specific wavelength can be chosen to better match the intended application.

In [94] the array of 532 nm lasers is being used, with optics that changes the dot projector into the line projector thus producing number of parallel lines. For the image acquisition two MS Lifecam HD 3000 camcorders are used. From the acquired images, the model of hand can be reconstructed. Fig. 1.17 shows the process of segmentation and feature extraction during the experiment. Fig. 1.18 presents a model reconstructed from the acquired data, demonstrating the concept.

The paper serves as the proof of concept and viability study of the method for the purposes of 3D hand geometry scanning, as such, database has not been created and method cannot be compared with other available approaches. The accuracy of method has been demonstrated by scanning calibrated object of known dimensions. For this purpose, a cube with a side of length 25.5 mm has been used and the root mean squared, and normalized root mean squared have been calculated.

1.5.4 3D acquisition - Stereo vision

Stereo vision is a process where, by combining information from two stationary cameras, we can determine a depth in an image via the method of passive triangulation. The passive

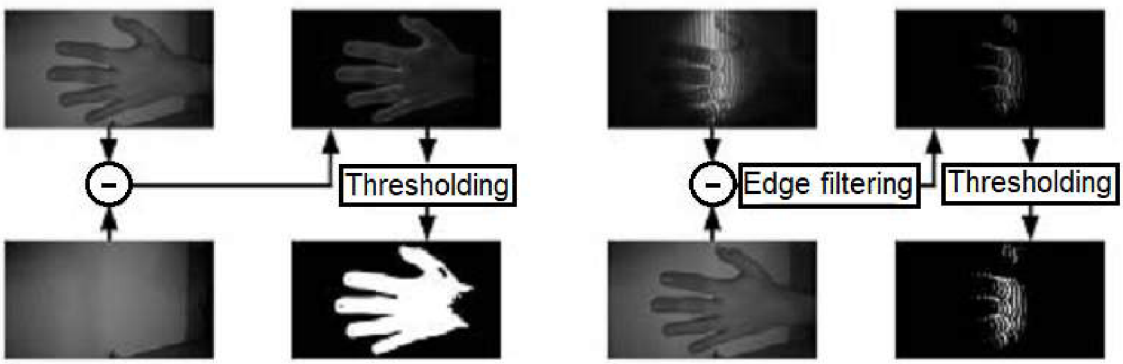

Figure 1.17: The process of extracting the deformation markers, created by diffracting line projection. Adapted from [94].

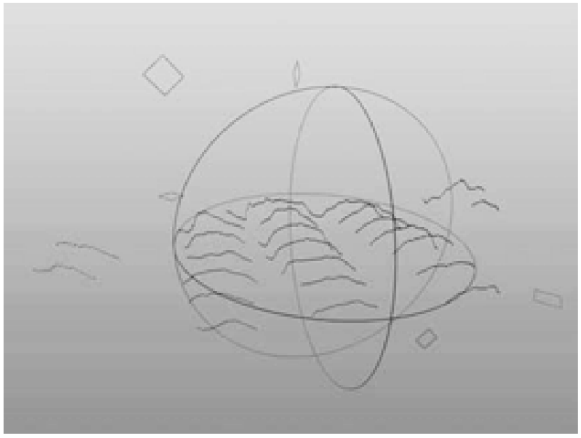

Figure 1.18: Resultant surface reconstruction from the lines projection [94].

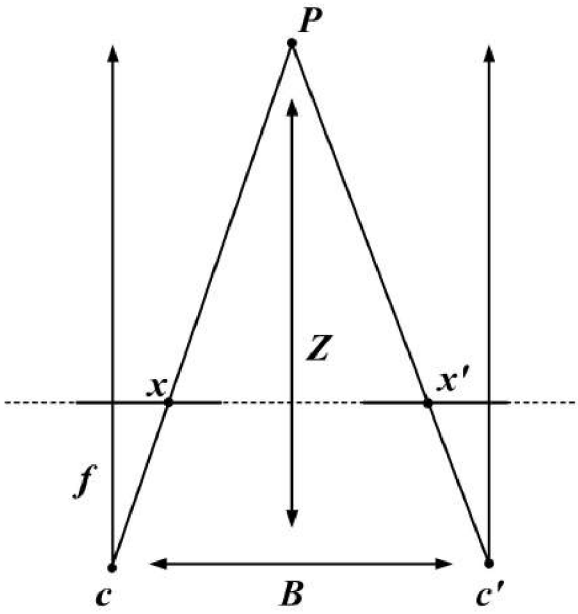

Figure 1.19: Based on parallax and known position of cameras c a c' distance Z can be determined.

triangulation uses the position difference of significant points on the pair of images, to determine the distance to sensors. If the resolution is high enough and therefore the number of point pairs is sufficient, a shape of the observed object, in our case the human hand, can be reconstructed.

On Fig. 1.20 basic principle of stereo vision can be observed. Points x and x' represent the real world coordinates projected on an image captured by its respective camera. Distance Z is then calculated using the known position and orientation of cameras and disparity d that represents the projected position difference caused by distance from observed point to camera. Following equations are used to calculate the disparity d and the corresponding distance to observed point.

$$d = x - x^{'} = f\frac{B}{Z} \tag{1.5}$$

$$Z = f\frac{B}{d} \tag{1.6}$$

Identification of points whose disparity will be used to determine the depth is a part of the so called correspondence problem [84], while the reconstruction of 3D model from these points is the part of the reconstruction problem [51].

At the moment, there are multiple commercially available 3D cameras, using the stereoscopy principle, with resolution high enough for the biometric purposes. For example, ENSENSO N10-304-18 can per documentation achieve the resolution in Z axis of up to 0.109 mm[7]. Unlike TOF based technology, the acquisition time can be as low as a time to capture one image, and under proper conditions, can offer higher precision, than active triangulation via the subpixel methods.

[7]Available at: https://en.ids-imaging.com/ensenso-n10.html

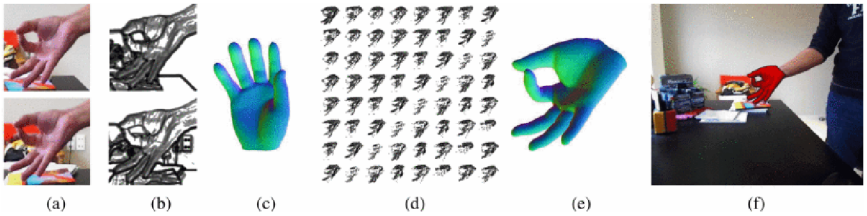

Figure 1.20: Example of stereo vision being used for hand related 3D information extraction. (a) The stereo pair input (b) the computed distinctiveness maps (cropped around the last hand position) (c) the skinned mesh of the employed hand model (d) the color consistency scores after the evaluation of 64 PSO particles for the last generation of the optimization process (the top scoring hypothesis is the top left), (e) the hand model configured in the highest scoring pose and, (f) the solution in (e) rendered and superimposed on the left image of the stereo pair[73].

The largest source of error for this technology is a lack of detectable significant point pairs on homogeneous or smooth surfaces, this however does not apply to human hand, as the papillary lines themselves, offer the necessary irregularities for significant point pair identification.

1.6 Fingerprint acquisition methods

As with the hand geometry in the preceding technology, in this subsection we shall introduce technological approaches to the problem of collecting biometry data, i.e. in this section we will focus on the description of the remaining sensor module of the fingerprint recognition system.

1.6.1 Capacitive scanners

Capacitive scanners are based on the technology of capacitive coupling to distinguish between the friction ridges and valleys along the scanned surface. The capacitive scanner is composed of matrix or a grid of electrodes. Regardless of the exact technology used, the electrodes, along with the glass cover that serves both as a dielectric and a protection measure the change in the electrostatic properties, that occur when the finger is placed on the surface. As the friction ridge will come into the direct contact with the glass surface, whereas the valley will be a nonzero distance away from this surface, difference will be measured, which can then be used to reconstruct a scan of the fingerprint.

This method of data collection is fact, as the change in the capacitance is, for practical purposes instantaneous and requires therefore no exposition time like the optical sensors. By the process of the operation, it also serves as a very rudimentary presentation attack instrument detector, since in order to bypass this sensor, the presentation attack instrument needs to have same surface distortions as a real fingerprint. However, major disadvantage of the technology is, that because of its fundamental principle, it needs to be essentially touched based and it is not possible to perform a touchless acquisition on any meaningful distance. It is fair to note that the finger does not need to actually come into contact with the sensor, and could be used as contactless, however, there is an exponential relationship between the capacitance and the distance between electrodes of the capacitor, hence surface

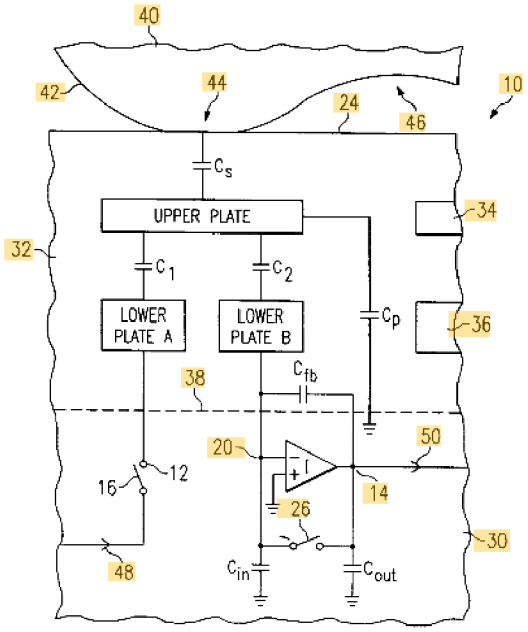

Figure 1.21: A schematic circuit diagram illustrating the invention [41].

reconstruction of the finger that is not in contact with the sensor (which leads to friction ridges being pressed against the surface and friction valleys being non zero distance from the surface) becomes impractical.

The capacitive based sensors are currently one of the most widespread technologies used in large number of mobile phones and as such is being developed by many companies. The basic principle of "pixel" acquisition in capacitive sensors is very well described in patent [41]. For our purposes, it is enough to understand, that the pixel measures the capacitance of the sensing capacitor, which changes based on the the nature of the contact with the fingerprint sensing surface. Due to the conductive nature of the human skin, the conductance mainly depends on the whether the capacitor is under the the ridge, i.e the skin is in contact with the sensing surface, or the valley, i.e. there is an air gap between the finger and the sensing surface. Performing a capacity measurement for each individual pixel along the sensor matrix, it allows us to reconstruct an image based on the decided threshold capacity. A simplified schematics can be seen in Fig. 1.21, where 44 is a ridge, 46 a valley 24 a sensing surface and C_s a sensing capacitor, the rest is out of the scope of this thesis and can be looked up in the reference patent.

There are many producers of the capacitive sensors such as Fingerprint Cards AB[8], Shenzhen CAMA Biometrics Co., Ltd[9] or Biometrics ASA[10] to name a few, and while parameters such as the sensing area, resolution and even circuitry of individual sensing elements may vary, and lead to various degrees of performance, there is a limit to modifications we can do in terms of requirement for touching the sensing area. To the knowledge of the author a touchless capacitive based sensor has not been produced, or even researched,

[8] www.fingerprints.com/technology/hardware/
[9] https://www.camabio.com/en/capacitive-fingerprint-module.html
[10] https://www.idexbiometrics.com/products/biometric-fingerprint-sensors/

let alone a way to expand this approach in the direction of on the fly acquisition. In order to achieve this goal, an alternative technology needs to be used, one based on optical approach.

1.6.2 Optical scanners

Optical sensors use charged coupled device (CCD) or active pixel sensor (APS) (often colloquially known as complementary metal-oxide semiconductor image sensors (CMOS)) image sensor to acquire the information necessary for the fingerprint identification. As the difference between the CCD and APS sensor is mostly unimportant for the purposes of this thesis, we will introduce them only briefly.

CCD technology was first described in a paper in 1970, where it was presented as a mean of memory storage a delay circuit or an imaging device. For the imaging purposes, a rows of capacitor arrays, are exposed to a lens directed image, the photoactive layer then enables conversion of incoming photons into the electric charge proportional to the level of incoming light. Using control circuitry, the array of pixels in a manner similar to shift register, is being gradually moved towards the measurement circuitry, where the light value of individual pixels can be measured. Repeating this process for all the columns then gives us the information necessary for the image reconstruction.

APS was first proposed in 1968, instead of using a capacitor like structure, to measure the light level, by measuring the charge like CCD. The APS image sensor uses a photodiode and a number of transistors to convert the light level to voltage which is then used for the image reconstruction. Today we most commonly encounter a setup, where a pinned photodiode is used as the active component and a circuit of 3 (3T), 5(5T) or 6(6T) complementary MOS transistors (CMOS) is used.

Regardless of image sensor technology used, what we are interested in is an image of a skin surface of such quality, that the friction lines structure can be recognized from it. In order to achieve that, the image has to meet several parameters to be usable for biometric identification.

As with capacitive sensors, the resolution is an important parameter, for optical sensor, we also should clarify, that we are interested in the spatial resolution, rather than the absolute resolution of the imagining sensor. In other words, the resolution parameter for scanner technology needs to be in relation to the unit of area or length. By custom, the biometrics technology utilizes a DPI parameter which, per intuition, represents a number of pixels, that will be used in a length of one inch, or more clearly spatial resolution of 500 DPI will represent the scanned area of 1 x 1 inch with 500 x 500 pixels on the sensor.

Resolution of the fingerprint sensor determines, what kind of fingerprint detail we can achieve using our system. While the algorithms used in this stage will be described later in this chapter, here we will only state, that based on the level of detail we require we will need to use certain minimal resolution. Jain in [47] shows, that different resolution allows us to extract different level of the detail and a group of features from the fingerprint. Level 1 details for example are extractable with 380 DPI resolution, Level 2 details will be confidently extracted from a 500 DPI, Level 3 detail requires 1000 DPI for an extraction. Publications [47] and [47] discuss how increase in resolution can improve on recognition rate even when using the level of detail that conventionally requires lower resolution.

The other important compound parameter, that directly relates to the quality of image for biometric purposes, is the contrast. As the contrast is not a directly designable parameter, we generally need to consider construction design and the light design to achieve our

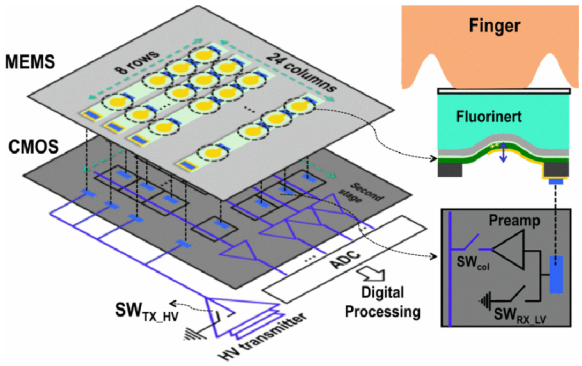

Figure 1.22: A schematic detail of the ultra sound fingerprint scanner [96].

desired image quality. In ideal circumstances, the system should be designed in a way, where friction ridges and friction valleys are represented in binary manner (1/0, 0/255). To get as close to this goal as possible, various techniques have been developed to accommodate it. Most notable and apparent is the use of frustrated total internal reflection (FTIR) [13][91], in this thesis we will consider a FTIR an event, where due to the presence of the ridge on the sensing surface, the total internal reflection is frustrated (doesn't occur) due change in refractive index of the border region, passes through and is absorbed by the skin. Since the areas of sensing surface with valley above them, will allow the total internal reflection, it will appear bright on the result image. Due to the clear contrast between ridges and valleys, this approach is widespread in the commercial applications.

1.6.3 Other scanners

To make the introduction of scanning technologies complete, we will finish the listing with ultrasonic and thermal imaging.

Ultrasonic fingerprint recognition is a novel approach utilized in biometric recognition [67][95][48]. It has however received a fair amount of attention, due to its being applicable in the smartphone technologies. Its main advantage being, that it can be placed under the display of the smart phone. Thus enabling a larger portion of the front side of the smartphone to be used for display only.

As can be seen on the image 1.22, the design utilizes an direct piezoelectric effect in order to generate an ultrasonic pulse, which one reflected from the epidermis present on the finger of the human skin, depending on whether a valley or ridge has been present, results in a different reflected wave. This wave is then translated into a charge and voltage utilizing the inverse piezoelectric effect. However, as this technology also requires a physical touch of the user's finger, it cannot be used in a touchless application either.

Use of thermal imaging for biometry application, has a long, but until recentrly not very widespread history [60][85][18]. It is fair to say, that this technology has never achieved a prominence due to its relative complexity, and inability to offer sufficient resolution until very recently [18]. While the inherent suitability to be used as a presentation attack detection technology, (as the scanned object must have a human-like temperature to be even scanned) as is, the technology is mentioned here simply for the purpose of completeness of this subject area.

This concludes an overview of technologies used for the purposes of the biometric recognition. While they all come with some inherent advantages that makes them suitable for various application. Only one comes out as suitable for the contactless recognition, optical, and even then it requires a departure from the FTIR principle which they use. However, as it is the one that allows an operation without the necessity of the touch, that is the technology this thesis will focus on from here.

1.6.4 Presentation attack detection

Presentation attack detection or anti-spoofing as known earlier is an area of research and a process which aims to ensure, that an intentional breach of biometric system via a body part analogue (an artificial fingerprint, hand ...) works[22]* and [44]* introduce this problem and available countermeasures. As the ability to fool biometric system in a simple way using an artificial object is unwanted and decreases a security value of a system vulnerable in this way. This thesis deals with this issue in relation to system development as well.

As stated in [23] number of items may be placed on the biometric scanner in an attempt to bypass it, such as:

1. Registered (enrolled) finger;
2. Unregistered finger (impostor's finger);
3. Severed fingertip of enrolled finger;
4. Genetic clone of enrolled finger;
5. Artificial clone of enrolled finger;
6. Others.

While all of these can be considered a presentation attack detection, in this thesis presentation attack detection will mostly consider the prevention of attacks by 3 and 5 as it deals with the idea, that maliciously acquired object is being used to bypass system and the one that can be prevented by introducing a presentation attack detection and/or processing the data available from the implicit sensor. The scenario 2 will be further discussed in the subsection on Doddingotn's zoo on it's own.

1.6.5 Touch based, contactless and on the fly recognition

While contactless and on the fly may seem similar perhaps even identical. In subsection 1.1.1 the methods of acquisition and recognition have been introduced in regard to the hand geometry. It is however essential to recognize that the publications mentioned are build on the touch based or simply touchless approach and even that in only two cases.

To the best of our knowledge, Svoboda [93] and Kanhangad [52] are the only touchless approaches. Svoboda utilised an Intel RealSense 3D camera to collect the data, while Kanhangad used a Minolta Vivid 910 Noncontact 3D Digitizer. Both works utilised their respective sensors' depth scanning properties; neither, however, is capable of on the fly acquisition. According to our recent survey, we could not find any work using on the fly hand biometric recognition except for the finger vein recognition by Kuzu [57]. In this technique, more user restrictive constraints were imposed during the acquisition process than with our approach [31]*.

For the purposes of the this thesis, the touch based technology, touchless and on the fly shall be defined thusly.

- **Touch based approach**, demands of user to come in contact with the device. Be it in order to stabilize the hand during the process of acquisition, or to utilize total internal refraction in order to improve the contrast of the placed hand during the

acquisition i.e. we consider system touch based, if the device cannot operate, or operates at significantly decreased performance, unless the user engages in a physical contact with the device.

- **Touchless** approach as the name suggest does not require the physical contact with the device during the acquisition phase of the recognition. We do not address the timing, or the circumstances when classifying system as such i.e. we consider the system touchless if the user is not required to come into the physical contact with the device during the recognition process without negatively impacting the performance of he system.
- **On the fly** approach is usually a subset of the touchless approach witch added constraints. In the on the fly system, we require that the device performs the recognition phases without requiring the user to come to a standstill. While it might be possible to develop a touch based on the fly device, both in the available literature, and in this thesis, the touchlessness is an implied part of the on the fly capabilities. Constraint, that is however not imposed on this device, is a cooperation factor. We assume a cooperating user will interact with the system and we do not require the system to perform recognition on the person who doesn't cooperate or is even avoiding the interaction. In this thesis, the users were all briefly instructed on how to interact with the device, but more on that in later chapters.

1.7 Camera related optical principles

In this section a necessary theory of the optical system design as well as the terminology will be presented, as it will be referenced later in the text. During the course of the PhD program, three optical sensing devices have been developed, all requiring a fundamental knowledge of the optical system design, in relation the actual real world application. The focus will be kept on the calculation of the necessary parameters to determine the optical paths and therefore the limits of various dimensions of the developed systems and calculation of necessary to determine the choice of the equipment that will be used in the developed system.

Field of view (FOV), represents a fundamental parameter that will define the dimensions of our system, as it gives us a relationship between the observable scene, **focal length** of rectilinear camera lens and the distance between a scene and the sensor. The angular field of view fov_θ can be derived from the lens formula and is defined by Eq. 1.7. This parameter is often stated by the camera lens manufacturer. The actual dimensions of the observable scene are than described by the field of view fov_w and is directly derivable from the Eq. 1.7. The formula for fov_w can be seen in Eq. 1.8.

$$fov_\theta = 2 \cdot tan^{-1}(\frac{h}{2f}) \tag{1.7}$$

$$fov_w = 2d_o \cdot tan(\frac{fov_\theta}{2}) \tag{1.8}$$

Where d_o is a distance from the lens to the target object also known as object distance, f is the **focal length** of our chosen camera lens and h is the size of the camera sensor. Due to difficulty of properly determining d_o (as the camera lens is actually a set of lenses), we use the fact that the uncertainty in this measure is substantially smaller then the object distance itself, which makes this approximation sufficiently accurate for our purposes. It

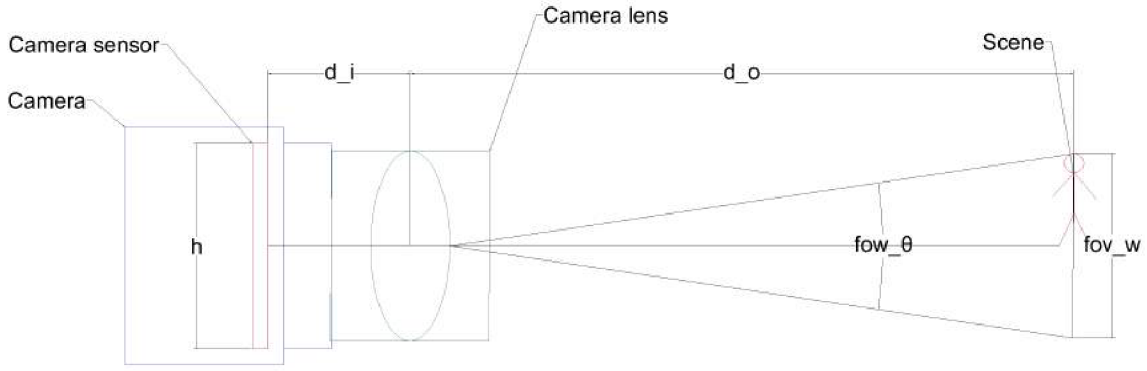

Figure 1.23: Illustration of camera, scene and the related parameters.

is worth mentioning that as most sensors do not have an equal width and height, the fov needs to be calculated for both dimension, or, as is often done, the diagonal fov to offer a better description of the scene.

"Given the distance to the target object and the size of the chosen sensor, the parameters of the camera lens need to be calculated. In this thesis the most significant parameter is without a doubt the z as it is the prime parameter that is directly related to the subsequent calculation of magnification and the depth of field. Both of which we will further elaborate on in this section. To determine the appropriate camera lens, the required field of view fov_w needs to be considered as well as the required magnification m. The **focal distance** [82] f can than be determined using the required object distance d_o, and ratio of sensor size to target size, also known as the magnification m [82] Eq. 1.9 presents the relationship between the **focal length** and the magnification. Since during the design process, we know the object size (the scanned object) and the image size (the size of our sensor) as we generally want an entirety our scene to be projected onto the entirety of the sensor in order to maintain highest possible resolution, we can use Eq. 1.10 to calculate our expected **focal length**.

$$m = \frac{d_i}{d_o} = \frac{h}{fov_w} = \frac{f}{d_o - f} \tag{1.9}$$

$$f = \frac{md_o}{(1+m)} \tag{1.10}$$

To better visualize, the Fig. 1.23 shows the relationship among these parameters. For the purposes of this thesis, we will consider these descriptors sufficient, despite making approximations and using often simplified relations, as the optical system is to be used for the acquisition of real world objects, which compared to the sensor size and camera dimensions can be considered significantly larger, making these approximations sufficient. As said previously, the fov is direction dependent, however, given we utilize a 2D sensor, the formulas above can be used to determine both the parameter for both directions. However as in this thesis we also utilize a line scanner camera, we need to define this relationship for line scanner setup as well.

Line scanner camera is an alternative camera system technology, most often utilised in the industrial application, specifically, in the application, where the scanned object is moving at a constant rate in relation to the system (typically conveyor belts or equivalent).

As mentioned in the biometry introductory section of this thesis, the swipe sensors for the fingerprint acquisition are an example of the line scan technology usage in the biometry field. Line scanner is able to acquire a one dimensional row of pixels at a rate specified by the particular sensor. As the scanned object moves over the sensor a number of these lines can be then used to reconstruct a typical 2D image. As the fov in the direction of the movement is no longer defined solely by the physical dimensions of the sensor. We start of using the Eq. 1.7, where we input the pixel size for the sensor size parameter. After that we introduce a modification to Eq. 1.8 that takes into consideration the movement characteristics of the system.

$$fov_{ls} = v_{ls}t + fov_w \tag{1.11}$$

Where the v_{ls} stands for the movement speed of the scanned object and t the total acquisition time. This formula describes the relationship between a dimension of the scene, parameters of our camera sensor and the distance traveled by the the moving object. This formula however does not express the amount of information we can actually resolve from this system. To determine this information we need to introduce the concepts of resolvability and the spatial resolution.

Resolvability of the system in this thesis will express the length or a dimension of a smallest detail, we will be able to distinguish from another one based on the parameters of the system. Spatial resolution, expresses a number of individual image elements (pixels), per unit of distance of the scene as seen in the acquired image. It is an expression of the level of detail we can expect from an image. This parameter is very important in biometry, as we could have seen in the biometry introduction, where for example the fingerprint acquisition systems tend to acquire images at 500 DPI spatial resolution, for sufficient level of detail.

The size of the smallest resolvable object w_r which can be distinguished in an image, can be determined from the absolute resolution of the used camera sensor r_s and the fov_w using the Eq. 1.12. The factor of 2 in the numerator ensures NS sampling theorem is met as the sampling rate needs to be at least a double that of the required detail.

$$w_{r1} = \frac{2 \cdot fov_w}{r_s} \tag{1.12}$$

The formula above is usable when we need to determine the resolvability in the sensor axis. Should we need to determine the resolvability in the movement axis in the case of the linear scanner system, we will need to factor in not only the speed of the axis, but the acquisition rate as well. The relationship between the resolvability limit and the camera parameters can be seen in the Eq. 1.13. Where the f_r is the acquisition rate of the camera used. As can be expected, higher the frame rate, the smaller is the size of the smallest resolvable system.

$$w_{r2} = \frac{2 \cdot v_s}{f_r} \tag{1.13}$$

For the completeness of this section, we will also include the relationship between the spatial resolution R, the parameters of the observed scene and the parameters of the camera both in terms of the camera sensor axis and linear scanner axis. The spatial resolution will be expressed in DPI as per convention in biometry. For this formula to be correct, fov_wneeds to be expressed in **mm** or the factor needs to be modified accordingly.

$$R_s = \frac{r_s}{fov_w} \cdot 25.4\ [dpi] \tag{1.14}$$

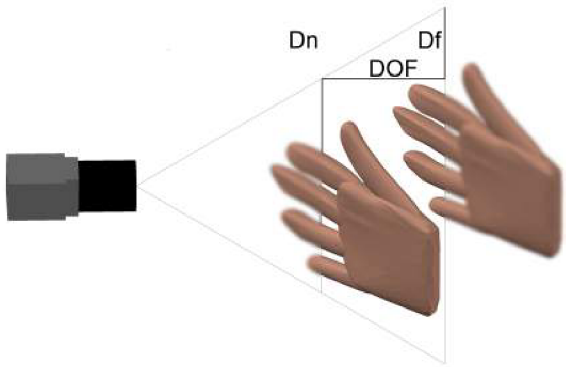

Figure 1.24: Visual representation of depth of field.

$$R_{ls} = \frac{f_r t}{fov_{ls}} \cdot 25.4\ [dpi] \tag{1.15}$$

Using these equations, we can determine the spatial resolution for both the standard and line scanner technology based systems. Given the requirements on the image quality, determined by the proposed application, this allows us to determine a large number of the parameters the equipment must meet. However, until now, we have disregarded a depth of the scanned object and assumed that the object will be in focus regardless of its orientation or shape, this assumption however, cannot be made should we require high level of detail, for example detail which is required for the on the fly acquisition. To be able to take the depth of the image in consideration, we need to introduce one more concept. The depth of field.

The **depth of field** (DOF) is the measure of the rage from the sensor, where per the definition of "in focus" we will expect the image to be in focus. See Fig. 1.24 for a graphical representation of the depth of field. In this thesis we will use following notation. The D_n will represent the distance from the sensor to the distance **from** which we consider the image to be in focus. The D_f will represent the distance from the sensor to the distance **to** which we consider the image to be in focus. DOF will then represent the length of an interval where we expect the scene to be in focus.

A reader may argue that the measure of focus is arbitrary and there is no discrete limit that differentiates between „in focus" and „out of focus", it is however important to remember, that since the sensing element in the camera (i.e. the pixel) has a finitely small size, it gives us a measure of maximum focus, a state, beyond which, any additional focusing will not perceptibly improve the image quality. This measure can be represented by the term a **circle of confusion** (COC), which is a maximum acceptable size of the spot of a resolved point in the scene. Given the pixel size, there is a spot size, which will be regarded by the sensing element as a point, and any additional decrease in will not result in a clearer picture. As this measure can be calculated for a given distance, and we know the size of the sensing element of the camera. To determine the DOF for a given system as we can see in Eq. 1.16, we need to determine the near distance D_n as defined in Eq. 1.17 and far distance D_f Eq. 1.18 where per our required COC c we expect the object to be in focus. To calculate these limits, we need to first calculate a hyperfocal distance H Eq. 1.19

$$DOF = D_f - D_n \tag{1.16}$$

$$D_n = \frac{s(H - f)}{H + s - 2f} \tag{1.17}$$

$$D_n = \frac{s(H-f)}{H-s} \tag{1.18}$$

$$H = \frac{f}{Nc} + f \tag{1.19}$$

where N is an f-number (the ratio of the system's focal length to the diameter of the entrance pupil [82]) of our chosen camera lens and s is the focus distance, or the distance from the sensor, to the object. Again, it is fair to point out, that these equations serve as an approximation suitable for our particular application due to the scale of the design and other equations would need to be used should the required precision be high (especially true when designing a much smaller system such as microscope), as is, the systems designed as parts of this thesis are substantial enough that the equations are sufficient.

This measure becomes very important during the development of the on the fly systems (OTF) where the depth at which the acceptable focus is achieved becomes a critical design parameter, as the unrestrained requires increased tolerance to better accommodate the user's biometry acquisition. Fig. 1.24 illustrates this issue, as we can see that neither of the shown hands have all five digits in focus, which would in turn negatively affect the recognition performance, had the system performed the acquisition during the illustrated moments, however if the acquisition was performed while the whole hand was in the focus, we would expect the performance to be better.

1.7.1 Camera calibration

In this subsection we will briefly introduce the process of camera calibration, especially stereo calibration and the related ideas and terminology as it is an approach explored and experimented on in this thesis for the purposes of 3D acquisition.

Camera calibration is process that describes the relationship between the recorded image from the camera, and the real scene, by defining intrinsic and extrinsic parameters. By defining these parameters, we may then undistort the image and perform standardized measurements along the entirety of the image plane. Whereas, prior to calibration, different parts of the image had a different transformation relationship to the captured scene.

Camera calibration, due to the requirement to use the acquired data for precise measurement is a field of research with long history, and is still being actively researched. Here we will mention works by Duane [25] which describes a photogrammetric approach to the calibration Genry [38] where the self calibration process in in introduced especially in relation to the stereo camera setup. More recent photogrammetry method along with introduction to the calibration in general was done in work by Zhang [116] many current calibration algorithms such as those in OpenCV [8] and Matlab [70] are based on this work.

Notation

For the purposes of this thesis, we will use notation as used by Zhang [116]. Any point in the scene we wish to capture can be described in terms of the $[x, y, z]$ world coordinates. By supplying/determining extrinsic camera parameters, $\mathbf{R}$ for rotation and $\mathbf{t}$ for translation, in relation to the world, we can acquire a camera coordinates $[x_c, y_c, z_c]$. By applying intrinsic parameters onto our chosen point, we are able to project this 3D point onto the 2D plane of our sensor, description of the intrinsic parameters of the camera is known as Camera Matrix and can be seen below. Camera matrix includes information about the **focal length** f_x, f_y and optical centers c_x, c_y the matrix is used as a transformation matrix.

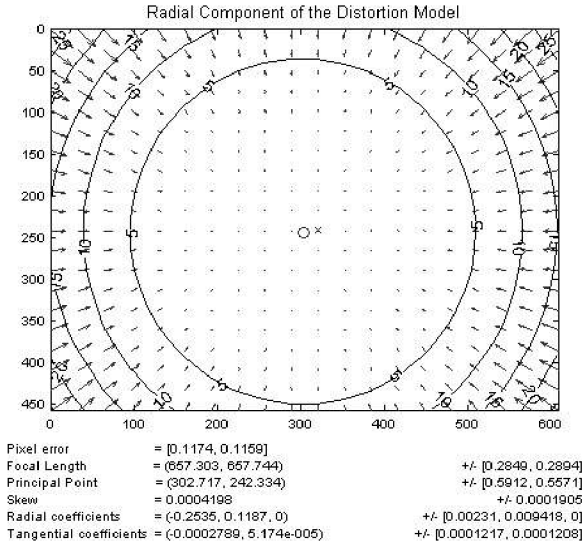

Figure 1.25: Representation of the distortion.

$$CameraMatrix = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{1.20}$$

These intrinsic and extrinsic parameters may be used to describe the relationship between the in-the-world point and the pixel coordinates, it however assumes no distortion due to lens, which is generally an incorrect assumption. To mitigate this, our calibration procedure needs to include calculation of distortion parameters. The most impactful, we will mention here, are tangential distortion, and radial distortion.

The **Radial distortion** is an effect, where the straight lines appear to be curved, especially at the edges of the image. This distortion is caused by the geometry of the lens, where the magnification varies, as we are getting further from the center of the image. Effect is especially pronounced on wide angle lenses and so called „fish eye lenses“. See Fig. 1.25 [115] for a visual representation of the radial distortion. Equations 1.21 and 1.22 represent formal definition of the radial distortion as seen in [116].

$$x_{radDIst} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \tag{1.21}$$

$$y_{radDist} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \tag{1.22}$$

where k_1, k_2 and k_3 are the distortion coefficients, and x and y are normalized image coordinates. The second type of distortion we will consider is the tangential distortion which occurs, when the camera lens is not parallel to the image plane, usually due to faulty construction, it is fair to point out that the components are never perfectly aligned, it is only the measure that differs based on the construction quality. Fig. 1.26 illustrates how this distortion occurs and how it is represented in the image. Equations 1.23 and 1.24 represent formal definition of the tangential distortion.

$$x_{tanDistort} = x + [2p_1 xy + p_2(r^2 + 2x^2)] \tag{1.23}$$

$$y_{tanDistort} = y + [p_1(r^2 + 2y^2) + 2p_2 xy] \tag{1.24}$$

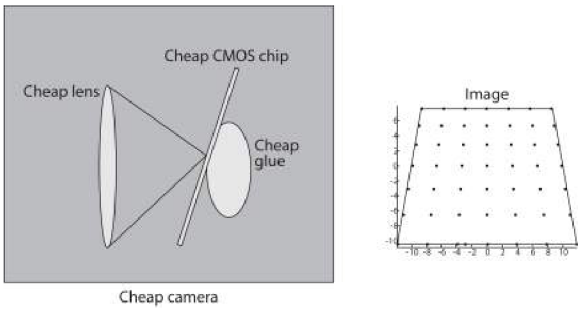

Figure 1.26: Caluse (left) and representation (right) of the distortion [61].

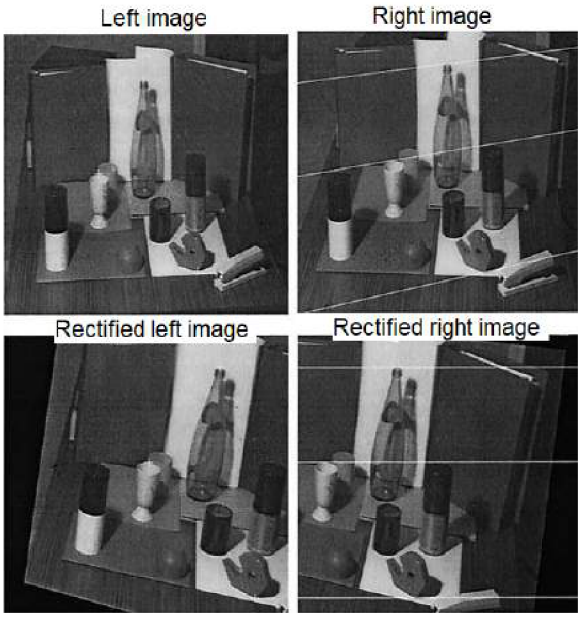

Figure 1.27: Rectification of the stereo images as done in [37].

These distortion coefficients (k_1, k_2, p_1, p_2, k_3) along with the camera matrix allows us to correctly assume the real position of individual point in the image in correspondence to the scene we capture, which may be important when we try to perform measurements for the biometry purposes, but is essential, when we attempt to reconstruct a 3D scene from the image as the image rectification would not succeed if the individual cameras were not calibrated properly.

Image rectification is a transformation of the stereo images in such a way, that the pairs of the conjugate epipolar lines become collinear and parallel to one of the image axes[37]. This transformation then simplifies the correspondence problem into a search in a 1D space, i.e. only along the epipolar line. As seen in the Fig. 1.27, the transformation of the two images allows the simplification of the correspondence problem, as in order for an algorithm to find the corresponding image pair, it needs only to search along the epipolar line and as the distance between two points is directly related to the depth information, we may use it for the subsequent 3D reconstruction.

Chapter 2

Design and realization of innovative biometry enabled devices

2.1 Biometry enabled line scanner

Despite cameras being widespread, decreasingly cheap, and their parameters soaring in the last couple of years, when the application demands highly detailed images of large areas of the human body, it can be quickly discovered, that even the available high-resolution cameras are insufficient given the requirement for spatial resolution and the field of view. It is then necessary to either come up with a way to utilize the less detailed imagery or to find a way to generate a sufficient number of high-quality 2D sub-images that can be reconstructed into the final image. Both methods have their drawbacks. What if, however, instead of using a traditional 2D camera sensor with a fixed number of pixels per area, a sensor with variable field of view was made, and the sub-images were constructed using a device that is capable of linear and constant movement.

In this section a design and a prototype of an optical scanning device, utilizing a line scanner instead of a traditional 2D sensor is presented. A device that can change the field of view based on the operators' needs and thus allowing for scanning body parts of various sizes without having to change the physical configuration of the device, while maintaining the high resolution per unit area. A device that can facilitate scanning of a finger, a hand or a whole arm without having to sacrifice the image quality.

In following sections a possible configuration of the device is demonstrated by capturing images of human hands and acquiring a small database. This decision has been made to allow the use of biometric quality evaluation to establish the performance of the device. Fingerprint quality evaluation algorithms are showing applicability as a biometric-enabled system. Dermatological application is demonstrated by using the acquired images to perform a measurement of common nevi (moles). Further uses in wound treatment and other biometrics such as hand geometry recognition and palmprint recognition are discussed as well [34]*.

2.1.1 Optical design

This subsection has been expanded from the [34]* a publication by the author of the thesis. In order to design an optical system, a basic requirements need to be designed, from which the setup can be proposed. The fundamental design requirements are as follows.

1. The spatial resolution of the image is to be at least 350 DPI.
2. The scanned area is to be at least 400 mm width × 180 mm length where length can be defined by the operator.

The first requirement allows to resolve objects of approximately 0.145 mm, and meets the minimal requirements for the fingerprint acquisition per the introduction in the theory. It ensures that both the biometric applications and the calibrated measurement for a medical imagining application can be performed. Both requirements are used to define linear scanner itself, as the minimum resolution can be derived from them per the Eqn. 2.1 .

$$R_{min} = 350(\frac{400}{25.4}) \approx 5,512 \tag{2.1}$$

Based on this requirement, a line scanner RAL6144-16gm Basler[1] with the resolution of 6144 px has been chosen as suitable, see Tbl. 2.1 for further details. Given the nature of the scanner, the intended use of the camera for industrial application is of added benefit, as it allows us a precise timing, unlike a camera intended for a mass consumption, will not be susceptible to the EM interference during imaging as much as consumer grade camera might if placed next to a stepper motor and power sources necessary in our system. The F-Mount of this particular model also allows us to use a full frame camera lens for better optical properties.

Table 2.1: The parameters of the Basler camera.

Name	raL6144-16gm
Interface	GigE
Chromaticity	Mono
Line rate	17 kHz
Shutter	Global

However, the previous calculation gives us only the necessary resolution in one axis, the sensor axis. To determine the resolution in the movement axis, we need to perform additional calculations, as per the design a line scan camera acquires only a single line at a time and we need a number of those to form an image. To reconstruct the complete image of the target, either the camera or the target needs to move along the movement axis at a predefined speed v_s and the acquisition must be performed at a rate of f which is dependant on the minimum resolvable detail w_r. The relationship between these parameters is defined by Equations 2.2 and 2.3.

$$f = \frac{1}{\tau} \tag{2.2}$$

$$\tau = \frac{1}{2}\frac{w_{r1}}{v_s} \tag{2.3}$$

where f [Hz] is the rate of acquisition in lines per second, τ [s] is the period between acquisitions, w_r[mm] is the resolvable width which is required and v_s [mm/s] is the speed

[1] Available at: www.baslerweb.com/en/products/cameras/line-scan-cameras/racer/ral6144-16gm/

of the system. The factor of $\frac{1}{2}$ is necessary to meet the Nyquist–Shannon (NS) sampling theorem [87] due to the movement of the system. Based on the minimum resolvable detail we demand, i.e. 0.145 and the designed speed of our device, the minimum frame rate of the line scan camera can be determined. Given the experimental nature of this device, and no lower bound for speed, at this point we can consider the requirement to be met by any device as long as the time necessary to perform an acquisition of necessary number of lines is given.

To find a suitable camera lens we need to determine the required focal length.

The focal distance f can be calculated using the required distance to the object, also known as object distance g, and ratio of sensor size to target size, also known as the magnification m as can be seen in Eqn. 2.4 [82].

$$f = \frac{mg}{(1+m)} \tag{2.4}$$

The size of the smallest resolvable object w_{r2} which can be scanned in the sensor axis can be determined from the resolution of used sensor r_s and the fov_{w2} using the Eqn. 2.5.

$$w_{r2} = 2\frac{fov_{w2}}{r_s} \tag{2.5}$$

where fov_{w2} is defined as a part of our requirements.

The object distance of the device is 535 mm, as further discussed in subsection 2.1.2 Mechanical design. Given the dimensions of the chosen sensor and the target area dimensions, the magnification needs to be at least 0.096 to guarantee the 350 DPI requirement. Using the Eqn. 2.4, the focal length can be calculated to be 47 mm. 50 mm camera lens has been used due to its availability. This guarantees a resolution of 350 DPI for the area of 417.1 mm. Camera lense AF Nikkor 50mm f/1.8D has been used in the device. The Chromasens Corona II type C[2] with LED-control unit XLC4-1[3] has been used for illumination. The Tbl. 2.2 includes the selected specification parameters of the selected illumination module.

Table 2.2: The parameters of the selected illumination module Chromasens Corona II.

Type	C
Module length [mm]	170
Focal length [mm]	190
Color	White
Illuminance [lux]	900,000

This system, considering the method of scanning, allows for objects up to 120 × 43 cm in size to be scanned while maintaining the 350 DPI spatial resolution or in other words, it functions as approximately 100 Mpx industrial-grade camera [34]*.

2.1.2 Mechanical design

A decision has been made to place the scanned object (i.e. the hand) 535mm away from the sensor. This sensor distance has been used, to allow for purchasing a camera lens of standardized focal length.

[2]Available at: https://www.chromasens.de/en/product/dark-field-illumination
[3]Available at: https://www.chromasens.de/en/product/led-light-controller

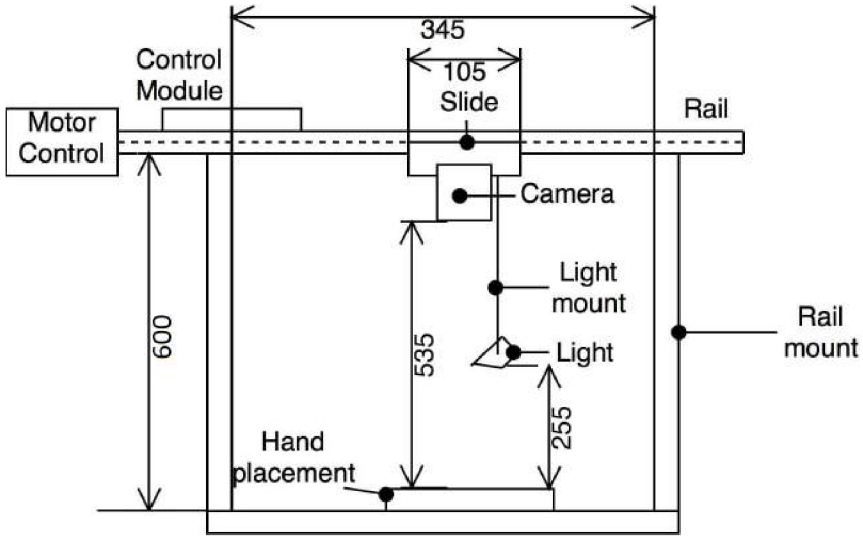

Figure 2.1: Construction design of the scanner in mm.

The Fig. 2.1 represents a front view of the hardware setup in schematics form, along with the physical dimensions in mm. The dimension 345 mm represents the maximum range the scanner can move and, therefore, also the maximum dimension of the resulting image along this axis. During the database collection, the users approached the device from the front, placed the hand onto the Hand placement platform and the scanner performed a complete image acquisition by moving from left to right. Further description of the database collection process will be included in the next chapter.

It is important to note that during the acquisition process, the device, including the scanned object, are completely stationary, except for the parts attached to the slide, by observing the Fig. 2.1 you can see that the camera and light are all affixed to the slide, and thus move at the constant speed with the slide itself. This ensures, that the projected line from the light source and the camera field of view are at all times and the same place, thus ensuring sufficient illumination for the camera acquisition itself.

For the linear movement system, *NL208TR-1200*[4] has been chosen. This system performs a lateral movement at the ratio of 4 mm per revolution of the axis. This device along with the stepper motor allows for a precise position of this camera system, as the relationship between the stepper motors step and the translation in the direction of the movement is precisely defined and allows an exact position both in term of relative movement as well as absolute movement.

The laboratory prototype can be seen on Fig. 2.2 on this image we can see all the key components including the stepper motor, linear movement system, camera, illumination, power sources and all necessary control elements. The white area at the bottom of the image has an outline of human hand and serves as a guide for the user to place the hand, ensuring that the image will include the hand in its entirety. Figures 2.3 and 2.4 show the attachment of the camera and the illumination respectively to the slide. Note that the camera mount has been designed with a stereo camera setup in mind for the future development.

[4]Available at: https://www.teatechnik.cz/linearni-osa-sroubem-nl/

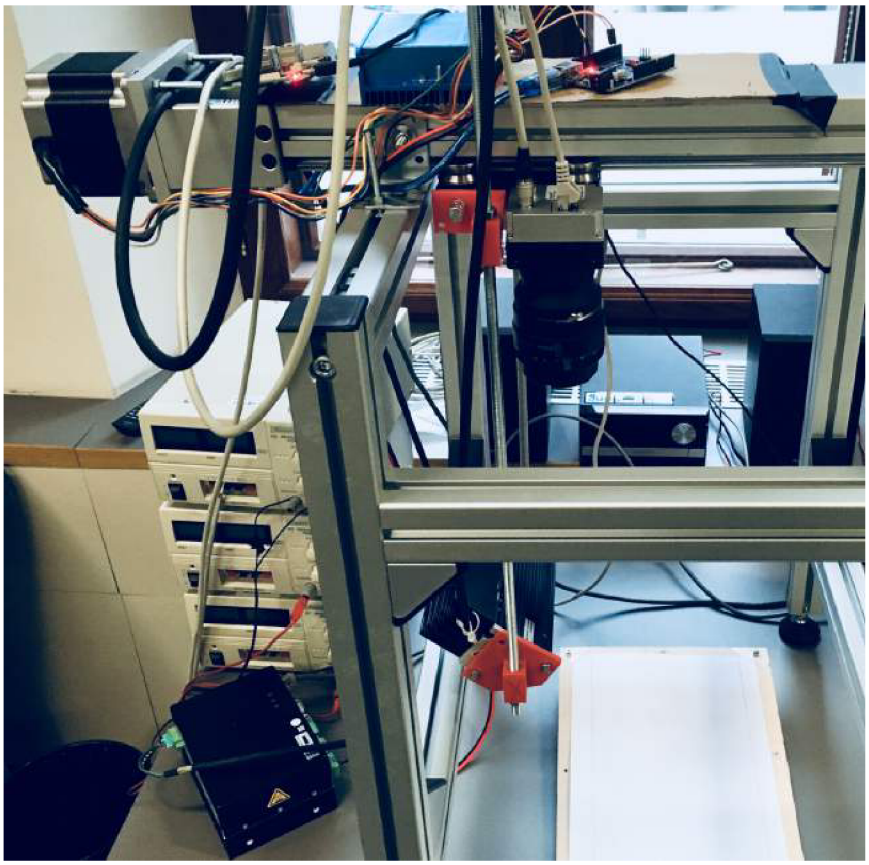

Figure 2.2: Linear scanner laboratory prototype.

At this point of the development, all the parts were purely functional and no effort has been made to affix various power supplies, and control devices to the system itself, i.e. unless the component needed to be at a fixed position as any time during scanning process, it was not mounted onto the device.

The notable exceptions to this where the mounts of the optical system, which have been designed to connect the component and ensure their position remains fixed during the scanning. On Fig. 2.3 a camera mount can be seen. A reader can notice, that the mount has been developed with future stereo application in mind, by allowing a second camera to be attached. Fig. 2.4 than shows a line light projector mount. This attachment was crucial, as it allowed the fine tuning the position and therefore ensure that the focus of the light module was projected to the appropriate position.

All the custom mount have been designed by the author and printed using the Polylactic Acid (PLA) material on the 3D printer. The models can be found in the appendices.

2.1.3 Electronic and electrical subsystems

The design pertaining to the electrical as well as communication aspects of the devices will be discussed in this section. Parts of this section have been published by the author in [34]*.

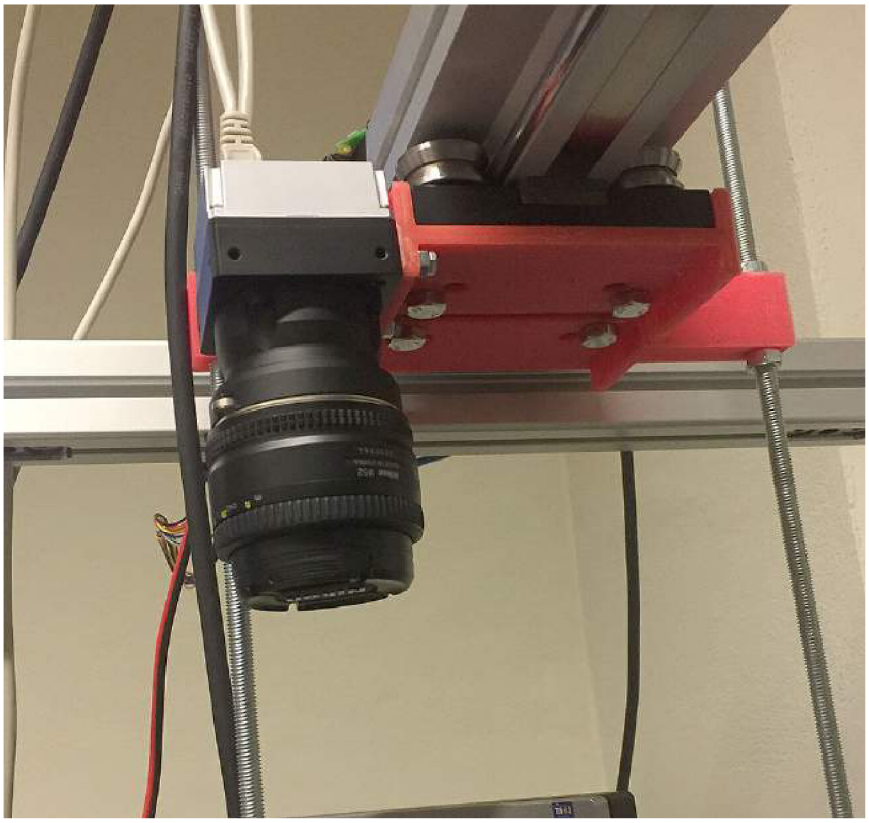

Figure 2.3: Detail of the camera mount serving as a fixture to the device.

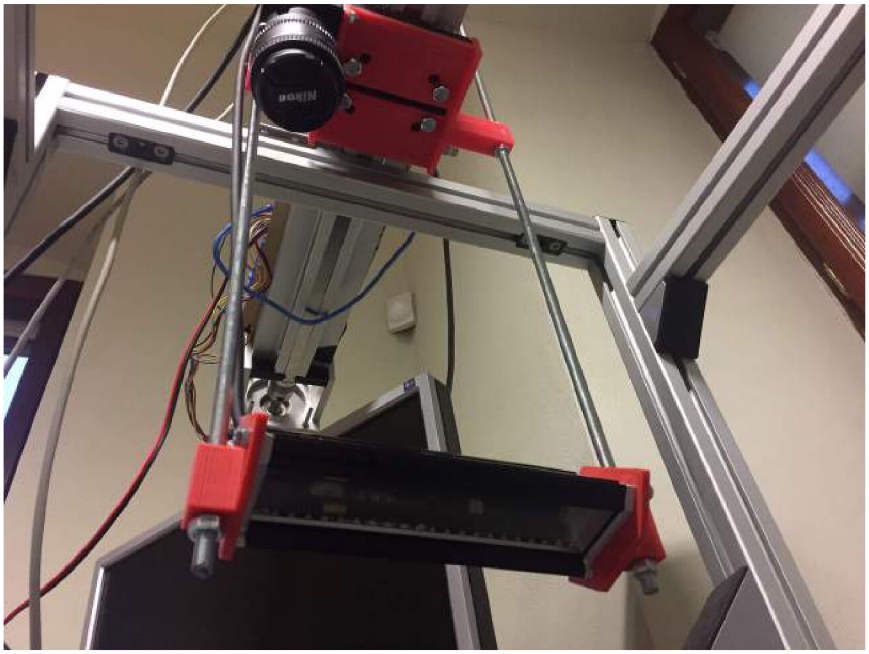

Figure 2.4: Detail of the illumination mount serving as a fixture to the device.

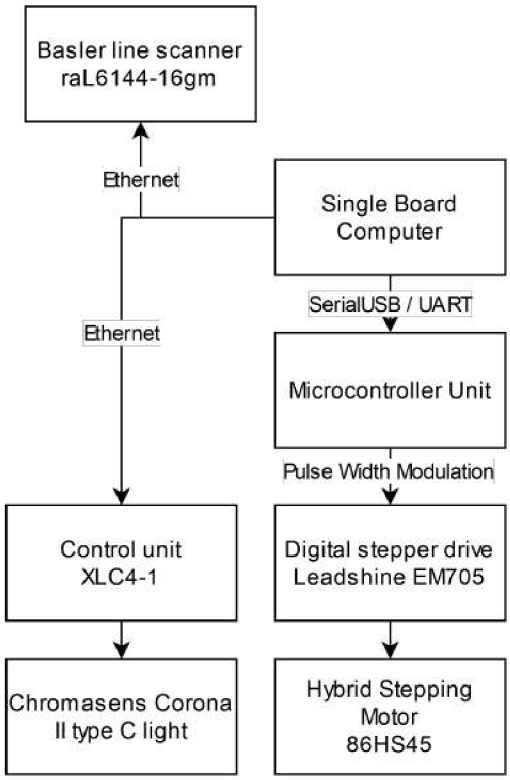

Figure 2.5: Functional diagram of the device, illustrating individual subsystem and communication interfaces.

The individual components used in this device, are listed along with they electrical properties in Tbl. 2.4.

In order to facilitate fast prototyping, instead of developing an embedded system, the proof of concept device has been developed using the of the shelf components, for computing and industrial components for the optical and mechanical subsystems, in order to ensure the repeatability and to eliminate the sources of error that may be present when using consumer electronics.

The main computing unit is a single board computer (SBC) Raspberry Pi 3 (RPi). This device has been chosen, due to sufficient computing capabilities, and the appropriate communication ports. This was especially valuable as the data transfer as well as control was to be performed over both LAN and WAN.

Due to RPi not being real time device, and the necessity to execute some operation, as will be discussed later, at precise times, it has been decided that the movement subsystem will be controlled by a separate microcontroller (MCU) Arduino Uno, RPi is therefore sending only high level instructions, and the actual motor control (using pulse width modulation) is left to the Arduino unit, thereby guaranteeing an absence of various high level exceptions and priority conflicts that would be unavoidable relying solely on RPi.

Fig. 2.5 illustrates how are the individual components connected and the interfaces utilized between them. Optional connection not part of the diagram is a connection between SBC and the LAN or WAN.

Table 2.3: The catalogue parameters of the linear movement axis.

Name	NL208TR1200
Maximum speed [mm/s]	200
Distance per revolution [mm]	4
Accuracy [± mm]	0.1
Length [mm]	1,200
Continuous force [N]	800

For power, the device utilized three Manson NP-9615 DC[5] regulated power supplies. Two were connected in series to generate 48 V and third one was used to provide 24 V, 12 V and 5 V voltage sources. This separation was done in order to shield the low powered devices from the 48 V branch, which was used to power the movement subsystem, which due to rapid oscillation necessary to drive the stepper motor might introduce an unwanted noise onto other branches. The RPi communicates with the Arduino over the USB ports available using UART communication protocols. The summary of electrical requirements for the used components can be found in Tbl. 2.4.

The XLC4-1 and CORONA II are part of the illumination subsystem and need to be powered at all times, in order to receive the instruction from the RPi. The communication protocol will be further covered in the section about the control systems, here it shall be only pointed out that the devices are powered from the 24 V source, and are the largest current drain when operating at maximum intensity and the communitation between RPi and illumination system is done using telnet protocol over Ethernet interface.

The conversion of rotational movement of the stepper motor into the linear motion necessary for the scanning has been realized using a linear motion system with a threaded rod NL208TR-1200[6] made by T.E.A. Technik - see Tbl. 2.3 for basic information about linear axis. An industrial solution has been used, in order to ensure the stability and predictability of the steps.

To further enhance this process, a motor driver that is capable of micro stepping is used. EM705 2 phase digital stepper drive along with 86HS45 stepper motor from Leadshine have been used to achieve that. The microstepping allows the system to perform a rotation movement of a specified fraction, which along with the known revolution to translation function allows us to make a sub mm step of known length. In our system a 12,800 mircrosteps per revolution have been used, giving our system a resolution of approximately 3 μm. Even smaller step was possible, given the equipment used, however as the required resolution per the design has been met and smaller step would decrease the speed, this option was rejected. The communication between the MCU and the motor driver is done using the Pulse Width Modulation (PWM) interface of the Arduino.

Both ends of the rail also contain a digital trigger which indicates, that the carriage has reached an end, this signal source is being monitored by the MCU and triggers a HW interrupt if activated.

[5] Available at: https://old-dev.manson.com.hk/products/detail/54
[6] Available at: https://www.teatechnik.cz/linearni-osa-sroubem-nl/

Table 2.4: The table of electrical parameters of individual components of the scanner

Component	Voltage [V]	Maximum current [A]
raL6144-16gm	12	0.375
XLC4-1	24	3
CORONA II	24	1.8
EM705	48	-
86HS45	48	4.2
RPI3	5	2.5
Arduino UNO r3	12	-

2.2 Stereo vision based 3D object scanner

One of the logical research direction, of the hand geometry based biometric recognition, was to move from the widely used 2D features, and include the 3D features. The reasoning behind this decision is twofold. One, the increased number of independent biometric features leads to an increase in biometric entropy [24][63], which in turns allow the development of system that can offer higher recognition rate, and support a database of more people. The second reason is an increase prevention against the presentation attack detection.

The proof of concept device has been developed as a part of the VRASSEO project [40]. While the project was targeted at the development of 3D weapon imaging. The stereo imaging approach used is a generic 3D imaging method and can therefore be used for the object of similar size, which hand guns and human hand are. The motivation behind the device development as well as the actual design and development of the existing device will be presented in this chapter. The description of the control algorithms and the evaluation of this device as mean for biometric imaging will be presented in the following chapters.

The intended goal of this project is to develop a scanner, that allows a collection of information necessary, to generate a 3D representation of the scanned object. This representation in turn is to be used for the purposes of generating synthetic databases that can be utilized as a training data for the neural networks. This database must be created with future orientation invariance in mind, in realistic conditions and diverse natural environments, in all imaginable orientations and positions. The database will consist of data for the creation of 3D weapon models. In this way, we gain absolute control over the quality level of the database, thanks to the possibility to combine background, orientation and position of the weapon, thus being able to create a database of almost any scale, which would be unfeasible in manual collection. The principle of creating 3D models is based on multimodal data fusion, where we combine depth and plain optical sensors to collect a set of information, such as texture information about the object being imaged and a depth map with the object in the scene, which we use to reconstruct a faithful model, where the different 3D analysis methods are synthesized together to improve the quality of the imaging. In this case, we specifically combine texture information, depth data from the depth camera, and depth data computed from stereo vision methods.

For an idea of how this device looks, see the Fig. 2.6.

2.2.1 Optical and mechanical design

Optical principles utilized in this device were discussed in subsection 1.5.4 of Theory chapter. To enable the 3D reconstruction form the acquired images, a device has been designed

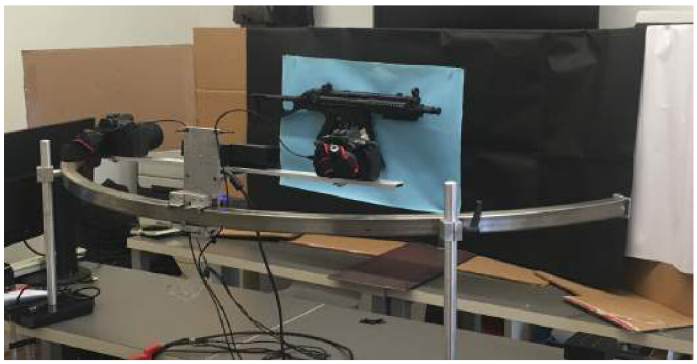

Figure 2.6: Overview of the stereo vision based 3D object scanner

so that by moving a trolley on a curved equidistant path from an object and acquiring the static 2D images along the way, sufficient database of data is gathered, so that we may (together with a priory knowledge regarding the dimensions of the device) reconstruct the 3D object utilising the correspondence problem.

To further increase the database size and to ensure that the sensor position for each pair of images is known exactly, two commercial Canon EOS 80D[7] have been chosen for the sensor, given its availability, resolution of 4,752 x 3,168 pixels and being modifiable in term of the lenses thanks to the common C-Mount. For the same reasons a camera lense CANON LENS EF 24-70mm f/2.8[8] has been chosen.

As can be noticed on the schematics on Fig. 2.7, the object distance between the U-Shaped rail and the scanned object is at most only 90 cm. This distance is further reduced by the design of the trolley. A crossbeam designed to serve as the mount for the cameras places both cameras 87.5 cm away from the object and 62 cm apart. This placement has been chosen, so that the angle between the crossbeam and the optical path of the camera is stable at 70 degrees, thus providing a known parameter.

Given the full resolution of camera of 6,000 × 4,000 px and the distance to object of approximately 87.5 cm, the lens focal length has been set in a manner that ensured that the pixel to mm ratio was at a 10:1, i.e. every 10 pixels corresponded to 1 mm of scene. This has been achieved by adjusting to focus to approximately 35 mm, which corresponds to the theory, given that the centre of the camera lens is placed ahead of the body of the camera. Fig. 2.8 shows an example of an acquired image along with a photo with defined measurements for comparison[9]. The relative error of 1.03% is caused mainly by the depth of an object. The resolution of the image was scaled down in this document.

For summary of camera parameters see Tbl. 2.5

Given the known placement of individual components and the parameters of the system, the 3D reconstruction methods may be used on the acquired data. The data acquisition will be further described in subsequent chapters, in this optical section, we will limit the description to the basic requirements and expected output.

[7] Details available at: https://www.usa.canon.com/internet/portal/us/home/products/details/cameras/eos-dslr-and-mirrorless-cameras/dslr/eos-80d

[8] Available at: https://www.usa.canon.com/internet/portal/us/home/products/details/lenses/ef/standard-zoom/ef-24-70mm-f-2-8l-ii-usm/

[9] Available at: https://us.glock.com/en/pistols/g19

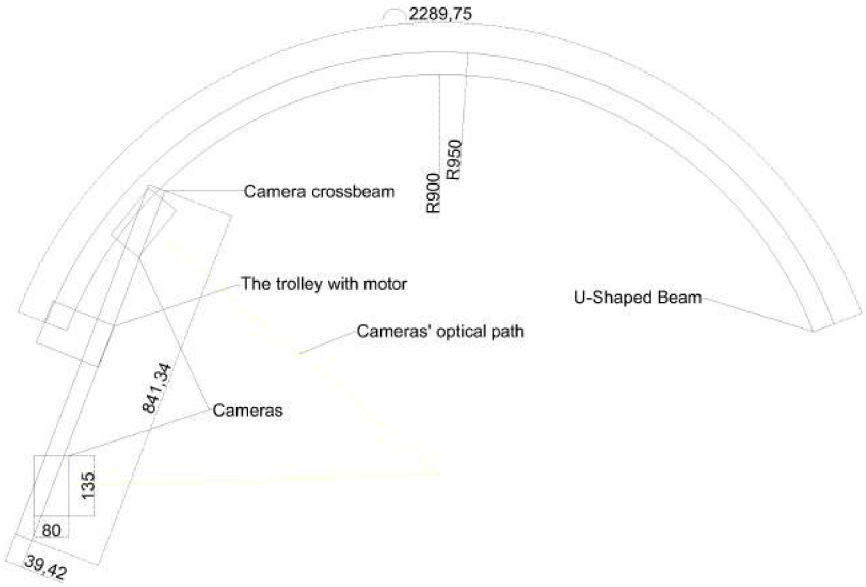

Figure 2.7: Schematics of the device, with placement of key components.

Figure 2.8: Image of G19 Glock with length slide (left). Image acquired by the device with the measured length of the slide in pixels (right).

Table 2.5: Camera setting during the operation.

Parameter	**Value**
Model	Canon EOS 80D
Height × width [px]	6,000 × 4,000
Exposure time [ms]	10
F-Number	8
Focal length [mm]	35

Figure 2.9: Image of FN .32 pistol by Fabrique-Nationale, from Camera (C) 0 Position along the rail (P) 1 (top left), C1 P1 (top right), C0 P2 (bottom left) and C1 P2 (bottom right).

In order gain sufficient data a number of images must be acquired. Per the theory and the researched literature, a sub 10° angle difference per acquisition should be sufficient for the purpose. It has been decided that 21 acquisitions corresponding to approximately 7° or 110 mm travel path per acquisition. This would result in 42 acquisitions per one passing of the trolley which as we know, is often the appropriate number to use.

To get a better idea of how the data are collected, see the Fig. y2.9 where 4 images taken during 2 acquisitions are shown, with the pair being side by side and the subsequent acquisitions are shown below each other. As can be seen the subsequent acquisitions show a very slight difference, whereas the use of cross beam allows for capturing a position that could not be achieved using only one camera.

From the images we can also observe two additional parts included for improved performance. Firstly, a blue screen has been included at the back of the device to simplify the subsequent data processing, by ensuring that even simple thresholding or only slightly customized one may be used for the separation of the foreground from the background. Blue has been used, as a white colored firearms were expected to be scanned. Secondly, an artificial illumination was used, in order to supply an unchanging illumination across the whole data collection, regardless of the sun disposition as well as to enable a shorter acquisition time and limit the effect of vibrations on the acquisition process.

2.2.2 Electronic and electrical subsystems

In this subsection a brief description of the electronics integrated in this device shall be introduced.

As is apparent from the previous section, the primary task of the included electronics, was to either

- Provide power to individual components
- Enable movement of the trolley along the ramp

Tbl. 2.6 lists the number and specification of the individual components used in the device. Except for the DC motor and its driver, each device had its own power-supply. The driver DRV8871 DC motor driver[10] and the motor, EMG30 with integrated encoder [11] were powered by Manson NP-9615 DC[12] regulated power supply, set to the output of 12 V, per the specification.

The EMG30 as can be seen in the specification, was equipped with the 30:1 gear box i.e. 30 turns of rotor resulted in one rotation of the shaft. Also included was a pulse encoder which outputted 360 pulses per one rotation of the shaft. As can be expected, the driver, DC motor, encoder and NI MyRIO platform were connected in the feedback loop and used to position the trolley in manner described in the subsection 2.2.1. Visual representation of this feedback loop can be seen in the Fig. 2.10. The NI MyRIO is in charge of the embedded side of the control algorithm for this device. The high level control is realized by a PC which communicates with the NI MyRio over Ethernet interface.

Further details and photo images of the devices are part of the technical report [40].

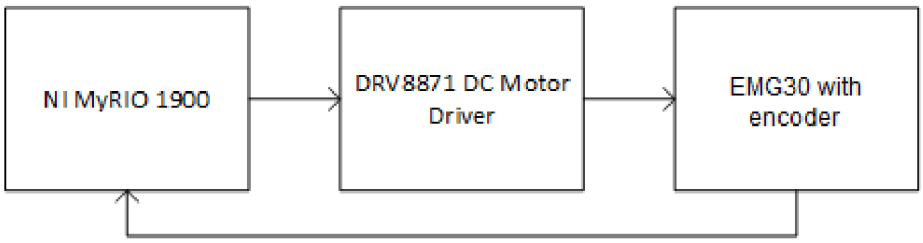

Figure 2.10: HW realization of the positioning feedback loop.

Table 2.6: (Maximum) power requirements of components of the device.

Component	Count	Voltage used [V]	Maximum current used [A]
Canon EOS 80D	2	8	3.00
Light sources	2	18	2.80
NI MyRIO 1900	1	12	1.17
DC motor and driver	1	12	0.53
Microsoft Kinect	1	5	1.50

2.3 On the fly enabled biometric system

In this section a design of the device that was used for the purposes of the touchless and on the fly biometric recognition. The device went through several iterations, and has been used

[10] Available at: https://www.ti.com/lit/ds/symlink/drv8871.pdf
[11] Available at: https://www.robot-electronics.co.uk/htm/emg30.htm
[12] Available at: https://old-dev.manson.com.hk/products/detail/54

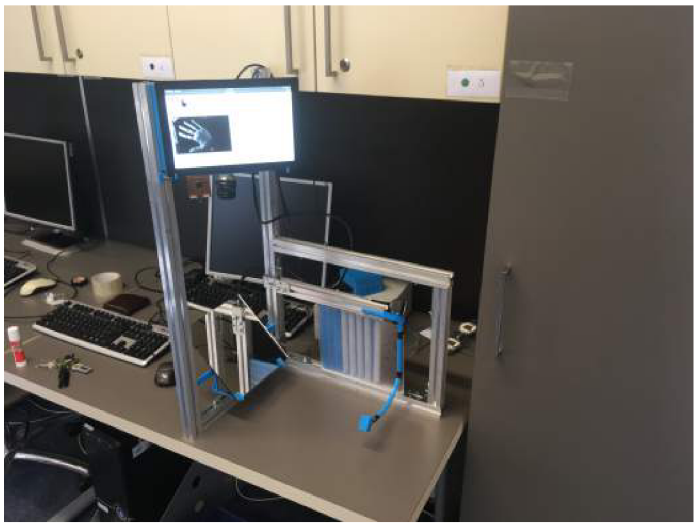

Figure 2.11: First rigid version of 3DHand device.

to acquire hand related biometric features. The most impactful publication was related to the hand geometry and other to fingerprints. As the optical mechanical and electronic development was common to both types of geometry, they shall be discussed together in this chapter. The data collection, processing and evaluation will be discussed separately in the following chapter.

Biometric-enabled systems have become highly desirable to authenticate and identify people due to the proliferation of mobile technologies. Their application in general settings, such as automated border gates using ePassports or at entrances to buildings or rooms, has become popular as well. Regrettably, where current biometry scanning technologies are reliably suitable for single-user devices, when hundreds or thousands of people need to be processed, shortcomings emerge. The requirement of physically coming to a halt during scanning poses a limitation to the throughput of the people processed. Additionally, since most of these systems require physical contact, hygienic concerns need to be considered, especially with contagious pandemics such as COVID-19.

As was presented in [31]*, a solution to the issue of limited throughput and hygienic concerns is a development of a biometric system that allows identification on the fly. Using the experimental setup of our design, it was demonstrated that hand-related biometrics can be acquired on the fly. Furthermore, using a random decision forest classifier and a house-collected database containing over 560 unique images, it was inaugurated, that the system is a viable solution for biometric identification on the fly.[31]*

Nearly all biometric systems today require users to be at a standstill status during scanning. Many biometric systems, especially those working with hand-related biometrics, such as fingerprints or hand geometry, require users to come into physical contact with the capturing sensor [24]. On the fly biometry is a technological subset of biometric systems. In this technology, users are not required to be physically at a standstill to interact with the device, and the biometric features can be collected on the fly. It is in line with on the fly authentication [39][24].

The majority of the biometric-enabled systems can be divided into two main categories, contact-based and contactless [24]. The contact-based systems require the user to come into

physical contact with the device. This may take the form of pushing a finger or the entire palm against a surface of specific sensors. It may also be done by placing the chin and forehead on the surfaces of a fundus camera for a retina biometric. This physical placement may be necessary either because the surface is the sensing part or it serves as a constraint positioning, facilitating the proper user scanning process by keeping the respective subject in place during the scanning process.

On the other hand, contactless biometry does not require the user to come into contact with the scanner. This attribute places further requirements on the system design side, but gives the user a less restrictive experience during the data acquisition, making the process faster. Facial [50], iris [17], and ear recognition [12] are notable examples of biometric systems that usually do not require the user to be in contact with the system during scanning. Worth noting, however, is that the user shall remain at a standstill for a specific time apart from a few exceptions. This requirement perceptibly lengthens the whole recognition process duration, even if the recognition process itself takes only a fraction of a second.

On-the-fly biometry eases user constraints during the scanning process by enabling data acquisition during user movement. This allows a much higher throughput rate, as the user does not have to be at a standstill during the scanning process and can just walk by. This is not to be confused with in-the-wild acquisition, demonstrated for facial recognition in [65], as for the on-the-fly scenario, a level of cooperation is expected from the user.

Apart from the facial recognition systems [81] and, recently, finger vein recognition [57], this approach has not been widely adopted due to several complexity limitations. This is due mainly to the technological principles the commonly adopted biometric systems depend on, where the users' physical contact and limited movement are necessary. Such limitations can be noticed in fingerprint-based identification systems, where aside from a few notable exceptions such as MorphoWave device[13] and research done in [104], the sensing technologies require the user's physical contact. Whether it is a resistive, capacitive, or even optical approach [97], the requirement to stay at a standstill and press the finger against the scanning surface of the device is essential [74]. In facial recognition, the issue of limited throughput, as apparent when using automated border gates, takes a shape of a considerable amount of time to be processed [71]. Development of an on-the-fly system that focusses on high-volume identification would alleviate these issues and could, therefore aid the implementation of faster biometric gates [64][58] and allow for the eventual spread of Biometric-Enabled Watchlists Technology [59]*[31]*.

In the paper[31]*, we proposed requirements that, when implemented in a device, would allow for a biometric system to be capable of collecting images on the fly with acceptable accuracy. In this chapter the optical, mechanical design will be described as well as the electronics designed and included for the purposes of the on the fly recognition. The database construction, control systems and evaluation will be discussed in the next chapter.

2.3.1 Optical and mechanical design

In this section an optical design of the device intended for the on the fly acquisition will be presented. Section has been separated into two subsection, as two major iterations of the device have been developed. The first intended as a proof of concept prototype has been used to collect a database and demonstrate that it is biometry enabled by performing a fingerprint recognition benchmark. This device has been realized using customer grade

[13] Available at: https://www.idemia.com/sites/corporate/files/product-downloads/file-2018-10/morphowave-compact-idemia-brochure-201909.pdf

optical element. The second iteration, a prototype, was build based on the experiences based on the first one and was to be used as a basis for a commercial device. This distinction influenced the choice of individual components as well as the overall design requirements that arose due to the optical, construction or simply design requirements, which will be discussed in the following subsections.

Proof of concept version

As was discussed in the section 1.7 equations 1.16 - 1.17 describe the relationship between the physical parameters of the optical system and the resultant depth of field. In order to develop a system that can capture necessary biometric information in one image, we need to ensure that during the moment of acquisition, the area of human hand that contains said information is in focus. Or in other words, is currently located in the interval defined by the depth of field.

This introduces a design problem, several factors need to be considered:

- The depth of field needs to be sufficient to enable in focus acquisition;
- the overall sensor exposure needs to be sufficient so that the ridges can be analysed;
- the illumination must not be so bright as to cause discomfort to users;
- the overall dimensions of the device must remain practical- The smaller the better.

A careful reader may have already spotted a problem with these requirements. Based on the theory discussed in the first chapter, we know that the depth of field can be increased either by closing the aperture of the camera lens, or by increasing the distance from the sensor to the object. Beyond certain distance, the object distance can no longer be practically increased as the device becomes impractically big, and we need smaller aperture to further increase the depth of field. However, by closing the aperture, the amount of light that reaches the sensor decreases. We can compensate for this, by enhancing the illumination, or opening the aperture. However we cannot enhance the illumination too much as beyond certain threshold, it will begin to cause discomfort to the user and the aperture can't be opened too much, as it results in shallower depth of field. Balancing these four parameters eventually lead to the development of the first proof of concept.

It has been decided that the first prototype must in theory allow for the resolution of at least 500 DPI, and that the acquisition area must be at least 240 mm of width and 200 mm of height. This has been deemed sufficient for the user to use without an undue risk of collision with the frame of the system. To denote this area, a screen with a void of sufficient dimension was to be built that signifies, where the acquisition area is. This type of design is meant to enforce the suggestion for the user as to where he is to expect the acquisition to take place. See the Fig. 2.11 to see a general setup including the guiding screen and the implied acquisition area. As this has been a proof of concept prototype and the triggering mechanism has been developed by utilising the remote trigger mechanism of the Canon camera, meaning the the trigger delay was not exact, a mirror has been included into the frame, which was designed to be used to determine the actual position, the depth of the hand at the moment of the acquisition. Placement of this mirror, required additional 50 mm of the width of the frame, defining the required field of view to be 300 $\times$ 200 mm.

Fig. 2.12 shows a schematic of the system setup used, indicating the location of the camera, gate and light sources. A commercial Canon EOS 50D[14] has been chosen for the sensor, given its availability, resolution of 4,752 $\times$ 3,168 pixels and being modifiable in term

[14] Available at: https://tinyurl.com/4uxhnn6v

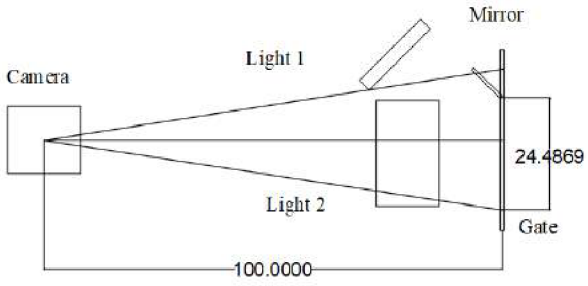

Figure 2.12: Schematics for laboratory setup for OTF fingerprint extraction, top view in mm.

of the lenses thanks to the common C-Mount. For the same reasons a camera lens CANON LENS EF 24-70mm f/2.8[15] has been chosen. In addition to already mentioned reasons, this lens was also deemed advantageous, due to manually adjustable focal length and aperture size, which ensured, that alternative setups may be tested, prior to purchasing fixed focal length lenses, for the future versions.

To test the viability of the setup for the biometric application, and more specifically a fingerprint recognition, spatial resolution of at least 400 DPI was desired. Given the above mentioned setup and using equations discussed in chapter 1.7 a spatial resolution of approximately 402 DPI was achieved at the object distance of 1,000 mm with the camera lens being set to the 70 mm per our calculations. The second factor that needed to be addressed was the DOF, again, given the equations in the theory chapter of the thesis, we were able to determine that given the distance and using the f-number $f/8$ we are able to resolve 0.025 mm circle of confusion in the objective distance rage between 963 and 1,039 mm, i.e. 76 mm.

The last parameter that needs to be solved is the acquisition or exposure time. In the x, y direction, the calculations have been made with the future 500 DPI in mind. To keep with this parameter in the z axis (towards the sensor), we need to estimate the drift in the x and y direction as the component of the forward movement along the z axis. Empirically, we have observed an approximate drift of 9%. If we assume a drift of 10%, i.e. assume that for every 100 mm, a user moves his hand in the forward direction, he drifts at most 10 mm in x and y direction, this gives us an ability to estimate a necessary minimum shutter speed, to achieve it. During the development we have decided that the device should support an arm swing speed of 50 cm/s. This per the assumptions above results in the required shutter speed to be at about 5 μs.

The last piece of equipment that needed to be addressed was the illumination. For this purpose a pair of Amaran H528S[16] light sources illuminating at 4,380 lux are used, the

[15] Available at: https://tinyurl.com/3er5z599

[16] Available at: www.bhphotovideo.com/c/product/1146129-REG/aputure_amaran_al_528s_daylight_led.html

Figure 2.13: The proof of concept device.

adjustable light output was deemed advantageous for the purposes of finding the appropriate light output to facilitate the high shutter speed used. Fig. 2.13 shows the actual laboratory setup indicating the camera, gate, and lighting locations. Fig. 2.14 shows an example image acquired by the device given the parameters described above.

Laboratory prototype

The laboratory prototype has been developed based on the experiences and knowledge gained during the experimentation with the proof of concept setup.

As stated at the beginning of this section, the goal of this device has been to construct an on the fly hand geometry recognition device, capable of demonstrating a novel biometry enabled approach to user recognition. While the methodology remains unchanged from the proof of concept, the consumer electronics, has been replaced with industrial grade optical components or the custom build electronics in the case of the illumination. The finished system was to offer a consistent output of predictable quality, to maintain the option of commercial development, the importance has been placed on utilising as much off the shelf components as possible without compromising the quality.

For this version operation requirements have been defined more rigidly.

The following requirements should be used by a biometric system using on-the-fly biometric identification to be feasible [31]*.

1. The scan quality should meet the requirements described in the biometric specification for personal identity verification, NIST SP800-76-2;
2. the device should be able to scan the image of a human hand, suitable for biometric identification, travelling unidirectionally at a speed up to 500 mm/s;
3. the device should be able to scan the image of a human hand, suitable for biometric identification, horizontally misaligned by no more than 35 mm;

Figure 2.14: Example of the device image acquisition.

4. (optional) the device should be affordable to allow widespread adoption.

It has been decided that the subset of components and their parameters have the most considerable bearing on meeting the requirements mentioned above. The system's main components that should be well-tuned are the camera, camera lens, triggering mechanism and light source. For the experimental setup, the following components have been selected to facilitate those requirements:

1. Basler acA4024-29um. The resolution of 4,024 × 3,036 px allows for a sufficient window of at least 500 DPI scan area. The camera also possesses a hardware trigger and sufficient shutter speed to facilitate requirements 2 and 3;
2. 35mm C-Series fixed focal length lens, with manual focusing and adjustable aperture, to facilitate requirement 3;
3. two Amaran H528S LED illumination sources to facilitate requirement 2;
4. custom made analogue laser triggering mechanism that serves as a signal generator for the camera operation.

The focal length of the experimental setup is 100 cm. This distance denotes a minimum length the camera needs to be from the scanned object. If the sensor and the object are placed on the straight line, it is also the lower bound length of the device itself. In reality this distance is even longer, as various mechanical fixtures and additional equipment needs to be present, further increasing the already fairly large size of the device. This clearly introduces a physical setup problem. Both in terms of space required, and during the transportation of the system. To make the laboratory setup more compact and transferable, a mirror is used.

By placing the mirror at a distance between the camera and the object, we can change the width/height ratio of the result device, making the device seemingly more compact and easier to transport, while maintaining the parameters of the optical path necessary for the device to work as designed. By positioning the mirror at an angle of 45°, we will avoid

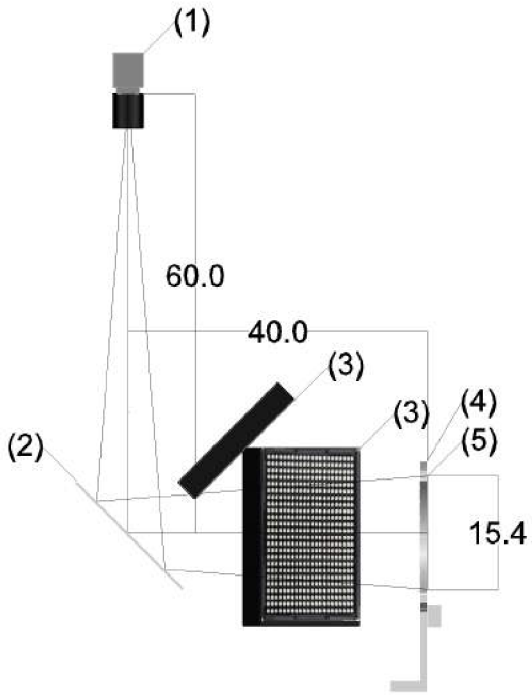

Figure 2.15: Side view schematics and optical path: (1) camera, (2) mirror, (3) illumination sources, (4) acquisition gate, and (5) field of view at the focus of the camera in mm [31]*.

deformation of the acquired images. Fig. 2.15 shows the component placements. Fig. 2.16 shows the 3D model of the experimental setup.

The photographs and exact construction details of the device are a part of the accepted patent DE102020200569A1, Einrichtung zur biometrischen Identifikation mit Hilfe von Fingerabdrücken und/oder Handcharakteristiken und Verfahren zur biometrischen Identifikation mit Hilfe dieser Charakteristiken [33]*.

As with the proof of concept setup, this version also includes a guiding structure that assists users in assuming a correct position. It has been significantly modified and improved based on the observation of the use of the previous version of above mentioned gate.

This new partition with a void, the acquisition gate, assimilates the shape of the capital letter 'C'; see the label (4) in Fig. 2.16. The area of the void has been further reduced. This is necessary, as the camera has only limited resolution and, therefore, only a limited area visible at the focal point. Unlike the previously used CANON 50D the total resolution is smaller, therefore to maintain the high spacial resolution the acquisition area needs to be reduced as well.

The image will be of a spatial resolution, as defined in Requirement 1. The gate has been designed and integrated as part of the device to make the device more intuitive to use. The rationale here is that the user can better assume an appropriate position. The user is more likely to use the device as intended, and the resultant image will be more likely at the expected quality.

The triggering mechanism is a light barrier attached to this gate, consisting of low power laser and light sensitive element. The design and description of this circuit will be further discussed in the following chapters. Here we shall only state, that this circuit serves as a

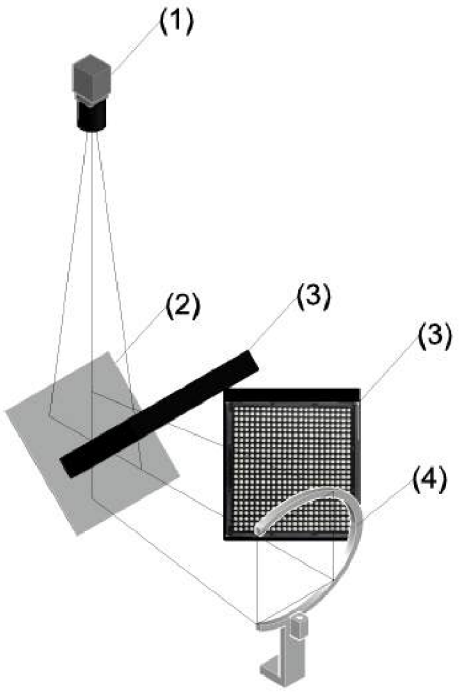

Figure 2.16: A 3D model of the experimental setup. (1) Camera, (2) Mirror, (3) Illumination sources, (4) Acquisition gate and (5) Field of view at the focus of the camera[31]*

trigger for the camera. As the Basler acA4024-29um contains a hardware trigger, we are able to perform the acquisition at the very accurately determined moment, i.e. when the user's hand is at precisely the right position.

Commercial prototype

Along with the Touchless Biometric Systems s.r.o.[17] and Touchless Biometric Systems AG[18] a commercial prototype of the above described setup has been developed. The goal of this development was a device that could be presented at a INTERSEC Dubai trade fair to be presented to a security community as well as to develop a setup, that might be commercialized.

In that regard, it has been decided, that some mechanical and electronic aspects need to be changed, more on that in the next subsections, which in turn changes requirements of the optical design. This took form primarily in changing the position and size of the mirror, position of the camera, which then necessitated the change of the camera itself.

This prototype has utilized following components:

- **acA5472-17um Basler camera**[19] with the resolution of 5,472 px × 3,648 px, this choice was based on the need to increase the field of view, while maintaining the spatial resolution, as well as improve the low light performance by increasing the

[17] Availbale at: https://www.tbs-biometrics.cz/index-en.html
[18] Available at: https://tbs-biometrics.com/en/
[19] Available at: www.baslerweb.com/en/products/cameras/area-scan-cameras/ace/aca5472-17um/

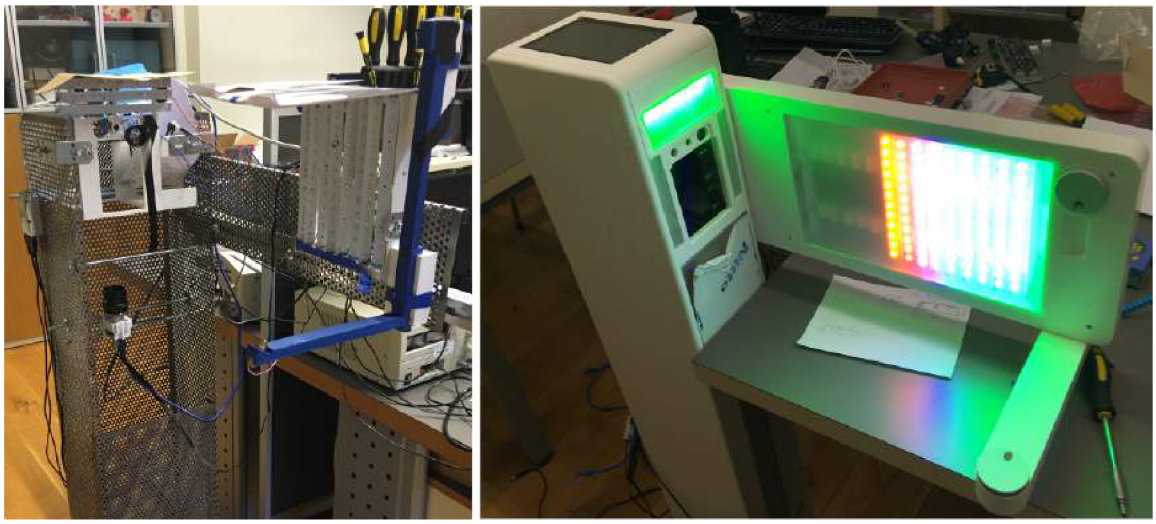

Figure 2.17: Custom build case with updated components and placement (left), final prototype with contracted case (right).

pixel size to 2.4 µm × 2.4 µm from the 1.85 µm × 1.85 µm of the acA4024-29um Basler camera.

- **35mm focal length lens, 1" sensor format Sensor**[20] a sensor suitable for our larger sized sensor. As can be noticed, the focal length remains unchanged, as the object distance remains the same.
- **73 × 116 mm, 4-6λ mirror**[21] this mirror has been chosen, to eliminate the artefacts caused by total internal reflection of the non-optical mirrors, by having a reflective top surface of the mirror. It was also of the size that fit into the casing of the commercial prototype.
- **Custom lighting** as the commercially available illumination was no longer sufficient for this device, a custom panels and circuits to control them have been developed and used in this device. Giving the system an ability to illuminate the object for just the duration of the acquisition and then be shut off, in other words, it should operate as a flash illumination. As the properties of the illumination circuit will be further elaborated on in the next subsection, in this subsection we will only state, that for the purposes of the biometric identification a light source with wavelength of 525 nm i.e. green, was chosen for the illumination. This choice has been made based on the theory, which postulates, that the green light source offers the best contrast due to the palm of the hand being red dominant.

On Fig. 2.17 the final version of the prototype can be seen. On the left is a development version, using the final selection of components, distributed according to finalized design plan, and on the right a custom build presentation version, with same components and components placement. The development version (on the left) is bare and offers a better overview of the placement of individual internal components, as well as the new gate concept. As can be seen, the 'C' shaped gate has been replaced per a design decision and only an 'L' shaped protrusion is present on the system. This design is partially necessary, as it

[20] https://www.edmundoptics.com/p/35mm-focal-length-lens-1quot-sensor-format/17863/
[21] https://www.edmundoptics.com/p/73-x-116mm-4-6lambda-mirror/2314/

enables a placement of the components of the laser gate, which is used for the triggering of the system. However, it has primarily been used to improve user experience, as the 'C' shape was too limiting and a number of people, especially less experienced users, hit it during the transition. The 'L' shape removes this drawback. It is however also one of the reasons why the used camera has a larger resolution, as due to decreased constraints, the user's hand at the moment of acquisition may be located at a larger area. In order to capture larger area with identical spatial resolution, the resolution of the sensor needed to be increased.

Another obvious difference is the replacement of commercially available illumination, and instead, a custom illumination has been developed, the characteristics of which shall be discussed in greater detail in following subsection. The placement of individual components can be seen in the Appendix A.1.

2.3.2 Electronic and electrical subsystems

The optical an mechanical aspects of the device were described in the previous subsection. In this subsection, the various electrical and electronic aspects of the devices shall be described, as they were being designed and developed, in various iterations of the system.

Optical gate triggering circuit

From the first version of the device the necessity for fast triggering was paramount. Subsequently it was decided that an optical triggering system shall be the most suitable approach, as its overall analogue nature allows for practically immediate triggering, with no need for processing the signal.

For this purpose a circuit has been designed to facilitate this function. On Fig. 2.18 we can see the key subpart of the schematics which is included in Apendix A.3. On this Fig. we can see a part of the circuit that performs the main function, where an output from a photo diode (outside of schematics on the left) leads to the input of the operational amplifier (OPAMP) (2) and is amplified by utilising the feedback loop over R25 resistor. The output from the amplifier (1) is lead to another OPAMP, which is setup as a threshold, and will output approximately 5 V, if the input is higher than the threshold and be grounded if not. The threshold level is defined by the voltage divider made up of resistors R28, R29 and R30 and allows for tuning the threshold between 0.32 V and 0 V per the Eqn. 2.6 below.

$$V_{thresh} = V_5 \cdot \frac{R2}{R1 + R2} \tag{2.6}$$

$$R1 = (R28 + R29_a) \tag{2.7}$$

$$R2 = (R30||R29_b) \tag{2.8}$$

where $R29_a$ and $R29_a$ represent partial resistances of the potentiometer $R29$ i.e. a resistance between ports 1-2 and 2-3 respectively of the resistor $R29$ from Fig. 2.18, symbol $||$ signifies a resistors connected in parallel.

This range has been chosen based on the technical documentation, of the photo-diode, the laser module and empirical observation and experimentation, which aimed to answer how to best accommodate the following criteria. In order for circuit to provide the required function, it must:

1. To trigger when the path of the optical beam is crossed by a hand;
2. to not trigger due to variations in the ambient light level in the room.

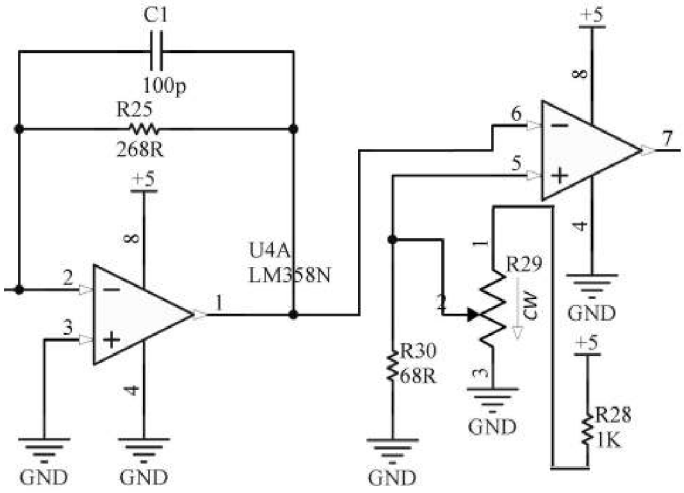

Figure 2.18: Snippet from the optical trigger schematics.

Why these conditions are important, can be seen on the simulated situations on Figures 2.19, 2.20 and 2.21.

Fig. 2.19 illustrates a well set trigger threshold, the voltage drop due to hand transition will cross the designated threshold and the circuit will generate a trigger as intended. This scenario is clearly the intended mode of operation, an will behave in the expected manner.

The Fig. 2.20 illustrates a situation where, for example due to ambient light being low or even absent, the voltage drop is insufficient to cross the set threshold and the trigger will not be generated. If during the installation, this behavior were to occur, the trigger threshold would be increased by adjusting the potentiometer R29 so that the combined resistance R2 increased and R1 decreased.

Fig. 2.21 illustrates an opposite issue, due to ambient light level being higher than usual, and the threshold voltage being set too high, the hand transition alone would not generate the trigger. In this scenario the behavior of the circuit would be unpredictable, which is clearly an unwanted behavior. If during the installation, this behavior were to occur, the trigger threshold would be lowered by adjusting the potentiometer R29 so that the combined resistance R2 decreased and the R1 increased.

Since the device in this iteration, did not have a system to account for the ambient light variations, the threshold was expected to be set manually during the device installation, and the device itself was designated for the indoor use only.

The source of the light beam in this iteration was a laser module ML-SHW-650-05-D[22] with 650 nm wavelength and 4 mW of power. Photodiode with increased spectral efficiency at the 650 nm SFH 203[23].

Due to time constraints, this circuit has not changed across all the iterations of the device, and was, in subsequent iterations, integrated into the main control board. Only modifications were in adding a terminal and protective elements, to protect the camera circuitry. To see the latest iteration, please see the Appendix A.3.

[22] Available at: http://www.shawohwa.com/en/products-list.php?tid=2&id=10
[23] Available at: https://tinyurl.com/vn8mw4nw

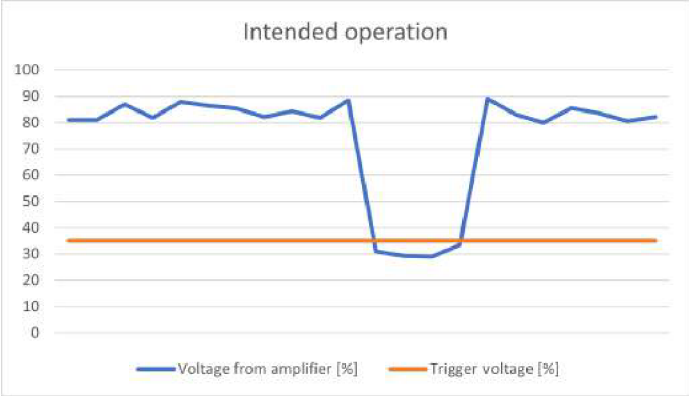

Figure 2.19: Example of voltage drop during hand transition, with correctly set trigger threshold.

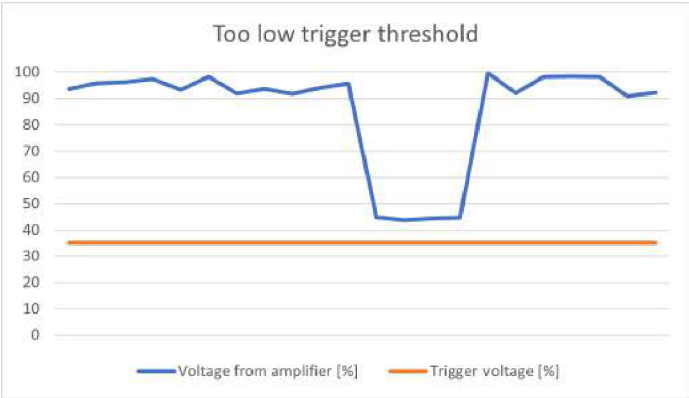

Figure 2.20: Example of voltage drop during hand transition, with trigger threshold being too low.

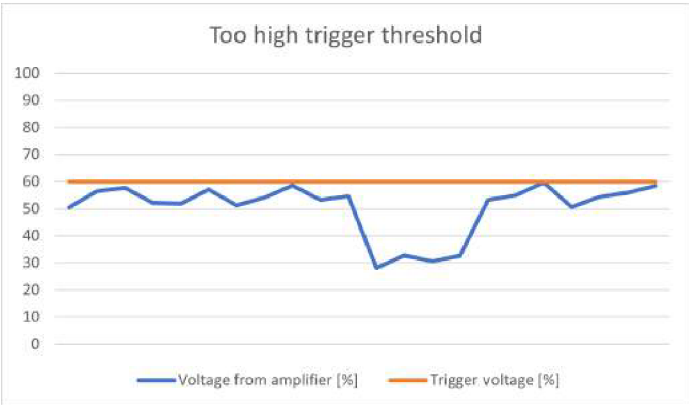

Figure 2.21: Example of voltage drop during hand transition, with trigger threshold being too high.

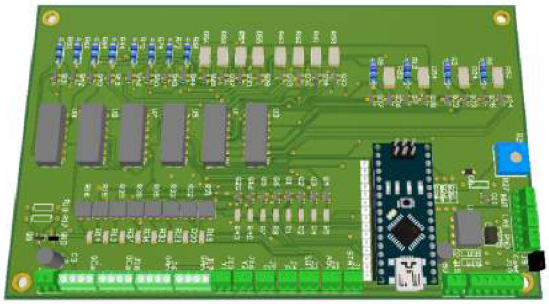

Figure 2.22: Simulated appearance of the control board.

Control board

When the demonstration prototype was being developed, the always-on commercial illumination was no longer suitable. Instead a source of light that triggers only during the moment of the acquisition, similar to flash illumination of camera was required. No commercially available solution seemed to be suitable for the conditions that were to be expected, therefore a decision has been made, to develop a main control board. For illustration Fig. 2.22 includes an appearance of the finished control board.

The board was to facilitate following functions:

1. Supply adjustable power to the custom illumination panels;
2. support multiple (two) power supply modes for illumination panels;
3. facilitate the optical trigger functionality;
4. supply power to the embedded controller;
5. allow embedded controller, to control the illumination mode;
6. allow embedded controller to observe the status of the optical trigger;
7. allow embedded controller to trigger illumination for a definable duration;
8. embedded controller will support an external triggering;
9. output a trigger signal (for camera and other electronics).

And following utility features:

- If the embedded controller fails, or is removed, the illumination must not be enabled;
- the circuitry must be designed in such a way that should the illumination stay on longer than expected either due to malfunction or an unexpected behavior, the board must not be damaged.

As can be imagined, many of these functionalities are combined, the following section describes how is this functionality achieved and a reference to corresponding circuit diagram.

Supply adjustable power to the custom illumination panels function serves an important function, as the level of illumination required for the short acquisition time is substantial and it needs to be ensured, that power level is both sufficient to achieve desired illumination and at the same time meets the tolerances of various components in the control board. Rationale behind the requirements for the subsystem being adjustable, is to enable an integrator, to take the environmental factors into consideration, e.g. if the room is well lit, the artificial illumination does not need to be at its full capacity. A voltage-follower

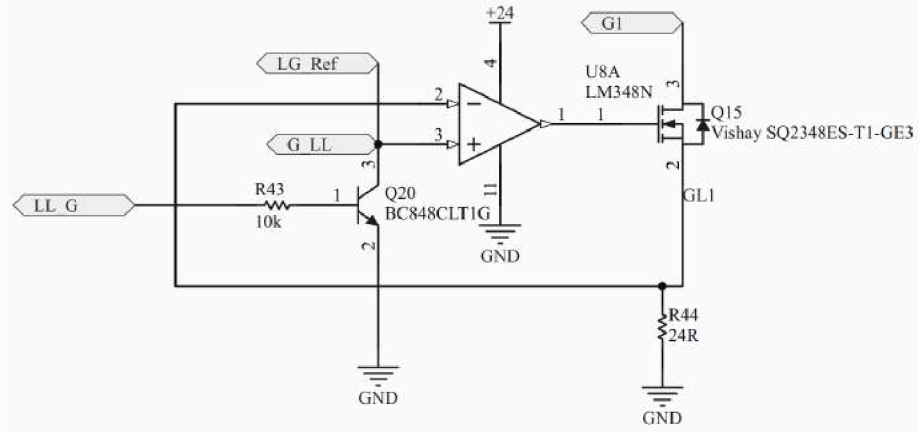

Figure 2.23: Voltage follower circuit used to power illumination panels.

circuit using operation amplifier has been chosen to facilitate this function. See Fig. 2.23 for an example of a circuit used to power one branch of the illumination panel.

On Fig. 2.23 we can describe a general principle of operation. A digital signal from an embedded controller sets port LL_G and either opens or closes NPN transistor Q20. When the transistor is opened only voltage equal to saturation voltage of the Q20 is on the +input of operation amplifier and the Q15 transistor remains closed as the minimum Gate-Source Threshold Voltage of SQ2348ES[24] is 1.5 V. When closed, LG_Ref value is on the OPAMP +input, this will lead to opening the Q15 transistor. Due to the feedback to OPAMPs -input the transistor will open only enough to ensure +input and -input match. As the reference voltage is generated using a potential divider, such as for example can be seen on Fig. 2.24 this allows us to set a reference voltage in the interval of 0-3.472 V. While the circuit is by definition voltage follower, the voltage is set, based on the current flowing through the R44 resistor. In the Fig. 2.23 this value matches the current of 145 mA maximum current for dual illumination panel, i.e. 72.5 mA per LED. As can be expected, this setting is intended for the low power mode. For complete power supply and reference voltage generators, see Appendix A.4, for LED drivers either Appendix A.6 or Appendix A.5.

Support multiple (two) power supply modes for illumination panels function is a feature that emerged a bit onto the development, where the users commented on the poor user comfort during the triggering of the illumination. The instant powering of light panels provided too large a difference in the illumination and the unprepared users complained about a discomfort of this exposure. To mitigate this, a dual mode functionality has been proposed. While the device is in standby mode, waiting for a trigger, the light panels are to be partially powered, in a so called low-light mode (LL mode) whereas during acquisition a high-light mode (HL-mode) was to be used, where LEDs would be driven by a higher current.

To add this functionality, circuitry already described previous paragraph, has been duplicated, and connected in parallel to the light panels. This based on the digital trigger enables either an LL-mode circuit or HL-mode circuit would be activated driving the cur-

[24] Available at: „https://www.vishay.com/docs/63706/sq2348es.pdf“

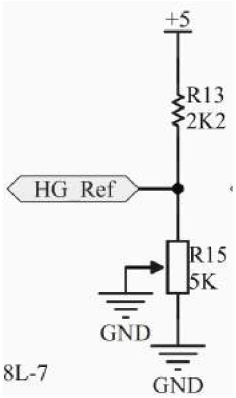

Figure 2.24: Circuit used to set a reference voltage for the illumination panel.

rent through LEDs. Appendix A.6 shows the circuits used for driving LEDs in LL-mode. Appendix A.6 shows the circuits used for driving LEDs in HL-mode.

Facilitate the optical trigger functionality function has already been described in Subsection 2.3.2. And circuit wise has been integrated into the control board without modifications. Please see said subsection for more information or Appendix A.3 for the related schematics.

Supply power to the embedded controller function is being managed by a commercially available 5V voltage supply. See Appendix A.4 where the power is being connected to the board via connector J15 and possibly Appendix A.2 where you can see how the 5V net is connected to the embedded controller.

Allow embedded controller, to control the illumination mode function is being facilitated directly by the embedded controller, see Appendix A.2 where the controller via pins D4, D6, D9 and D10 can trigger the HL-modes for red (LED_R), blue (LED_B), yellow (LED_Y) or green (LED_G) light panels via the dedicated nets respectively. D3, D5, D8 and D11 then similarly can trigger LL-modes. Appendix A.6 and Appendix A.6 show how the above mentioned nets connect to the individual driver circuits.

Allow embedded controller to observe the status of the optical trigger function is being facilitated by the INT0 net. The signal is generated in the optical trigger circuit, and is connected to and is received by the embedded controller at port D2, see Appendix A.2 for more details. As can be seen in the Appendix A.3 this signal is capped at 3.3 V by the Zener diode to ensure that the digital input port of the MCU does not get damaged by a voltage surge.

Allow embedded controller to trigger illumination for a definable duration function is primarily performed by the firmware. On the HW all that was required, was to keep the various delays of individual components and ensure, nothing can unexpectedly influence it in a significant manner. But overall, this issue has been prevented from emerging by primarily using analogue circuit design, with the MCU being used dominantly for simple

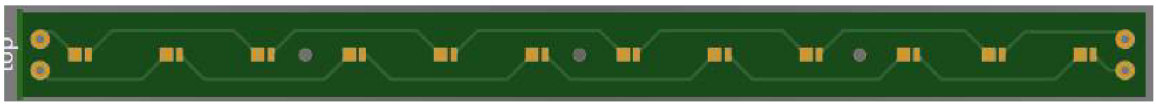

Figure 2.25: PCB of single light panel sized 150 × 15 mm.

triggering and/or timing. Especially since the optical trigger circuit and LED drivers refrain from using any digital circuitry, potential sources of unwanted delays have been kept to a minimum.

Embedded controller will support an external triggering function. As can be seen in Appendix A.2 the D7 pin is connected to a J3 port. The D7 pin supports HW interrupt function and can be used for an external triggering. In this particular case it is to ensure that the light is triggered, when the camera reports a start of an acquisition to the main computing unit.

Output a trigger signal (for camera and other electronics) function has been discussed as part of Function 6. The optical trigger circuit generates a digital signal when the laser beam is crossed by a hand. Ports J14, J9 and J10 all output this signal, the signal is same for J10 and J9, where J10 is to be used for camera and J14 is inverted and used by control units.

Light panel

Light panel or illumination panel is the singular segment of the custom illumination system. A series of these panels, namely eight was designed to be used to illuminate the hand during the acquisition by a camera. While panels with red, blue, yellow and green LEDs have been designed and installed, for the purposes of this thesis only green is relevant and will be further discussed. Circuit can be seen in the Appendix A.7. Fig. 2.25 is a large panel to be used in the side of the system and Fig. 2.26 a small panel to be used at the front of the device, both boards are shown to get a better idea how the light panel looks and operates, for an idea how the panels are fitted in to the device, see Apendix A.1 (in yellow), keeping in mind, that there were 3 front facing light panels placed on the front of the device and 5 were place on the side wing of the device (the first 5 from the right). Each panel was powered by a two LED drives as discussed in the control board subsection previously, one for each illumination mode.

In the demonstration prototype RF-GNRI35TS-DK-2T[25] micro LEDs were used, 12 per panel, 2 × 6 LEDs with two group connected in parallel and LEDs, within one group connected in series, thanks to this configuration a 24V power source could be used to achieve maximum stated illumination. This setting doubles the current draw per panel as opposed to having all LEDs connected in series, however, individual components have been chosen to ensure it could be run for theoretically unlimited duration without a risk of damaging the related circuitry.

[25] Available at: „www.refond.com/upload/resources/rf-gnri35ts-dk-2t-e1/RF-GNRI35TS-DK-2T-E1.pdf“

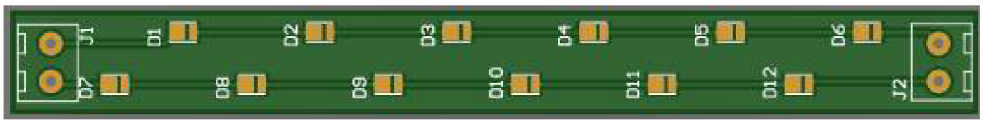

Figure 2.26: PCB of single light panel sized 130 × 15 mm.

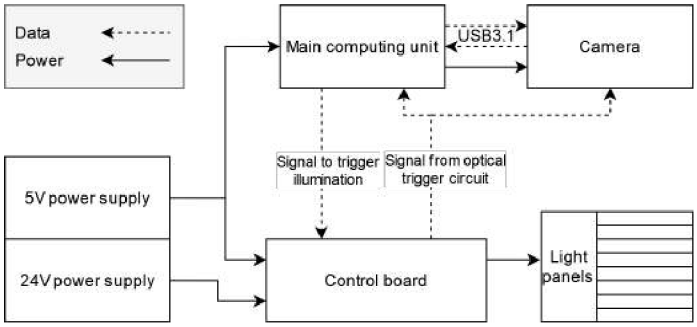

Figure 2.27: Power and data distribution along various system components.

Data and power distribution

In this subsection a summary of electronics used and the manner in which they are connected to each other is presented. To offer a overview of components used, as they were being introduced in various preceding subsections a list follows:

- Main computing unit [26];
- control board with Arduino Nano embedded controller[27];
- camera[28];
- 8 light panels;
- 24V power supply;
- 5V power supply.

The way these components are powered and communicate with each other is shown on Fig. 2.27. The camera uses USB 3.1 interface for both data communication with the main computing unit, as well as to draw power. The control algorithm itself will be described in a dedicated section, in this section the description will be limited to stating, that the control board transmits a trigger signal both to the main computing unit and camera and the main computing unit transmits a camera ready status to the embedded controller, which in turn triggers the light panels in a manner described previously.

This marks the end of the Design and realization of innovative biometry enabled devices chapter, the next chapter will focus on the control algorithms, data collection and evaluation.

[26] Available at: „https://up-shop.org/up-squared-series.html“
[27] Available at: „store.arduino.cc/products/arduino-nano-33-iot“
[28] Available at: „baslerweb.com/en/products/cameras/area-scan-cameras/ace/aca5472-17um/“

Chapter 3

Algorithms implementation and solutions evaluation

In this chapter all data related procedures performed for the systems introduced in the previous chapter, if applicable, control algorithms, data collection, data processing and data evaluation methods will be introduced and discussed. Publications related to this are, but not connected to any device developed as part of this thesis shall also be discussed here.

This chapter will primarily focus on processing related to the Biometry enabled line scanner, On the fly hand geometry and fingerprint recognition capable device and Biometric-enabled watchlists technology.

3.1 Biometry enabled line scanner

In order to enable the scanning function of a scanner and the subsequent recognition, two separate problems need to be dealt with software wise. The hardware needs to be controlled in a predictable and known manner, and the images need to be processed using biometry approach. Both of these problems have been tackled and were presented on the conference and published as part of the proceedings [34]*. In this sections, we will discuss the solutions to these two problems.

3.1.1 Control system implementation

Used hardware components and a manner in which they communicate and are powered has been discussed in Chapter 2, in this subsection the control algorithms which govern the behavior of the lights, linear motor, and linear scanner shall be discussed.

The scanning process on the basic level is fairly simple, as discussed in theory chapter, the linear scanner needs to begin an acquisition with such a frame rate, that due to movement speed of the linear motor, the number of lines acquired will match the desired spatial resolution. In order to achieve it, three systems must operate in tandem.

Light Control is a system, that handles the operation of the components related to the illumination. The high level operation was facilitated by a single board computer (SBC).

As explained in [34]*, the SBC was communicating with the XLC4-1 unit using Telnet interface over the Ethernet.

In the process of communication with the XLC4-1 unit, a script has been developed to transmit a command over the Ethernet interface to the unit. The script deals with

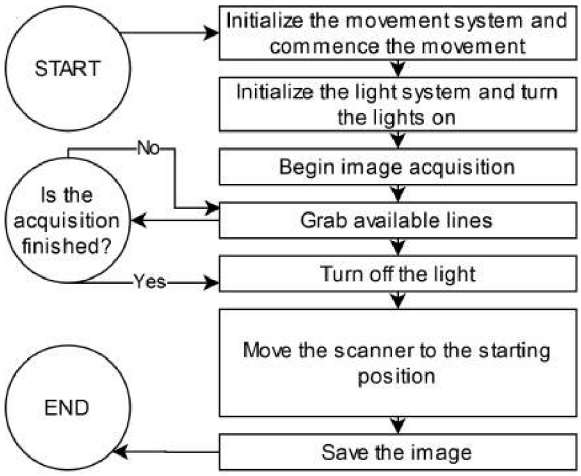

Figure 3.1: Control algorithm overview, as implemented for biometry enabled line scanner.

opening and closing of the network connection and adding start and stop sequences to a sent command. Following commands have been utilized to provide functionality as proposed in Fig. 3.1 and were transmitted using said script.

- *IY E 1400* is used to set the current used in the light module, were *E* defines a port and specifies a device used, and 1,400 is the required current in mA.
- *LC A 1* is used to change the state of the light module. *A* defines the port used (device independent) and 1 signifies that the module is to be turned on.
- *LC A 0* analogue to command above, but instead turns the illumination off.
- *IF* and *TE* are used to verify the system readiness and monitor the temperature of the illumination module.

Movement control system handles the operation of all components involved in realisation of the linear movement. The high level operation is again realized using SBC, but a micro-controller unit (MCU) is used to receive this instruction and pass it further down.

As explained in [34]* and keeping further elaborated in section 2.1.3 there is a relationship between the time passed and the distance traveled by the carriage given the setting of the device, it can be defined thusly:

$$d\,(t) = \frac{(4 \cdot (62,500 \cdot t))}{12,800}[mm] \tag{3.1}$$

Where time *(t)* is in seconds 62,500 is the frequency of the driving PWM, 12,800 is the number of steps per revolution of the stepper motor and 4 finally the carriage movement per revolution of the motor in mm.

We can use this function to precisely set the distance that is to be traveled at a constant rate of movement, thus ensuring that the spatial resolution captured by the camera will be as per the design.

The prototype for the database collection used a scanning time of 10 s, to cover a distance of 195.31 mm (discounting a ramp up and down time, which were used to avoid damaging the device).

A script has been developed which allows a simple control of the rail via the MCU. The script establishes a connection using the USB port and using serial communication protocol it can transmit a set of command codes. The commands are defined thusly:

- *Rxxxxx* a command that prompts the carriage move in one direction (to the right) for X milliseconds, where X is the 5 bytes (*xxxxx*) following the direction byte converted to integer, where *xxxxx* is being treated as a 5 digit long integer.
- *Lyyyyy* a command that prompts the carriage move in the other direction (to the left) for Y milliseconds, where X is the 5 bytes (*yyyyy*) following the direction byte converted to integer, where *yyyyy* is being treated as a 5 digit long integer.

If the the carriage reaches an end of the rail, a signal generated by the rail initiates an interrupt function. This function stops the movement of the carriage and resets the position counter if at the starting position or just stops the carriage movement if at the end position. After that should the command be received, which would move it past this edge position, control algorithm will disregard it and print corresponding message.

Linear scanner control The line scanner is controlled by SBC using the GigE interface and Basler Application Programming Interface (API). The script developed for this purpose handles proper initialization, data grabbing and image construction. All triggering and frame grabbing is performed using software scripts.

The prototype due to standard dimensions of the target performs acquisition of 5,120 lines at a rate of 500 lines per second. Considering the movement speed of the linear movement system, the chosen rates cause a line width to be 0.039 mm. Due to the apparent difference of the vertical pixel size from the horizontal pixel size, the image is upscaled by a factor of two in the x-axis prior being saved [34]*.

In order to perform this function a console application has been prepared, which calls the necessary functions from Pylon API for camera control and data retrieval and OpenCV library for image construction and saving. Fig. 3.2 shows a state diagram describing the general working principle of this application. As can be seen, it accepts only three arguments (which are passed via the aforementioned script) which set the save path for the image, an exposure time, which sets the duration for the aperture to be opened for a single line, and an acquisition rate, which specifies the number of lines to be captured per second.

The script provides high level functions with simplified interface which then call the application. It is the same script that also invokes the light and movement functionalities.

3.1.2 Data collection

Based on the hardware and software capabilities of the device and in accordance with the theory researched, the device has been set to generate images with the resolution of 6,144 × 5,120 px, this correspond to the scanned area of 417.1 × 195.3 mm due to mechanical properties of the device, mainly the object distance and focal length of the camera lens. The spatial resolution of this system should therefore be equal to 375 DPI in the scanner axis and 666 DPI in the movement for each image. Which in per the theory is below a customary and in some cases required spatial resolution of 500 DPI, but is above the minimum resolution required for fingerprint based recognition. Fig. 3.3 which has been also presented in [34]* shows two examples of image the device can produce (downscaled in the thesis).

The database collection was done without adhering to any specific condition, in an room with both daylight and artificial light present. Volunteers were invited to approach

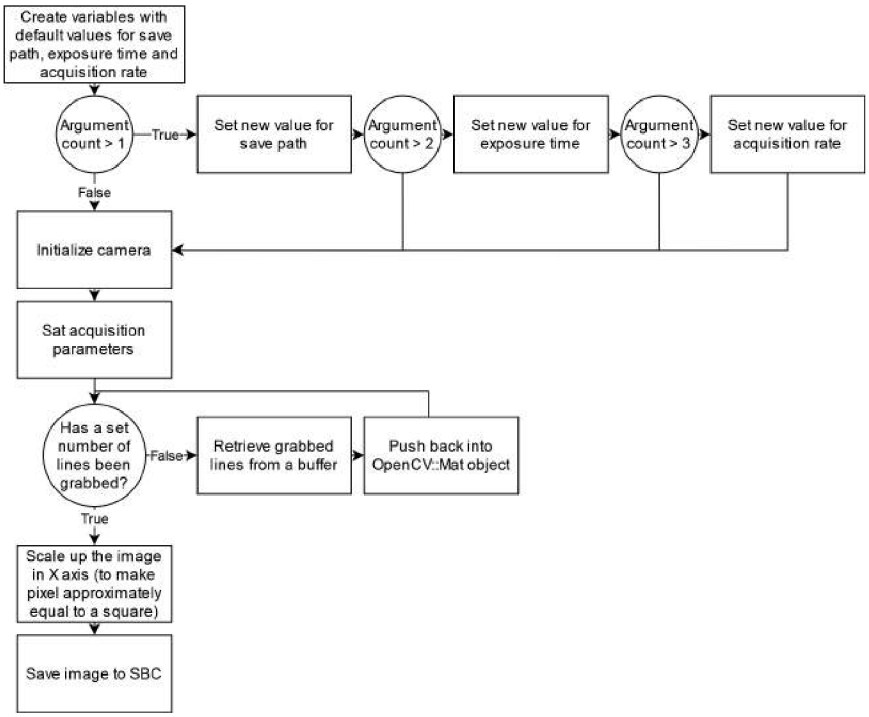

Figure 3.2: State diagram of camera controlling application.

the device and were briefly instructed on the use of the device. One demonstration has been performed by the operator at the beginning. Both instruction and demonstration focused on the preferred hand placement and stressed the importance of limited movement during the scanning process. Users were invited to undergo scanning of palmar and/or dorsal side of their hand. Some volunteered to offer both sides, some preferred to have only one side scanned.

During the course of the database collection, total of 145 images were acquired from 77 volunteers. Volunteers were of mixed gender in the 20-30 age group. 56 images contained a representation of hand from the palmar side, remaining 89 from the dorsal side. Database has been annotated only to differentiate images from individual users, but otherwise has been irreversibly anonymized. Second collection session has been planned but was canceled due to small number of repeat volunteers. The scaled down and cropped database acquired via this device can be seen in the Apendix B.

3.1.3 Data evaluation

It is essential to remind the reader that the device was developed primarily as a scanning method and that the biometric quality evaluation serves primarily to gauge the image quality, however, since it still fits the main theme of this thesis, evaluation both of the biomedical as well as biometry application shall be discussed in this section.

Biomedical application As stated in [34]*, the device is mainly intended to be used for monitoring hand diseases and disorders, such as eczema, as well as various skin growths, such as nevus, to monitor their size and area. In the previous section, it was shown that

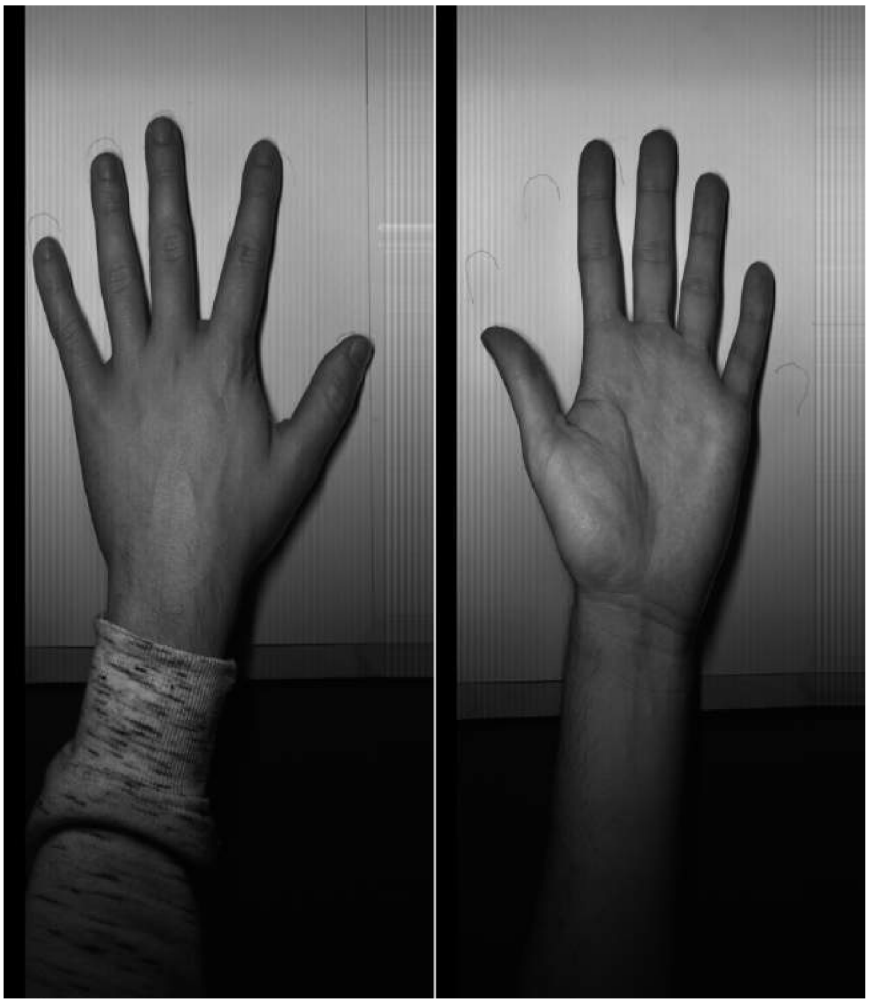

Figure 3.3: Example of dorsal hand surface image (left) and palmar surface image (right).

the device based on its optical design is able to distinguish ridges that have a width from 0.2 to 0.5 mm [24]. Based on fixed optics and calibrated system, precise measurements can be performed, which is a condition for this application.

To demonstrate the detail and a proof of utility. Two examples have been chosen from the data set. On Fig. 3.4 and 3.5 a detail of a dorsal side of the hand has been extracted and presents two nevi. Fig. 3.5 shows a result after the image has been locally normalized and adaptive thresholding algorithm has been applied.

Due to the system being calibrated, we can determine that a bounding box of 34 $\times$ 29 px and 24 $\times$ 19 px that can be used to surround the nevi corresponds to size of 1.23 $\times$ 1.13 mm and 0.87 $\times$ 0.74 mm respectively. The analysis of the size of the nevi can be immediately performed as well as saved and compared during subsequent scanning for any change ie. should the size or the number of the nevi change at a later scanning, this information would be easily accessible preventing an exhaustive manual observation and annotation.

In the same manner, second possible application it to monitor a wound recovery process. This can be seen in Fig. 3.6, where we can see an already recovered past wound. At this

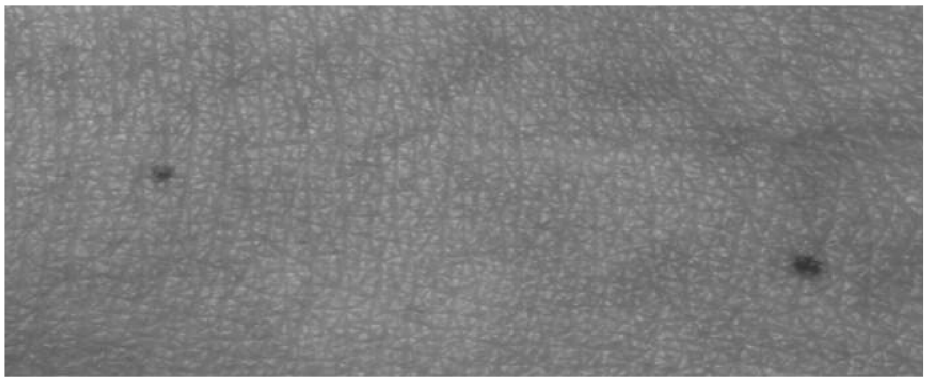

Figure 3.4: Detail of nevi (moles) from user 7 dorsal hand side, left hand.

Figure 3.5: Detail of nevi from user 7 dorsal hand side, left hand after thresholding.

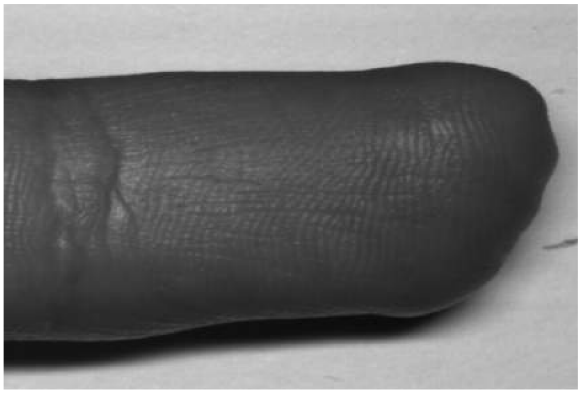

Figure 3.6: Detail of recovered wound from user 8 palmar hand side.

image, we can confirm that the the friction ridges are restored on the remainder of the finger. The rate of change, or history of the recovery, can be readily observed.

Fingerprint recognition as a biometry application is due to the spatial resolution a possibility.

Due to large scanned area (417.1 × 195.3 mm) and the relative small area the last digit of a finger occupies, a preprocessing, which aims to localize and subsequently extract fingerprints into a set of subimages, is required. As the fingerprint enhancement algorithms would be working with a needlessly large area otherwise.

Hand extraction process starts with coarse image alteration. In this step, a rectangle of 7,800 px to 4,300 px is isolated from position $x = 4{,}500$ px and $y = 225$ px. Cropped images contain only the hand and a part of the wrist. This area has been determined based on the guide marks on the device that suggest to user where the hand should be placed and the empirical observation. Images were rotated so that fingers pointed upwards, which is an expected orientation in some fingerprint quality assessment tools. In the last step of this preprocessing is application of a simple threshold algorithm, to ensure that finger tips are easily recognizable.

The localization itself is done using convolution-based localization algorithms, 5 position with highest response and with specified separation are returned the Region of Interest (ROI) is then extracted based on this information. After getting the finger images, histogram equalization and local thresholding methods are used to normalize the image and to increase the contrast of the ridges. Result can be seen in the Fig. 3.7.

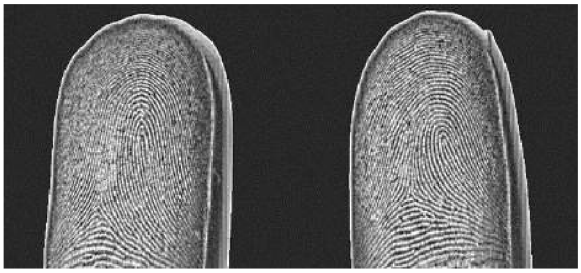

Figure 3.7: Example of two processed fingerprints from user 27.

280 images containing 230 fingerprints were acquired from the images. The fingerprints were then evaluated using VeriFinger (version 10) [1] and FiQiVi [19]. The median VeriFinger quality was 32 (maximum 60, minimum 10), FiQiVi showed a median of 38 (maximum 57 and minimum 19). For biometric recognition the VeriFinger developer Neurotechnology recommends the quality of at least 40 for identification and 30 for verification. This has also been experimentally verified by [4].

While the quality of the fingerprints may seem low, it is an impressive score. Firstly, the device to facilitate the large area acquisition used a 375 DPI resolution, which is lower than it is usual for fingerprint sensors. Secondly, unlike the almost absolute majority of sensors, this device does not use a scanning surface against which the finger is to be pressed. Thus, it reduces the available scanning area due to the finger geometry, as it is not pressed against a glass surface, as well as decreases the contrast due to the absence of total internal reflection. This concept can be seen in [69]. Thirdly, this evaluation contained also thumbs which were usually acquired at an angle. Lastly, only basic image pre-processing was done with no fingerprint enhancement algorithms being used. Therefore, the fact that despite all the disadvantages, the device still manages to meet the fingerprint biometry baseline is a relatively strong proof of its viability.

Hand geometry recognition as a biometry application can also be performed as discussed in [34]*. The images collected by the scanner are of sufficient quality to perform recognition based on hand geometry if other biometrics are unavailable, for example, due to injuries such as a burn or skin diseases.

[1] Available at „https://www.neurotechnology.com/mm_100_sla.html“

As the acquisition process is calibrated and the images can be used to perform measurements of the target object, absolute measurements may be used as a feature set for the recognition algorithms, such as described by [77] and [105] or relative measurement, as described by [90].

Majority of algorithms perform segmentation and work with a binarized image of the hand, which these images can also be converted to. The feature set can then be constructed from such features as outlined in the Fig. 3.8, given the knowledge that pixel's height corresponds to 0.039 mm and the pixel's width corresponds to 0.036 mm. Using this information, any selected length of the image may be recalculated into a real-world dimension.

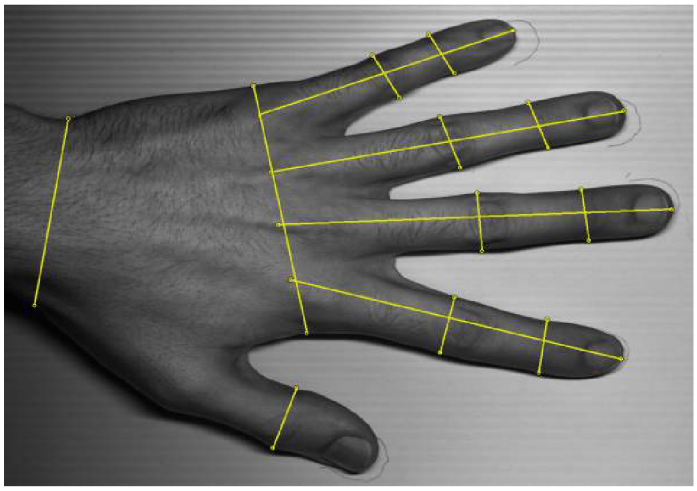

Figure 3.8: Proposed approach to hand geometry recognition.

3.2 On the fly biometry enabled device

In order to develop an on the fly biometry recognition enabled system, several separate problems need to be resolved software wise. The hardware needs to be controlled in a predictable and known manner, the images need to be processed using biometry specific algorithms and finally the validity of the methods needs to be confirmed by evaluating the results of the biometry recognition. In this sections, we will discuss the solutions to these problems, as they were published by the thesis author in IET journal [31]* with focus on on the fly hand recognition, and international patent [33]* with focus on the fingerprint recognition and acquisition.

3.2.1 Control system implementation - On the fly hand geometry device

Used hardware components and a manner in which they communicate has been discussed in Chapter 2, in this subsection the control algorithms which govern the behavior of the camera, for the purposes of data collection shall be discussed.

As this system has been conceived primarily as tool for the academic research, as such, the control algorithms have only a minimum necessary functionality, to facilitate the data collection process. See below the necessary steps.

Camera initialization which included enabling the HW triggered acquisition, exposition time, gain and gamma. Except for trigger mode, which was necessary for the proper device operation and exposition time, which by design needed to be <600 μs as per the design discussed in Chapter 2, all the settings have been set manually to achieve visually optimal performance, prior the database collection.

Data storage initialization consisted of object creation, which kept the information about the user identifier, as well as the index of the image to be taken. This object would later facilitate the image storage into the database with a file name indicative of the above mentioned facts.

Image collection was a process, which facilitated the retrieval of an image from a camera, once available. Since, the process occurred based on the external stimuli (the camera buffer indicating a change in state), it's main task was to await this event. For a better idea of the process, see the overall state diagram of the collection algorithm on Fig. 3.9

Image saving was performed based on the current state of the data storage object. Depending on the identifier of the user as well as the index of the current image, the image was saved in to the database.

The description above outlines only the main steps, a number of sub-tasks was necessary to generate a database, in which the images of same user could be identified and differentiated from those of other users. For an overview of the process, see Fig. 3.9.

3.2.2 Data collection and processing

As discussed in [31]*, the the data pipeline consisted of two distinct tasks, a database collection and database annotation. See below for more details.

Database collection consisting of multiple acquisitions of 99 users has been collected. The ideal hand movement was demonstrated to each user, and the requirements as they related to goals were briefly explained. The user was instructed to walk towards the device at a comfortable pace. As the user approaches the device, he/she must stretch and extend the arm away from the side, aiming the palmar region of the hand towards the acquisition gate. As the user passes the system, the hand must pass through the acquisition gate to be captured by the camera. It has been observed, that a small number of users failed to successfully follow the instructions during the first run, resulting in an image of poor quality, or no image at all, depending on the nature of the fail. After a verbal instruction was issued in order to avoid the error in future attempts, no user failed to follow the instructions again. The instructions as presented, were focused on the following key points:

1. The palm should be flying in parallel with the acquisition gate as it passes through. In other words, the hand should have the same orientation as the gate plane.
2. The fingers are to be stretched and in line.
3. The speed of the hand movement through the device should be comfortable, it may be variable.

The required hand position with respect to the acquisition gate as per Points 1 and 2 is shown in Fig. 3.10

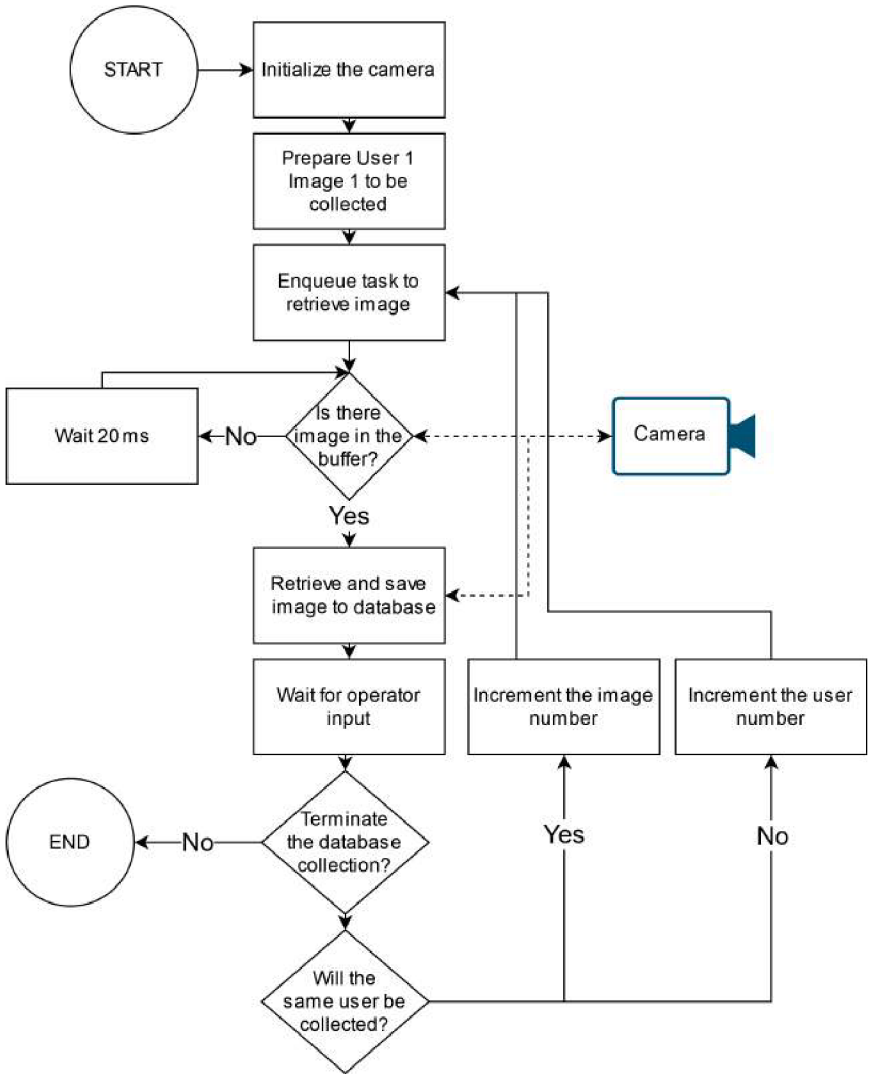

Figure 3.9: Database collection control algorithm for the on fly hand geometry recognition device.

At least five hand images were requested to be taken from each user, and additional acquisitions were encouraged to those willing. A total of 573 images have been captured, and nine images have been removed due to incorrect/unexpected triggering, leaving 564 unique images for the analysis. A sample of image collected after a correct use and trigger is shown in Fig. 3.11. Examples of images removed due to unexpected triggering and an incorrect triggering, can be seen in Fig. 3.12. The former was caused by a volunteer who swiped his hand to see „How it works“ and latter by student who instead moving his hand from back to front moved it from front to back, resulting in an out of focus image.

The database by the design of the collection has been anonymized, with each entry having a unique number in the format *Usr_XXX_YYY.png* where *XXX* signified a three digit index of a user, starting at 001 last being 099 and *YYY* being a three digit image index which corresponds to the scanning number index starting at 000 and ending at a number

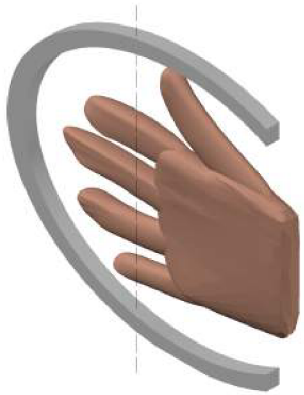

Figure 3.10: The required hand position with respect to the acquisition gate during acquisition.

Figure 3.11: Image of hand as correctly collected by the on-the-fly device.

dependent on the number of acquisitions user underwent, i.e. should the user number 55 agree to be scanned 10 times, the name of the last image would be *Usr_55_010.png*

Database annotation has been performed manually, as the goal of the publication [31]* was to evaluate the practicality of the proposed on-the-fly system for biometric identification and the automatic feature generation was out of the scope of the paper.

Significant points on the hand image have been annotated manually, to establish a ground truth scenario to benchmark against. For the improved visual inspection and increased annotation accuracy, the images have been enhanced using the CLAHE algorithm [117]. CLAHE is the algorithm, that has been used in [86] for fingerprint enhancement. Thirty-seven significant points have been annotated in each image of the palmar hand side using the VGG Image Annotator (VIA) proposed by Visual Geometry Group (VGG) [28].

Figure 3.12: Image of incorrect triggering collected by the on-the-fly device (left). Image of unexpected triggering collected by the on-the-fly device (right).

Significant points at the hand's edge have been chosen to ensure their simple identifiability, uniqueness and limited variability caused by hand movement. This led us to choose the interphalangeal creases to identify the individual finger joints. Significant palmar creases indicate the sides of the palm as they reach the hand edge, and the wrist crease is used for wrist width information. These chosen points were easily identifiable across the hand database, as can be seen in Fig. 3.13.

Fig. 3.13 shows both the effect of CLAHE enhancement and the chosen 37 feature points, which mostly correspond to the edge positions of the finger creases and palm creases and the tips of the individual digits. Annotation was somewhat time-consuming. To annotate 564 valid images and took four people approximately four hours each to complete, where each image took approximately 90 seconds to annotate.

During the annotation stage, nine images were identified and discarded, as they contained seriously out-of-focus hand images (incorrect triggering) or no hand at all (unexpected triggering). Overall, however, the images have been judged to be of acceptable quality especially for hand geometry recognition, where the requirements on resolution are much lower than on images intended for fingerprint recognition. The scanned objects were in focus, and the acquisition gate defined the position. The eliminated images were caused exclusively by incorrect user behaviour during scanning.

Upon closer observation of the images, as shown in Fig. 3.13, we noticed an outline of the hand next to the expected hand image, making a ghost-like appearance. The mirror's glass surface caused this ghost effect as a fraction of light didn't pass through the glass and instead has been reflected. This was an important observation, and the device iteration for fingerprint recognition would instead incorporate a mirror with a reflective top surface, rather than a reflective bottom surface, to the reflection and remove the optical artifact [31]*.

3.2.3 Recognition classifiers

As discussed in [31]*, the end of previous section produces, data which in accordance with the biometry theory should hold the information that can be used to identify one hand from another. In this section, the actual biometry feature selection, recognition and evaluation shall be discussed.

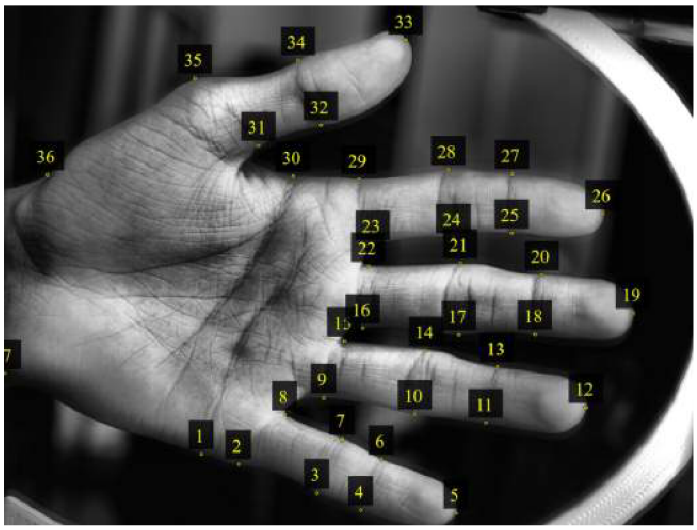

Figure 3.13: Annotated image of a hand enhanced by CLAHE algorithm.

The feature selection as discussed in theory, is a process, where we identify parameters or a group of parameters that show a non-random correlation between their value and the user's identity. During hand geometry identification, the distances between the chosen points on the hand's surface are used to determine the user's identity [24].

As discussed in previous chapter, $\boldsymbol{n} = 37$ significant points has been labelled on the image of the palm, as shown in Fig. 3.13, these points have been annotated as **p1** − **p37**. Each of these points is described as coordinates in the image in the form of $\boldsymbol{p}$=$(p_{(n_0)}, p_{(n_1)})$, where $p_{(n_0)}$ corresponds to the $\boldsymbol{x}$ coordinate and $p_{(n_1)}$ corresponds to the $\boldsymbol{y}$ coordinate on the image. It is self evident that a distance among some of these significant points can serve as a biometric feature.

Exhaustive feature selection, where each of the n feature points is connected to every other point, would yield us $T_n = 703$ features, as per the n^{th} triangular number, see Eqn. 3.2.

$$T_n = \frac{n(n+1)}{2} = \frac{37(37+1)}{2} = 703 \tag{3.2}$$

However, even a cursory observation indicates that many, if not most, of these features are arbitrarily defined by the random position and orientation of the individual fingers. For example a distance between tips of individual fingers hold very little biometric information. Aside from tenuous indications of the user's phalangeal dexterity, they bear no significant relevance to the user's identity. Therefore, a subset of distances has been chosen as depicted in the Appendix A.1, while the rest have been discarded as unnecessary or ineffective. For example, ineffective distances between feature points **p13**->**p18** and **p22**->**p33** have been discarded at the pre-processing stage.

The Euclidean distance between the feature points p_n and p_m is measured and saved as a unique feature f_n. Doing so for each of the point pairs defined in Appendix A.1 creates feature set 1 (Fs_1) of features f_0-f_{130}. By normalising Fs_1, we get feature set 2 (Fs_2). Fs_2 is a set of relative distances calculated as a quotient of the features in Fs_1 divided by two × lengths of the middle finger. This can be approximated by the summation of the Euclidean distances between points **p16**->**p19** and **p19**->**p22**, as shown in Eqn. 3.3.

Features f_{132}-f_{262} are, therefore, for values of $n = 0...130$, defined as in Eqn. 3.3. The divisor used for Fs_2 has also been used as feature f_{131}.

$$f_{131} = \sqrt{\sum_{i=0}^{1}(p_{16_i} - p_{19_i})^2} + \sqrt{\sum_{i=0}^{1}(p_{19_i} - p_{22_i})^2} \tag{3.3}$$

$$f_{n+132} = \frac{f_n}{f_{131}}, n \in 0...130 \tag{3.4}$$

In total, there are 263 features selected as relevant and to be used in the classification algorithm.

By comparing Apendix A.1 with Fig. 3.13, we can observe that most features consist of distances between individual feature points within individual digits and distances between neighboring feature points. This selection has been decided upon as these distances show the lowest intra-class variability due to the hand and individual finger positions as per the theory.

Classification algorithm has been chosen based on the database size and the number of features used. Either the support vector machine (SVM) [16] classification algorithm or random forest classifier (RFC) [9] was deemed a good option. In order to avoid over-fitting, the RFC has been chosen as a more suitable classifier for our particular application due to categorical and unbalanced data. As indicated in [72], the RFC is particularly efficient in these cases.

Fig. 3.14 shows that the approach to train the classifier and the technique used to optimise its parameters are straightforward. We have iteratively adjusted the Number of estimators parameter for various Maximum depth parameters. Once the increase in the number of estimators introduces no significant improvement in performance, the process is stopped. The classifier adjustment as well as performance measurement has been performed using two cross validation approaches.

The entire feature set has been used to train the random forest classifier. Considering the database size (564 entries) and the feature set size (131 features), it was feasible to perform the Leave-One-Out Cross Validation (LOOCV) [92] for every entry in the database. The use of LOOCV was previously used successfully for biometric application in [15].

The parametrisation was focused on adjusting the RFC's critical parameters, namely the maximum depth and the number of estimators.

The final parameters that have been used in our classifier were implemented using Scikit-learn [75] and its Random Forest Classifier and can be summarised as follow:

- Maximum depth: 11
- Number of estimators: 150
- Maximum number of features considered for the split: 2
- Random state: 1
- Features removed: none

Fig. 3.15 shows the performance of the classifier during the parameter optimisation stage and its relation to the number of estimators.

Feature pruning techniques have been utilized in an attempt to improve the classifier performance. Prominent features are selected on the basis of the mean decrease in impurity,

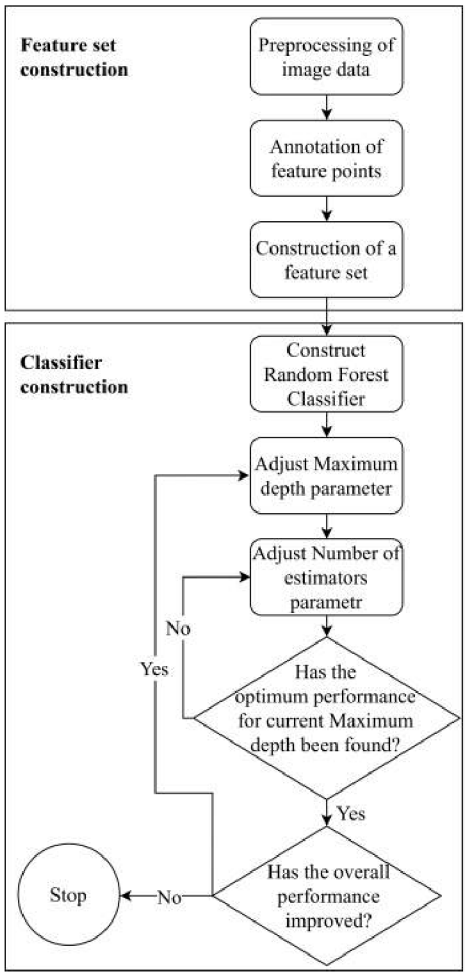

Figure 3.14: Parameter optimisation of the classifier.

which is calculated using Scikit-learning tools [75]. During the LOOCV the ratios of individual importance are summed for each constructed classifier. Fig. 3.16 shows the sorted array of the sums of the importance ratios. The red bars indicate the features of relative distances and the blue bars refer to the absolute distances.

There are two observations worth noting. We noticed that some features are over three times as important as the others. It is also to be noted that the relative distances clearly offer less benefit than the absolute distances.

Regarding our database, even the least important feature contains information contributing to improving the overall classification performance. Therefore, we have concluded that while the importance of the individual features varies, no features will be removed on the basis of these observations.

It is worth pointing out that removing the least important parameters has led to a decrease in the number of estimators necessary for achieving optimal performance.

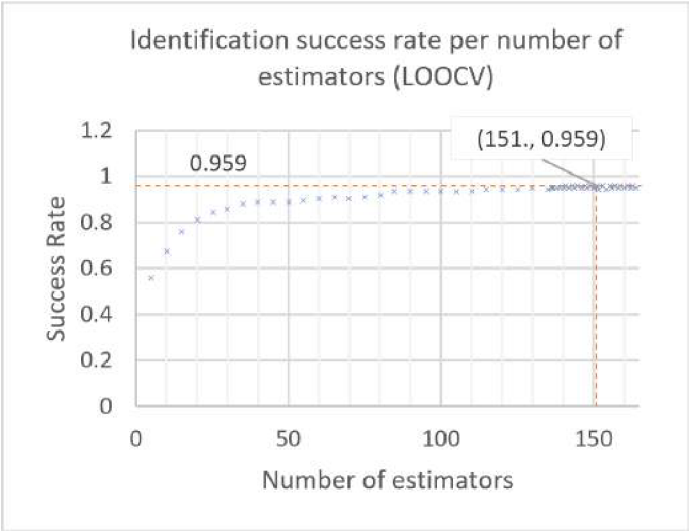

Figure 3.15: Classifier optimization using estimator number.

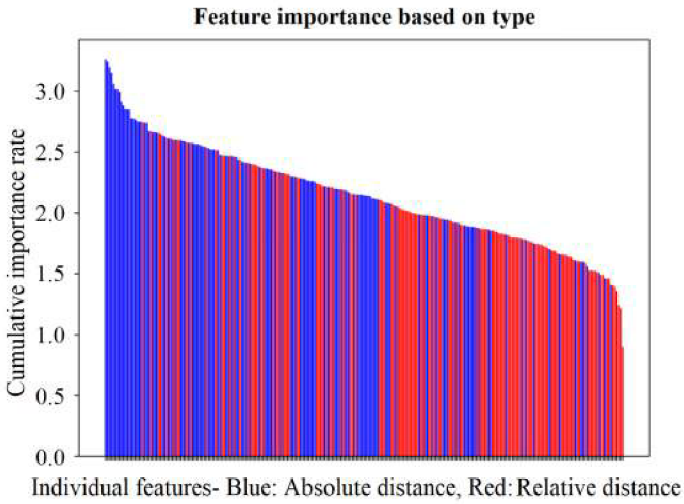

Figure 3.16: Importance of individual features. Absolute distances (blue) and relative distances (red).

3.2.4 Recognition evaluation

As discussed in [31]*. In Fig. 3.15, we can see the final performance of the classifier. As the classifier is a set of decision trees, the results take only the form of correct (TP – True Predictions) or incorrect (FP – False Predictions) classifications. We have calculated the success rate (SR) as defined in next section and standard deviation (SD) [6] as metrics of our system from the classifications. Performance measures such as false accept rate, false reject rate and equal error rate, cannot be calculated for our classifier.

Table 3.1: Best performing classifier results

CV method	TP	FP	SR [%]	SD
LOOC	542	23	0.9593	0.1976
RS K-Fold	2712	113	0.9600	0.0146

The performance has been measured using the LOOCV and a Repeated Stratified K-Fold cross-validation (RS K-Fold) [55] with 5 folds and 5 repetitions. The mean values of SR and SD shown in Tbl. 3.1 are considered for the individual iterations.

Leave out one Cross-Validation is a type of validation to measure the performance of the chosen recognition algorithm. It has been chosen as one of two, to give a more complex idea.

The LOOC RFC is trained using matrix X_{train} which consists of the vectors of features of each database entry except one and a vector of results y_{train}, formed from the entirety of our database with one entry left out. The entry left out is used as a test set i.e. X_{test} and y_{test}. Using the trained classifier, a prediction y_{pred} has been calculated based on the full feature entry X_{test}.

$$y_{pred} = RFC(X_{test}) \tag{3.5}$$

The prediction y_{pred} was compared to the expected result y_{test} and the result has been noted as a change in score according to Eqn. 3.6. Where we increment the true prediction (TP) index if the prediction matches the expected result and false prediction (FP) if it does not. The same steps are repeated for every entry in the database.

$$\{TP, FP\} = \begin{cases} \{TP+1, FP\} & \text{if } y_{pred} = y_{test} \\ \{TP, FP+1\} & \text{if } y_{pred} \neq y_{test} \end{cases} \tag{3.6}$$

The success rate (SR) is the ratio of TP and the total number of entries. The TPs and FPs are then used for the calculation of the mean values of SR and SD.

Repeated Stratified K-Fold Cross-Validation is the second validation algorithm used.

As the LOOCV has a large variance according to the theory, the 5-Fold Stratified Cross-validation was repeated five times to offer a result with smaller variation at the cost of using a smaller training set.

Repeated Stratified (RS) K-Fold cross-validation splits the data into five folds (groups), where one fold is used as a test set data while the rest is used as training data. The advantage of RS K-Fold is that the data is split in a way that takes into consideration the different number of entries per individual class [55].

Utilising the Scikit-learn and its Repeated Stratified-K-Fold implementation *sklearn.model_selection.RepeatedStratifiedKFold*, TP and FP have been calculated using the Eqn. 3.6 mean SR and SD across five iterations have been calculated and depicted in Tbl. 3.1.

Computational benchmark has been performed, in order to evauluate the possible utility as the on they fly biometric device.

Table 3.2: Identification time in 565 entries large database.

Measurements No.	Average time [ms]	SD [ms]
565	56.28	0.18

The classification algorithm has been benchmarked using an x64 based PC, with Intel® Core™ i7-8750H 2.2GHz and single-slot 8GB 2,667 MHz SODIMM. No optimisation or acceleration was performed, and the run time of the Python script was measured using Python function *time.perf_counter()*. The measurement was taken during each loop of the LOOC verification. The recognition time in Tbl. 3.2 represents the time it takes to determine a person's identity given the database of 565 people. It is worth repeating that in this paper, we have not developed algorithms that determine the position of the feature points, as the ground truth hand-annotated database was used. The average time figure only describes the classification time and may approximate the expected performance of a fully autonomous system.

Performance comparison is challenging to make. To the best of our knowledge, the proposed on the fly system is novel, especially regarding the acquisition process, for hand-related biometrics, there is no publication that could serve as a suitable comparison. Only available direct comparison can be made to on the fly facial recognition systems, however, due to different biometry and acquisition method used, the comparison has very low information value.

The largest issue, is the impossibility to directly compare the recognition algorithms from other publications with ours. This is due to majority of available publications having used stationary, touch-based system to generate the database (often including structural pins for hand guiding), making our unrestrained hand database unsuitable for algorithms, expecting restrained hand images. It has been therefore decided to compare the overall performance of our system, i.e. the identification success rate, with the identification success rates presented in the other publications. We believe this comparison gives a generally fair comparison with state of the art, despite the differences.

Despite additional complexities added by the on-the-fly acquisition, our system yields a comparable performance. In State of the Art section in theory, Tbl. 1.1 we compare our results with the results reported in other publications. It is worth mentioning that only [93] and [53] utilise contactless systems, and none is capable of on-the-fly acquisition.

The reported results in [2][7][89] are similar to our proposed system. However, they used a small-sized database. Better performance than our system has been reported in [5]. However, the focus of that paper is on the classification algorithms. In contrast, our work's novelty lies in the acquisition process and overall improvement of the methodology of biometric data acquisition, with classification being used solely as a demonstration of its viability. In follow-up work, we focus on classification performance improvements.

The work [53] uses a much more extensive database, making the results more concrete, and we cannot make a specific claim that our performance would be better had our database size been similar. However, we have included this comparison. It is only one of a few that performs contactless recognition based on hand geometry, thus allowing us to get a clearer idea of how the systems compare.

Tbl. 1.1 shows on-the-fly biometric performance comparison available in the literature. While the finger vein biometry system appears to offer slightly improved performance,

we feel that the user comfort and the corresponding throughput would be decreased. In contrast, the on-the-fly face recognition is more user friendly but offers lower performance.

As the acquisition methods are different, it is worth pointing out that the use case may vary as the different approaches offer different advantages and disadvantages. The most apparent advantages of the proposed system have already been mentioned in the introductory section. The fact that the hand does not touch the device makes the whole acquisition process more hygienic, which leads to less maintenance as the device does not have to be cleaned regularly. Touchless submission of biometrics also addresses the issue of having to clean the various optical parts that are being touched. The on-the-fly approach makes the whole recognition process faster, as more people can pass through the system per unit of time.

Among the disadvantages of the on-the-fly system, we mostly need to consider the increased user cooperation requirement. Despite the fact that using the on-the-fly system is intuitive and readily taught, the constraining nature of the scanning process could lead to recognition failure due to incorrect hand position at the time of the scanning, fast passing speed, wrong hand angle, or improper absolute hand positioning. From the hardware standpoint, it is also fair to mention that strong external light sources are required for a short exposure time to illuminate the hand.

While our system performs comparably to other hand geometry-based systems, the biometry itself may be unsuitable for some applications. As explored in other works, the biometric entropy of the hand geometry is low when compared to the other biometrics such as fingerprint or iris recognition.

3.2.5 Multimodal hand based biometry recognition

The on the fly biometry enabled device developed as part of this thesis, offers an intriguing opportunity, which has been discussed in the theory chapter. Since the images obtained are of the entire hand, which can be used for the hand geometry recognition as discussed in previous sections and [31]* and at the same time the device has been developed and tested for the purposes of the fingerprint recognition, proved by the accepted patent [33]*. We have means to utilize both of these feature sets and employ the principles of multimodal biometry as discussed in theory and [29]*.

While the utilization of images from the on the fly device for purposes of multimodal biometry has not been performed because of the pending patent. In [59]* the multimodal biometry approach has been explored for fingerprint and face recognition and as expected, the combination of two biometries has a positive effect on the overall recognition rate.

As stated in [59]* for purposes of rating multiple modalities.

An appropriate metric for the biometric-enabled recognition is the Doddington risk categorisation which is well studied for various modalities [20][79], including the databases which we used in the experiments [106].

Databases used as well as reasoning for using them is also provided in [59]*.

In this thesis, we simulated the watchlists using databases of both fingerprint and facial biometrics. FRGC V2.0 [76] and faces-in the-wild (LFW) databases [46] contain facial images of different people in various qualities. The large-scale facial FRGC2.0 (50,000 recordings with 4,007 subject sessions) and LFW (13,000 images of unconstrained web photos of 5,749 different subjects) databases are well suited for the watchlist modelling for the following reasons:

1. They contain facial images of different quality;

2. they reflect different sources of data about the person of interest, such as surveillance; camera.
3. they represent satisfactory statistics as a model of the watchlist landscape in Doddington risk categorisation.

The chosen fingerprint database is the NIST Special Database 4 fingerprints (NIST SD4) [102] due to its availability, size and quality. We also appreciated that the database is evenly distributed over each of the five classifications with 400 fingerprint pairs from each class. The NIST SD4 contains eight sets of 250 fingerprint images to a total of 2,000 images which is large enough for our experiments [59]*.

We used the commercially available biometric recognition software from Neurotechnology: Verifinger for fingerprints (one of the most accurate in the FpVTE 2012 NIST competition [101]) and Verilook for facial images (one of the most accurate in the facial recognition vendor Test 2013 NIST competition [42]) [59]*.

The thesis aims towards developing a suitable watchlist framework, and a general formula for the cost of using this framework is proposed as [59]*.

$$Cost = C_{FN} * FNR * P_G + C_{Fp} * FPR * (1 - P_G) \tag{3.7}$$

where C_{FN} and C_{FP} is a cost of false negative classification or false positive clasification respectively, i.e. the cost of person on watchlist being admitted through the check and cost of person not on watchlist being detained for example. P_G is the probability of random user being genuine. The FNR and FPR are as can be expected dependent on the threshold of the recognition algorithm [59]*.

[59]* provides an interesting observation about the fusion of fingerprint and face based biometric data, (which however for certain applies only for the selection of databases, as used in the paper). Given a weaker threshold for NIST fingerprint database, a risk of FRR is eliminated, at the cost of increased FAR, however in combination with LFW or FRGC face database with stronger threshold, the FAR is eliminated, meaning that together, for these databases, perfect recognition would be achieved, demonstrating in this case advantages of multimodal biometry.

Aside from the increased recognition rate, the multimodal approach offers additional significant advantage, which would also apply to the on the fly hand recognition device, and that is an improved presentation attack detection. As discussed in [22]* and [44]* the spoofing of both fingerprint and hand geometry recognition systems is an existing issue. By combining two biometries, the complexity of necessary presentation attack instrument increases, thus making the device more resistant to this sort of security breach.

Conclusion

This thesis sets out to explore and expand the field of the hand based biometry. To achieve this goal, it asked a question whether the current way of acquiring data, is the best we can do, and proposed, that perhaps there is a better way.

Thesis researches the ways in which the human hand and it's biometric features have been scanned, how the images are processed and how the biometric data are extracted and evaluated. The current method have been considered and the biggest shortcoming has been identified not in the individual algorithms, but in the philosophy of data retrieval in general.

Three devices have been developed as part of the thesis in an effort to either improve on the data that can be used for the recognition or by offering an added benefit to the acquisition process.

The development of the stereo vision based 3D object scanner would fall into the former category. It aimed to offer a 3D scanning approach, that would allow to extract not only the 2D hand geometry data, but 3D as well. While the object scanning approach itself has been verified and published in technical report, as part of a project, its added value to biometry remained unconfirmed as more promising avenues were pursued.

Such as Biometry enabled line scanner, which offered a functionality, where the scanned area would be adjustable, without having to change the camera sensor. This function has been facilitated, using the linear motion system and line scanner. Thereby allowing for a constant spatial resolution, while increasing the total resolution by adjusting the distance traveled by the line scanner.

In order to verify it's utility as a biometric device, a database has been collected and fingerprint quality evaluation algorithms have been applied. The use as a hand recognition based device and medical device has been discussed and its use supported by supplied theoretical research. The device, the database construction and the biometric and medical applications have been presented and published at a conference.

The scanner also gave way to the research avenue which forms the core of this thesis. If biometric data can be acquired using a moving camera, an inverse should also be possible, to extract biometric information from a moving object, i.e. to extract the hand related biometric features on the fly.

This thought has been developed over multiple iterations. The theory has been developed to define requirements on a system based on an assumed hand movement characteristics. The proof of concept prototype has been constructed to demonstrate the viability of this approach based on the developed theory. The laboratory prototype and later a functional prototype have addressed the issues of previous versions.

The laboratory prototype was able to facilitate the hand recognition, verified on a database constructed using this device, base on the fingerprint quality algorithms, it would facilitate only a limited fingerprint recognition. Where the first claim has been published

in the IET Biometrics Journal and the later was submitted and accepted by a conference. The IET publication offers a comprehensive comparison to other on the fly recognition algorithms as well as other generic hand geometry based recognition methods, and it performs comparably despite the added complexities supplied by the on the fly aspect.

The functional prototype, ie. the final version was able to provide images at quality sufficient for reliable on the fly fingerprint recognition. Description of this device and the functionality has been submitted as a patent application which has been later accepted.

As the above mentioned device has successfully demonstrated an ability to acquire multiple biometry feature types (hand geometry and fingerprints). It by definition is capable of supporting a multimodal biometric recognition. The advantages in terms of improved recognition and presentation attack detection have been discussed and published as part of IET Biometrics Journal publication and book chapters.

The thesis proposes, that perhaps, the current contact based data acquisition, when it comes to hand based recognition, now serves as a bottleneck to the mass recognition, it presents hygienic concerns and overall does not technologically fit into 21^{st} century and touch based mass identification system definitely does not belong to current/post COVID era. This thesis therefore decided to focus on implementing touchless on the fly principles into the biometric process.

Fulfillment of the dissertation goals

The theory, research into the state of the art and proposed development paths have been published as a chapter in a Springer published book [30] a later revised and expanded as a chapter in IET published book [29].

I. Development of biometry enabled optical scanning devices
   As a part of this thesis, three optical systems have been developed with the intent of utilizing them for a biometric application.

   Stereo vision based 3D object scanner a device developed for a 3D object scanning, reported as a part of a project [40]. While used only for weapon scanning where the 3D reconstruction has been demonstrated, it has been deemed unsuitable for the hand based recognition, and was not pursued further in favor of more promising development avenues.

   Biometry enabled line scanner was designed and constructed to offer a scalable scanning device usable for medical or biometry use. The concept has been presented on a BIODEVICES conference and published in it's proceedings [34]*.

   On the fly enabled biometric system is a novel solution to data acquisition and biometric recognition. It's hand recognition capabilities were published in IET Biometrics [31]*, its fingerprint recognition viability has been submitted to and accepted by an International Conference on Mechatronics and Electrical Systems [35]*. Along with an industrial partner the concept has been developed into a functional prototype [32]*. based on this prototype a patent application has been submitted and the patent has been awarded [33]* and has been introduced to the security community during INTERSEC 2020.

II. Applications and evaluation of novel biometry enabled devices

Stereo vision based 3D object scanner required a development of control algorithms for the movement of the camera mount along the U-shaped pipe. As well as the communication prototype which was used to call the low level control functionalities as well as the data retrieval from the camera. Description of these algorithms has been published as a part of the technical report for the related project [40].

Biometry enabled line scanner required a development of control algorithms for linear motion systems, as well as scripts for operation of camera system and illumination. The fingerprint quality algorithms have been utilized, to demonstrate its biometry utility, and its medical application has been presented on a BIODEVICES conference and published in it's proceedings [34]*.

On the fly enabled biometric system required an implementation of control algorithms for the camera system, as well as the firmware for triggered illumination. The viability of the laboratory prototype as an on the fly fingerprint recognition device has been demonstrated by performing a fingerprint quality evaluation on the database collected on the device. The paper describing the process and results has been submitted to and accepted by an International Conference on Mechatronics and Electrical Systems [35]*.
The viability of the laboratory prototype as an on the fly hand geometry based recognition device has been demonstrated using novel hand geometry based recognition algorithm. The algorithm has been proposed, developed, verified on a database collected on the new device and compared to the state of the art in the IET Biometrics journal [31]*.
It has been proposed, that due to device's ability, to collect several biometry feature types, it is a viable mean of performing multimodal biometry. The advantage in terms of improving the recognition rate as well as utilisation of biometric watchlists has been presented and accepted as a publication in an IET Biometrics journal [59]*. The use of multimodal biometrics as a presentation attack prevention has been discussed in a book by IET publications [44].

Bibliography

[1] Abe, N. and Shinzaki, T. On the fly finger knuckle print authentication. In: International Society for Optics and Photonics. *Biometric and Surveillance Technology for Human and Activity Identification XI.* 2014, vol. 9075, p. 907504.

[2] Adán, M., Adán, A., Vázquez, A. S. and Torres, R. Biometric verification/identification based on hands natural layout. *Image and Vision Computing.* Elsevier. 2008, vol. 26, no. 4, p. 451–465.

[3] Al Fiky, F. M. and Ageed, Z. S. A new features extracted for recognizing a hand geometry using BPNN. *International Journal of Scientific & Engineering Research.* 2014, vol. 5, no. 9.

[4] Alsmirat, M. A., Al Alem, F., Al Ayyoub, M., Jararweh, Y. and Gupta, B. Impact of digital fingerprint image quality on the fingerprint recognition accuracy. *Multimedia Tools and Applications.* Springer. 2019, vol. 78, no. 3, p. 3649–3688.

[5] Angadi, S. and Hatture, S. Hand geometry based user identification using minimal edge connected hand image graph. *IET Computer Vision.* Wiley Online Library. 2018, vol. 12, no. 5, p. 744–752.

[6] Bland, J. M. and Altman, D. G. Statistics notes: measurement error. *Bmj.* British Medical Journal Publishing Group. 1996, vol. 312, no. 7047, p. 1654.

[7] Boreki, G. and Zimmer, A. Hand geometry: a new approach for feature extraction. In: IEEE. *Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05).* 2005, p. 149–154.

[8] Bradski, G. The OpenCV Library. *Dr. Dobb's Journal of Software Tools.* 2000.

[9] Breiman, L. Random forests. *Machine learning.* Springer. 2001, vol. 45, no. 1, p. 5–32.

[10] Chakraborty, S. and Rao, K. Fingerprint enhancement by directional filtering. In: IEEE. *2012 9th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology.* 2012, p. 1–4.

[11] Chen, H., Valizadegan, H., Jackson, C., Soltysiak, S. and Jain, A. K. Fake hands: spoofing hand geometry systems. *Biometric Consortium.* 2005.

[12] Chen, H. and Bhanu, B. Human ear recognition in 3D. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* IEEE. 2007, vol. 29, no. 4, p. 718–737.

[13] Chennankara, B., Xu, W., Lin, F., Drake, M. and Fiddy, M. Optical fingerprint recognition using a waveguide hologram. *Applied optics*. Optical Society of America. 1995, vol. 34, no. 20, p. 4079–4082.

[14] Chikkerur, S., Cartwright, A. N. and Govindaraju, V. Fingerprint enhancement using stft analysis. *Pattern recognition*. Elsevier. 2007, vol. 40, no. 1, p. 198–211.

[15] Coakley, M. J., Monaco, J. V. and Tappert, C. C. Keystroke biometric studies with short numeric input on smartphones. In: IEEE. *2016 IEEE 8th international conference on biometrics theory, applications and systems (BTAS)*. 2016, p. 1–6.

[16] Cortes, C. and Vapnik, V. Support-vector networks. *Machine learning*. Springer. 1995, vol. 20, no. 3, p. 273–297.

[17] Daugman, J. How iris recognition works. In: *The essential guide to image processing*. Elsevier, 2009, p. 715–739.

[18] De Roose, F., Mainguet, J.-F., Gallaire, D., Martinent, A., Torchia, A. et al. 14-1: Flexible Large-Area Multi-Fingerprint Sensors based on Thermal Mass Detection. In: Wiley Online Library. *SID Symposium Digest of Technical Papers*. 2020, vol. 51, no. 1, p. 176–179.

[19] Dejmal, D. *Analýza systémů pro měření kvality otisku prstů*. Brno, CZ, 2019. Bachelor's thesis. Brno University of Technology, Faculty of Information Technology. Available at: https://www.fit.vut.cz/study/thesis/19344/.

[20] Doddington, G., Liggett, W., Martin, A., Przybocki, M. and Reynolds, D. *Sheep, goats, lambs and wolves: A statistical analysis of speaker performance in the NIST 1998 speaker recognition evaluation*. National Inst of Standards and Technology Gaithersburg Md, 1998.

[21] Doublet, J., Revenu, M. and Lepetit, O. Robust grayscale distribution estimation for contactless palmprint recognition. In: IEEE. *2007 First IEEE International Conference on Biometrics: Theory, Applications, and Systems*. 2007, p. 1–6.

[22] Drahansky, M., Kanich, O. and Dvorak, M. Spoofing methods in hand-based biometrics. In: *Hand-Based Biometrics: Methods and Technology*. Institution of Engineering and Technology, March 2018, p. 367–398. DOI: 10.1049/pbse008e_ch15. Available at: https://doi.org/10.1049/pbse008e_ch15.

[23] Drahanský, M., Kanich, O. and Dvořák, M. Spoofing methods in hand-based biometrics. In: *Hand-Based Biometrics: Methods and Technology*. The Institution of Engineering and Technology, 2018, p. 367–398. IET Book Series on Advances in Biometrics. ISBN 978-1-78561-224-4. Available at: https://www.fit.vut.cz/research/publication/11694.

[24] Drahansky, M., Orság, F., Dvořák, R., Hájek, J., Váňa, J. et al. *Biometrie* [print]. Computer Press, s.r.o, january 2011. 0–0 p.

[25] Duane, C. B. Close-range camera calibration. *Photogramm. Eng.* 1971, vol. 37, no. 8, p. 855–866.

[26] Dubuisson, M.-P. and Jain, A. K. A modified Hausdorff distance for object matching. In: IEEE. *Proceedings of 12th international conference on pattern recognition*. 1994, vol. 1, p. 566–568.

[27] Duta, N., Jain, A. K. and Mardia, K. V. Matching of palmprints. *Pattern Recognition Letters*. Elsevier. 2002, vol. 23, no. 4, p. 477–485.

[28] Dutta, A. and Zisserman, A. The VIA annotation software for images, audio and video. In: *Proceedings of the 27th ACM international conference on multimedia*. 2019, p. 2276–2279.

[29] Dvorak, M. and Drahansky, M. Hand shape recognition and palm-print recognition using 2D and 3D features. In: *Hand-Based Biometrics: Methods and Technology*. Institution of Engineering and Technology, March 2018, p. 283–307. DOI: 10.1049/pbse008e_ch12. Available at: https://doi.org/10.1049/pbse008e_ch12.

[30] Dvořák, M. and Drahanskỳ, M. 3D Hand Geometry Recognition. In: *Biometric-Based Physical and Cybersecurity Systems*. Springer, 2019, p. 103–123.

[31] Dvořák, M., Drahanskỳ, M. and Abdulla, W. H. On the fly biometric identification system using hand-geometry. *IET Biometrics*. Wiley Online Library. 2021, vol. 10, no. 3, p. 315–325.

[32] Dvorak, M., Goldmann, T., Drahansky, M., Dvorak, R., Goldmann, T. et al. *Zařízení pro snímání biometrických charakteristik lidské ruky*. 2020.

[33] Dvorak, M., Goldmann, T., Drahansky, M., Dvorak, R., Goldmann, T. et al. *Einrichtung zur biometrischen Identifikation mit Hilfe von Fingerabdrücken und/oder Handcharakteristiken und Verfahren zur biometrischen Identifikation mit Hilfe dieser Charakteristiken*. 2021.

[34] Dvorák, M., Kanich, O. and Drahanskỳ, M. Scalable Imaging Device using Line Scan Camera for Use in Biometric Recognition and Medical Imaging. In: *BIODEVICES*. 2021, p. 160–168.

[35] Dvorak Michal, D. M. and Waleed, A. B. On-the-fly fingerprint acquisition method. In: *2019 4th International Conference on Mechatronics and Electrical Systems, Praha, CZ*. Asia Society of Researchers, 2019, p. 1–6.

[36] Ferrer, M. A., Fàbregas, J., Faundez, M., Alonso, J. B. and Travieso, C. Hand geometry identification system performance. In: IEEE. *43rd Annual 2009 International Carnahan Conference on Security Technology*. 2009, p. 167–171.

[37] Fusiello, A., Trucco, E. and Verri, A. A compact algorithm for rectification of stereo pairs. *Machine vision and applications*. Springer. 2000, vol. 12, no. 1, p. 16–22.

[38] Gennery, D. B. Stereo-camera calibration. In: *Proceedings ARPA IUS Workshop*. 1979, p. 101–107.

[39] Girault, M., Poupard, G. and Stern, J. On the fly authentication and signature schemes based on groups of unknown order. *Journal of Cryptology*. Springer. 2006, vol. 19, no. 4, p. 463–487.

[40] Goldmann, T., Dvořák, M. and Drahanskỳ, M. *Zařízení pro snímání zbraní*. 2017.

[41] Gozzini, G. *Capacitive pixel for fingerprint sensor*. Google Patents, october 21 2003. US Patent 6,636,053.

[42] Grother, P. and Ngan, M. Face recognition vendor test (FRVT) performance of face identification algorithms, NIST interagency report 8009. *National Institute of Standards and Technology*. 2013, p. 1–138.

[43] Han, C.-C. A hand-based personal authentication using a coarse-to-fine strategy. *Image and Vision Computing*. Elsevier. 2004, vol. 22, no. 11, p. 909–918.

[44] Heidari, M., Goldmann, T., Dvorak, M. and Drahansky, M. Antispoofing and multispectral (optical) methods in hand-based biometrics. In: *Hand-Based Biometrics: Methods and Technology*. Institution of Engineering and Technology, March 2018, p. 337–365. DOI: 10.1049/pbse008e_ch14. Available at: https://doi.org/10.1049/pbse008e_ch14.

[45] Hong, L., Wan, Y. and Jain, A. Fingerprint image enhancement: algorithm and performance evaluation. *IEEE transactions on pattern analysis and machine intelligence*. IEEE. 1998, vol. 20, no. 8, p. 777–789.

[46] Huang, G. B., Mattar, M., Berg, T. and Learned Miller, E. Labeled faces in the wild: A database forstudying face recognition in unconstrained environments. In: *Workshop on faces in'Real-Life'Images: detection, alignment, and recognition*. 2008.

[47] Jain, A., Chen, Y. and Demirkus, M. Pores and ridges: Fingerprint matching using level 3 features. In: IEEE. *18th International Conference on Pattern Recognition (ICPR'06)*. 2006, vol. 4, p. 477–480.

[48] Jiang, X., Lu, Y., Tang, H.-Y., Tsai, J. M., Ng, E. J. et al. Monolithic ultrasound fingerprint sensor. *Microsystems & nanoengineering*. Nature Publishing Group. 2017, vol. 3, no. 1, p. 1–8.

[49] Joo, S.-I., Weon, S.-H. and Choi, H.-I. Real-time depth-based hand detection and tracking. *The Scientific World Journal*. Hindawi. 2014, vol. 2014.

[50] Kamgar Parsi, B., Lawson, W. and Kamgar Parsi, B. Toward development of a face recognition system for watchlist surveillance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. IEEE. 2011, vol. 33, no. 10, p. 1925–1937.

[51] Kanade, T. and Okutomi, M. A stereo matching algorithm with an adaptive window: Theory and experiment. *IEEE transactions on pattern analysis and machine intelligence*. IEEE. 1994, vol. 16, no. 9, p. 920–932.

[52] Kanhangad, V., Kumar, A. and Zhang, D. Combining 2D and 3D hand geometry features for biometric verification. In: IEEE. *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. 2009, p. 39–44.

[53] Kanhangad, V., Kumar, A. and Zhang, D. A unified framework for contactless hand verification. *IEEE transactions on information forensics and security*. IEEE. 2011, vol. 6, no. 3, p. 1014–1027.

[54] Kim, S., Ban, Y. and Lee, S. Tracking and classification of in-air hand gesture based on thermal guided joint filter. *Sensors*. Multidisciplinary Digital Publishing Institute. 2017, vol. 17, no. 1, p. 166.

[55] Kohavi, R. et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Montreal, Canada. *Ijcai*. 1995, vol. 14, no. 2, p. 1137–1145.

[56] Kong, A., Zhang, D. and Kamel, M. A survey of palmprint recognition. *Pattern recognition*. Elsevier. 2009, vol. 42, no. 7, p. 1408–1418.

[57] Kuzu, R. S., Piciucco, E., Maiorana, E. and Campisi, P. On-the-fly finger-vein-based biometric recognition using deep neural networks. *IEEE Transactions on information Forensics and Security*. IEEE. 2020, vol. 15, p. 2641–2654.

[58] Labati, R. D., Genovese, A., Munoz, E., Piuri, V., Scotti, F. et al. Advanced design of automated border control gates: biometric system techniques and research trends. In: IEEE. *2015 IEEE International Symposium on Systems Engineering (ISSE)*. 2015, p. 412–419.

[59] Lai, K., Kanich, O., Dvořák, M., Drahanskỳ, M., Yanushkevich, S. et al. Biometric-enabled watchlists technology. *IET Biometrics*. IET. 2017, vol. 7, no. 2, p. 163–172.

[60] Le Noc, L., Picard, F., Jerominek, H., Alain, C., Pope, T. D. et al. Fingerprint recording with bolometric detectors. In: SPIE. *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Defense and Law Enforcement II*. 2003, vol. 5071, p. 442–452.

[61] Lehtola, V. V., Kurkela, M. and Rönnholm, P. Radial Distortion from Epipolar Constraint for Rectilinear Cameras. *Journal of Imaging*. 2017, vol. 3, no. 1. DOI: 10.3390/jimaging3010008. ISSN 2313-433X. Available at: https://www.mdpi.com/2313-433X/3/1/8.

[62] Li, J., Feng, J. and Kuo, C.-C. J. Deep convolutional neural network for latent fingerprint enhancement. *Signal Processing: Image Communication*. Elsevier. 2018, vol. 60, p. 52–63.

[63] Lim, M.-H. and Yuen, P. C. Entropy Measurement for Biometric Verification Systems. *IEEE Transactions on Cybernetics*. 2016, vol. 46, no. 5, p. 1065–1077. DOI: 10.1109/TCYB.2015.2423271.

[64] Liu, X., Pedersen, M. and Charrier, C. Image-based attributes of multi-modality image quality for contactless biometric samples. In: IEEE. *2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN)*. 2016, p. 106–111.

[65] Liu, Z., Luo, P., Wang, X. and Tang, X. Deep learning face attributes in the wild. In: *Proceedings of the IEEE international conference on computer vision*. 2015, p. 3730–3738.

[66] Lopes Silva, P., Luz, E., Moreira, G., Moraes, L. and Menotti, D. Chimerical dataset creation protocol based on Doddington zoo: A biometric application with face, eye, and ECG. *Sensors*. Multidisciplinary Digital Publishing Institute. 2019, vol. 19, no. 13, p. 2968.

[67] Lu, Y., Tang, H., Fung, S., Wang, Q., Tsai, J. et al. Ultrasonic fingerprint sensor using a piezoelectric micromachined ultrasonic transducer array integrated with complementary metal oxide semiconductor electronics. *Applied Physics Letters*. AIP Publishing LLC. 2015, vol. 106, no. 26, p. 263503.

[68] Mahalanobis, P. C. On the Generalised Distance in Statistics. *Proceedings of the National Institute of Sciences of India*. 1936, vol. 2, no. 1, p. 49—55.

[69] Maltoni, D., Maio, D., Jain, A. K. and Prabhakar, S. *Handbook of fingerprint recognition*. Springer Science & Business Media, 2009.

[70] MATLAB. *Version 7.10.0 (R2010a)*. Natick, Massachusetts: The MathWorks Inc., 2010.

[71] Oostveen, A.-M., Kaufmann, M., Krempel, E. and Grasemann, G. Automated border control: a comparative usability study at two European airports. In: *8th International Conference on Interfaces and Human Computer Interaction (IHCI 2014), Lisbon, Portugal*. 2014.

[72] Pal, M. Random forest classifier for remote sensing classification. *International journal of remote sensing*. Taylor & Francis. 2005, vol. 26, no. 1, p. 217–222.

[73] Panteleris, P. and Argyros, A. Back to rgb: 3d tracking of hands and hand-object interactions based on short-baseline stereo. In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2017, p. 575–584.

[74] Parziale, G. and Chen, Y. Advanced technologies for touchless fingerprint recognition. In: *Handbook of Remote Biometrics*. Springer, 2009, p. 83–109.

[75] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B. et al. Scikit-learn: Machine learning in Python. *The Journal of machine Learning research*. JMLR. org. 2011, vol. 12, p. 2825–2830.

[76] Phillips, P. J., Flynn, P. J., Scruggs, T., Bowyer, K. W., Chang, J. et al. Overview of the face recognition grand challenge. In: IEEE. *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*. 2005, vol. 1, p. 947–954.

[77] Pititeeraphab, Y. and Pintavirooj, C. Identity Verification Using Geometry of Human hands. In: IEEE. *2018 11th Biomedical Engineering International Conference (BMEiCON)*. 2018, p. 1–4.

[78] Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A. et al. Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*. Elsevier. 1987, vol. 39, no. 3, p. 355–368.

[79] Poh, N., Kittler, J. and Bourlai, T. Quality-based score normalization with device qualitative information for multimodal biometric fusion. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*. IEEE. 2010, vol. 40, no. 3, p. 539–554.

[80] Putra, I. and Sentosa, M. A. Hand geometry verification based on chain code and dynamic time warping. *International Journal of Computer Applications*. Citeseer. 2012, vol. 38, no. 12, p. 17–22.

[81] Raghavendra, R., Yang, B. and Busch, C. Robust "On-the-Fly" person identification using Sparse Representation. In: IEEE. *2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*. 2013, p. 1–4.

[82] Ray, S. F. *Applied photographic optics: Lenses and optical systems for photography, film, video, electronic and digital imaging*. Routledge, 2002.

[83] Ross, A., Jain, A. and Pankati, S. A prototype hand geometry-based verification system. In: *Proceedings of 2nd conference on audio and video based biometric person authentication*. 1999, p. 166–171.

[84] Roy, S. and Cox, I. J. A maximum-flow formulation of the n-camera stereo correspondence problem. In: IEEE. *Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271)*. 1998, p. 492–499.

[85] Schippers, P., Dobmeier, M. P., Kuehlewein, R., Bustgens, B. and Urban, G. Thermal flexible foil-fingerprint sensor. In: IEEE. *The 13th International Conference on Solid-State Sensors, Actuators and Microsystems, 2005. Digest of Technical Papers. TRANSDUCERS'05*. 2005, vol. 1, p. 579–580.

[86] Sepasian, M., Balachandran, W. and Mares, C. Image enhancement for fingerprint minutiae-based algorithms using CLAHE, standard deviation analysis and sliding neighborhood. In: *Proceedings of the World congress on Engineering and Computer Science*. 2008, p. 22–24.

[87] Shannon, C. Communication in the Presence of Noise. *Proceedings of the IRE*. Institute of Electrical and Electronics Engineers (IEEE). jan 1949, vol. 37, no. 1, p. 10–21. DOI: 10.1109/jrproc.1949.232969. Available at: https://doi.org/10.1109/jrproc.1949.232969.

[88] Shu, W. and Zhang, D. Automated personal identification by palmprint. *Optical Engineering*. 1998, vol. 37, p. 2359–2362.

[89] Singh, A. P., Thakur, R. K., Kumar, A. and Baksh, R. User authentication using hand images. *International Journal of Science and Research*. 2014, vol. 3, no. 3, p. 317–322.

[90] Siswanto, A., Tarigan, P. and Fahmi, F. Design of contactless hand biometric system with relative geometric parameters. In: IEEE. *2013 3rd International Conference on Instrumentation, Communications, Information Technology and Biomedical Engineering (ICICI-BME)*. 2013, p. 199–203.

[91] Smith, N. D. and Sharp, J. S. Accessible biometrics: A frustrated total internal reflection approach to imaging fingerprints. *Science & Justice*. Elsevier. 2017, vol. 57, no. 3, p. 193–198.

[92] Stone, M. Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society: Series B (Methodological)*. Wiley Online Library. 1974, vol. 36, no. 2, p. 111–133.

[93] Svoboda, J., Bronstein, M. M. and Drahansky, M. Contactless biometric hand geometry recognition using a low-cost 3D camera. In: IEEE. *2015 International Conference on Biometrics (ICB)*. 2015, p. 452–457.

[94] Svoboda, J., Klubal, O. and Drahanskỳ, M. Biometric recognition of people by 3D hand geometry. In: IEEE. *The International Conference on Digital Technologies 2013*. 2013, p. 137–141.

[95] Tang, H., Lu, Y., Fung, S., Tsai, J., Daneman, M. et al. Pulse-echo ultrasonic fingerprint sensor on a chip. In: IEEE. *2015 Transducers-2015 18th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS)*. 2015, p. 674–677.

[96] Tang, Y., Jiang, L., Hou, Y. and Wang, R. Contactless fingerprint image enhancement algorithm based on Hessian matrix and STFT. In: IEEE. *2017 2nd International Conference on Multimedia and Image Processing (ICMIP)*. 2017, p. 156–160.

[97] Uchida, K. Fingerprint-based user-friendly interface and pocket-PID for mobile authentication. In: IEEE. *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*. 2000, vol. 4, p. 205–209.

[98] Varchol, P. and Levicky, D. Using of hand geometry in biometric security systems. *Radioengineering-Prague-*. Citeseer. 2007, vol. 16, no. 4, p. 82.

[99] Villegas, O. O. V., Domínguez, H. d. J. O., Sanchez, V., Maynez, L. O. and Orozco, H. M. Biometric human identification of hand geometry features using discrete wavelet transform. *Discrete Wavelet Transforms–Biomedical Applications*. 2006, p. 251–266.

[100] Wang, M.-H. Hand recognition using thermal image and extension neural network. *Mathematical Problems in Engineering*. Hindawi. 2012, vol. 2012.

[101] Watson, C., Fiumara, G., Tabassi, E. et al. Fingerprint vendor technology evaluation, NIST interagency report 8034. *National Institute of Standards and Technology*. 2014, p. 1–233.

[102] Watson, C. I. and Wilson, C. L. NIST special database 4. *Fingerprint Database, National Institute of Standards and Technology*. Citeseer. 1992, vol. 17, no. 77, p. 5.

[103] Weinberger, K. Q. and Saul, L. K. Distance metric learning for large margin nearest neighbor classification. *Journal of machine learning research.* 2009, vol. 10, no. 2.

[104] Weissenfeld, A., Zoufal, A., Weiss, C., Strobl, B. and Domínguez, G. F. Towards mobile contactless 4-fingerprint authentication for border control. In: IEEE. *2018 European Intelligence and Security Informatics Conference (EISIC).* 2018, p. 73–76.

[105] Wirayuda, T. A. B., Kuswanto, D. H., Adhi, H. A. and Dayawati, R. N. Implementation of feature extraction based hand geometry in biometric identification system. In: IEEE. *2013 International Conference of Information and Communication Technology (ICoICT).* 2013, p. 259–263.

[106] Wittman, M., Davis, P. and Flynn, P. J. Empirical studies of the existence of the biometric menagerie in the frgc 2.0 color image corpus. In: IEEE. *2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06).* 2006, p. 33–33.

[107] Wong, M., Zhang, D., Kong, W.-K. and Lu, G. Real-time palmprint acquisition system design. *IEE Proceedings-Vision, Image and Signal Processing.* IET. 2005, vol. 152, no. 5, p. 527–534.

[108] Woodard, D. L. and Flynn, P. J. Finger surface as a biometric identifier. *Computer vision and image understanding.* Elsevier. 2005, vol. 100, no. 3, p. 357–384.

[109] Yager, N. and Dunstone, T. The biometric menagerie. *IEEE transactions on pattern analysis and machine intelligence.* IEEE. 2008, vol. 32, no. 2, p. 220–230.

[110] Yoruk, E., Konukoglu, E., Sankur, B. and Darbon, J. Shape-based hand recognition. *IEEE transactions on image processing.* IEEE. 2006, vol. 15, no. 7, p. 1803–1815.

[111] Zhang, D., Kong, W.-K., You, J. and Wong, M. Online palmprint identification. *IEEE Transactions on pattern analysis and machine intelligence.* IEEE. 2003, vol. 25, no. 9, p. 1041–1050.

[112] Zhang, L., Zhang, L., Zhang, D. and Zhu, H. Online finger-knuckle-print verification for personal authentication. *Pattern recognition.* Elsevier. 2010, vol. 43, no. 7, p. 2560–2571.

[113] Zhang, L., Zhang, L., Zhang, D. and Zhu, H. Ensemble of local and global information for finger–knuckle-print recognition. *Pattern recognition.* Elsevier. 2011, vol. 44, no. 9, p. 1990–1998.

[114] Zhang, T., Zeng, X., Zhang, Y., Sun, K., Wang, Y. et al. Thermalring: Gesture and tag inputs enabled by a thermal imaging smart ring. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* 2020, p. 1–13.

[115] Zhang, W. 05-Camera lens and distortion. *Camera lens and distortion.* Chuanzhi College. 2000. Available at: http://robot.czxy.com/docs/camera/chapter01/05-distortions/.

[116] Zhang, Z. A Flexible New Technique for Camera Calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. December 2000, vol. 22, p. 1330–1334. MSR-TR-98-71, Updated March 25, 1999. Available at: https://www.microsoft.com/en-us/research/publication/a-flexible-new-technique-for-camera-calibration/.

[117] Zuiderveld, K. Contrast limited adaptive histogram equalization. *Graphics gems*. Academic Press. 1994, p. 474–485.

Appendix A

On the fly biometry enabled device

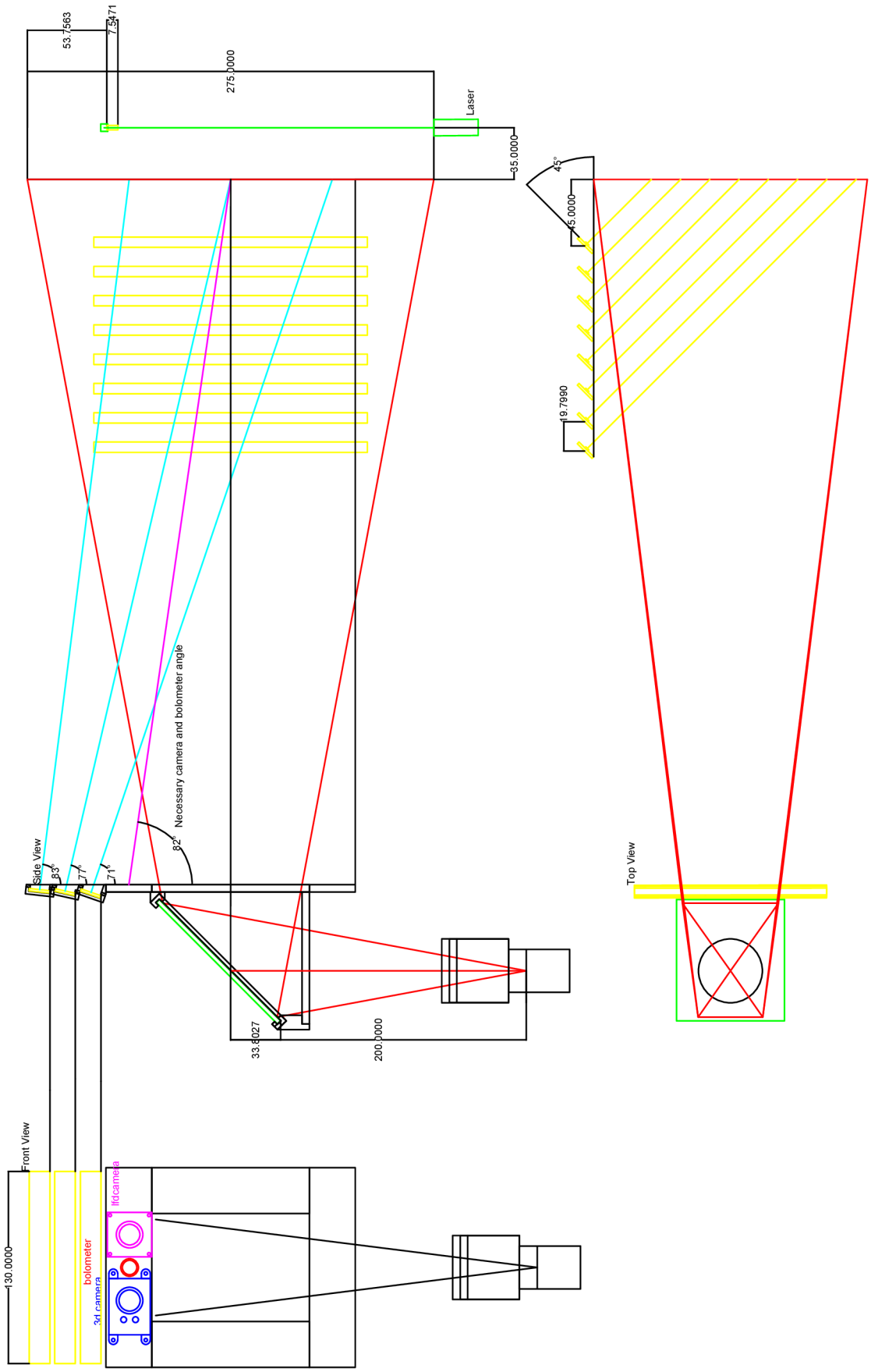

Figure A.1: The placement of optical elements of the On the fly commercial prototype.

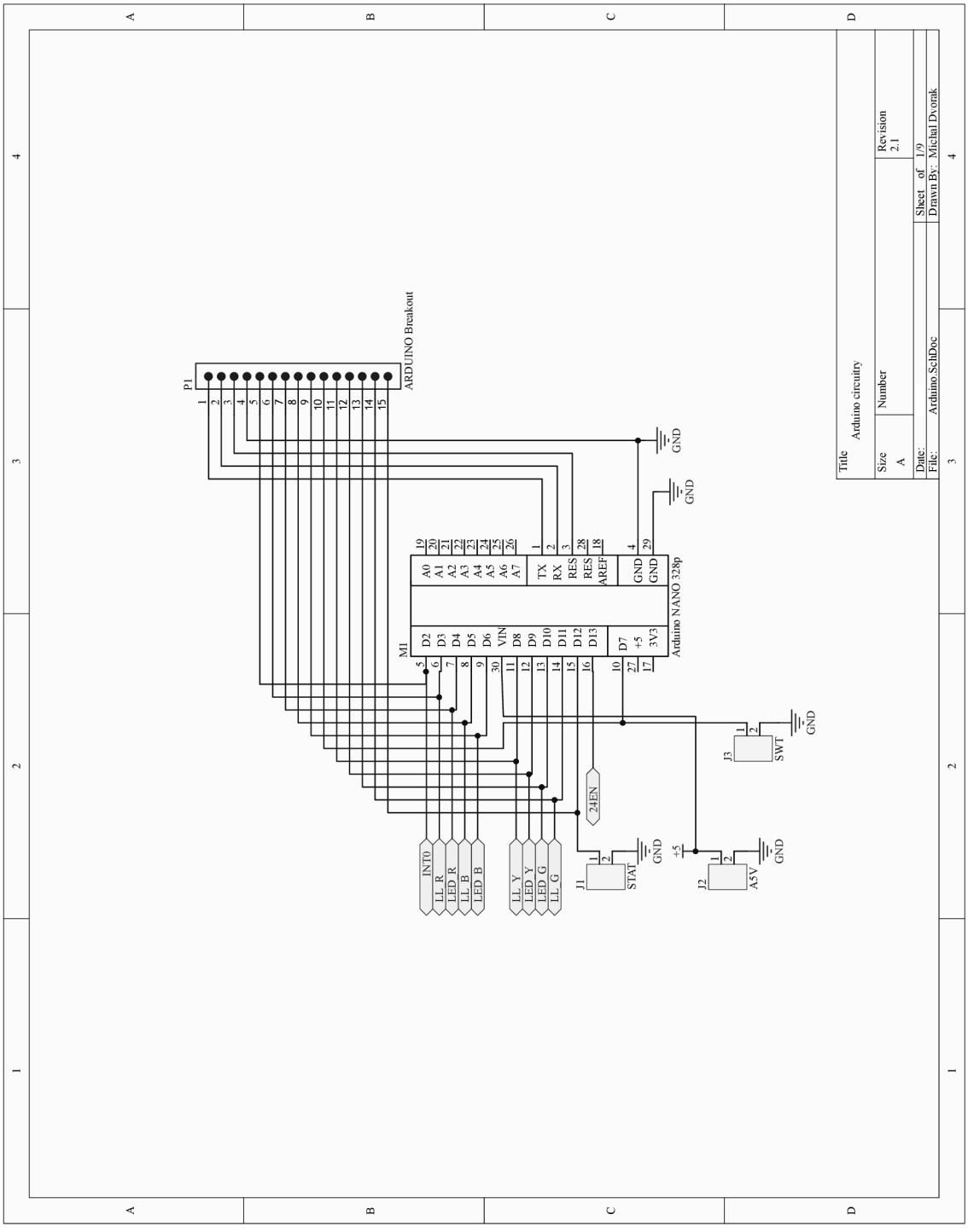

Figure A.2: MCU controller circuit of the control board of the On the fly commercial prototype.

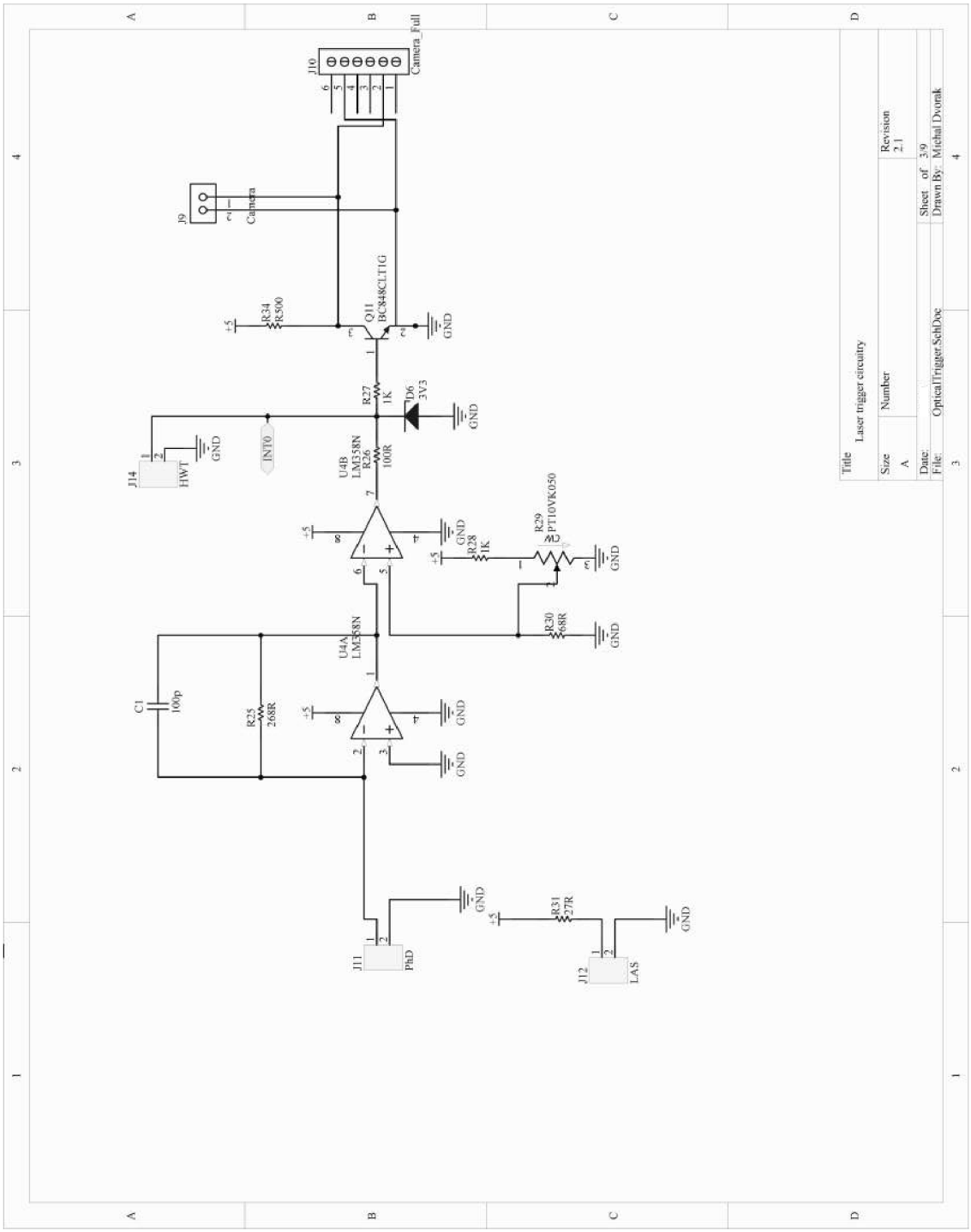

Figure A.3: Optical trigger circuit for the OTF Hand device.

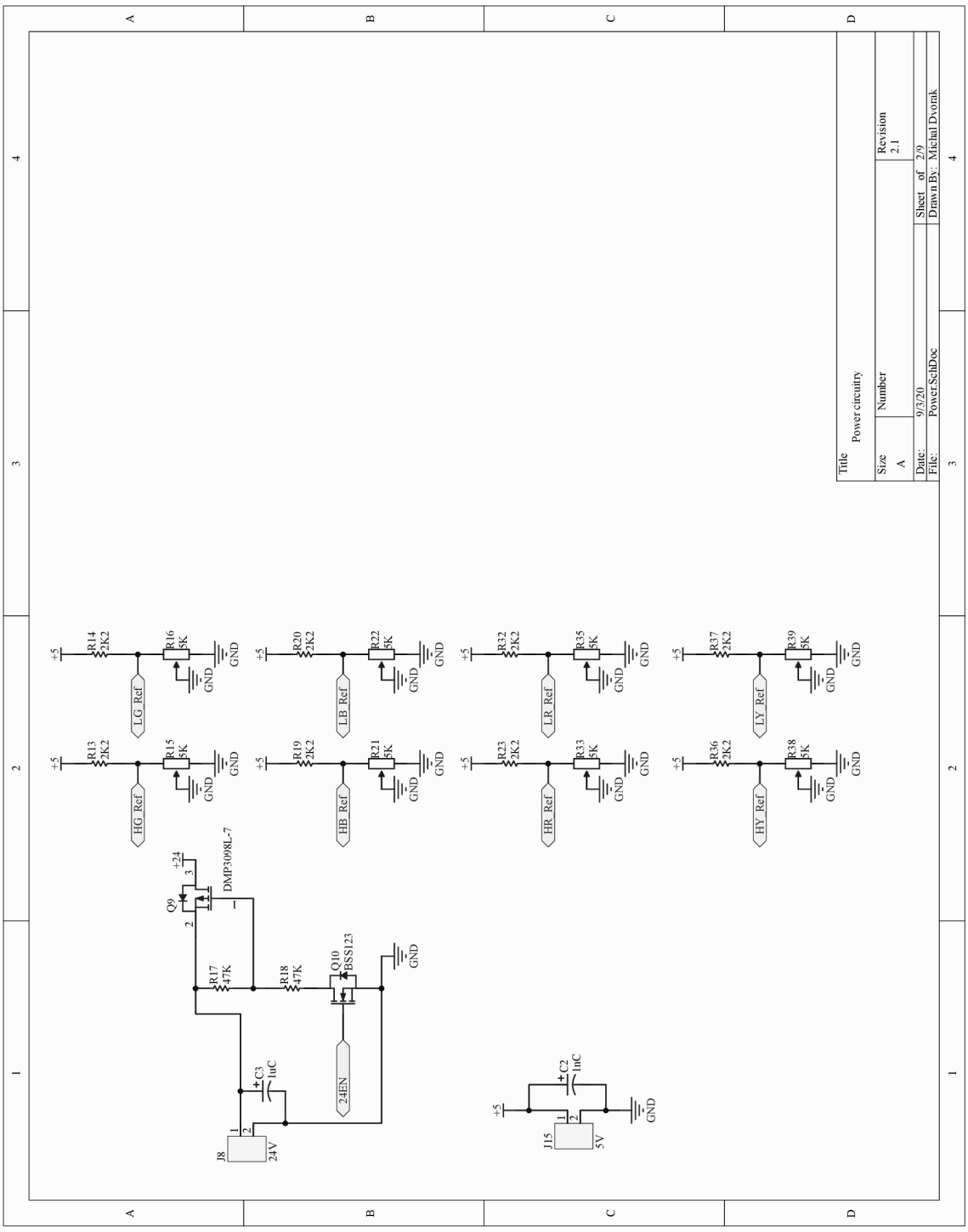

Figure A.4: Power circuit and the reference set circuit for individual color panels for the OTF Hand device.

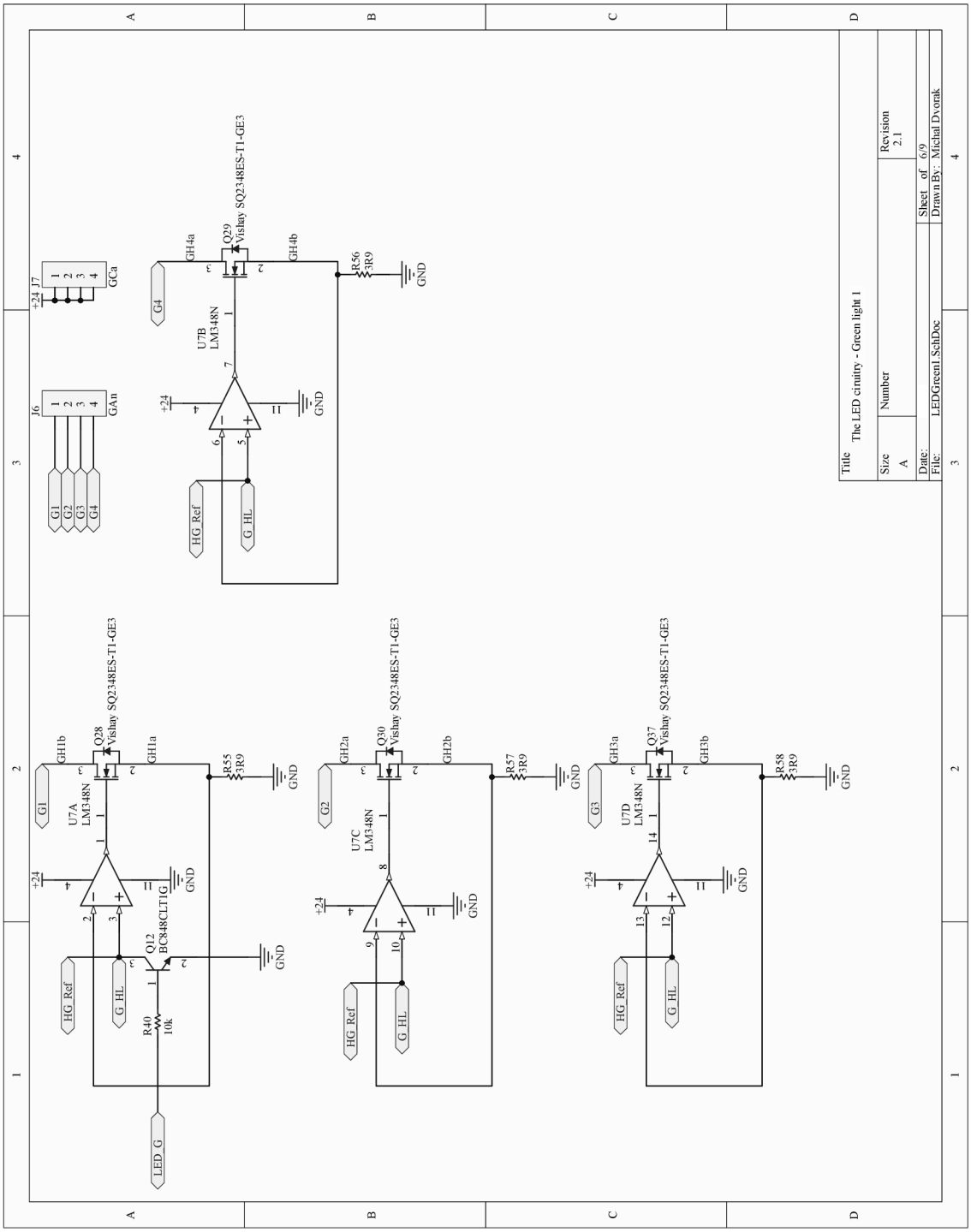

Figure A.5: Current follower circuit for the high level green panel.

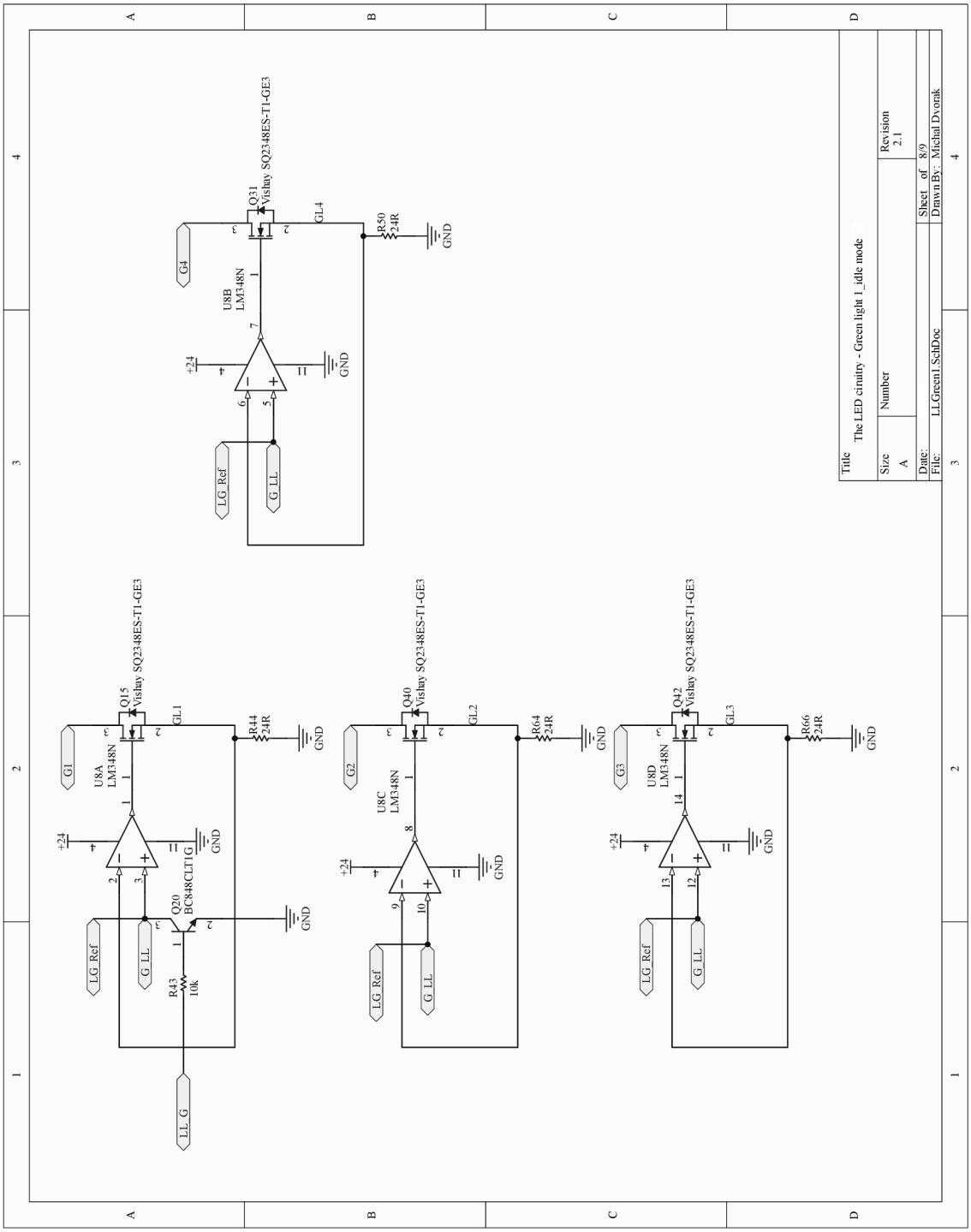

Figure A.6: Current follower circuit for the low level green panel.

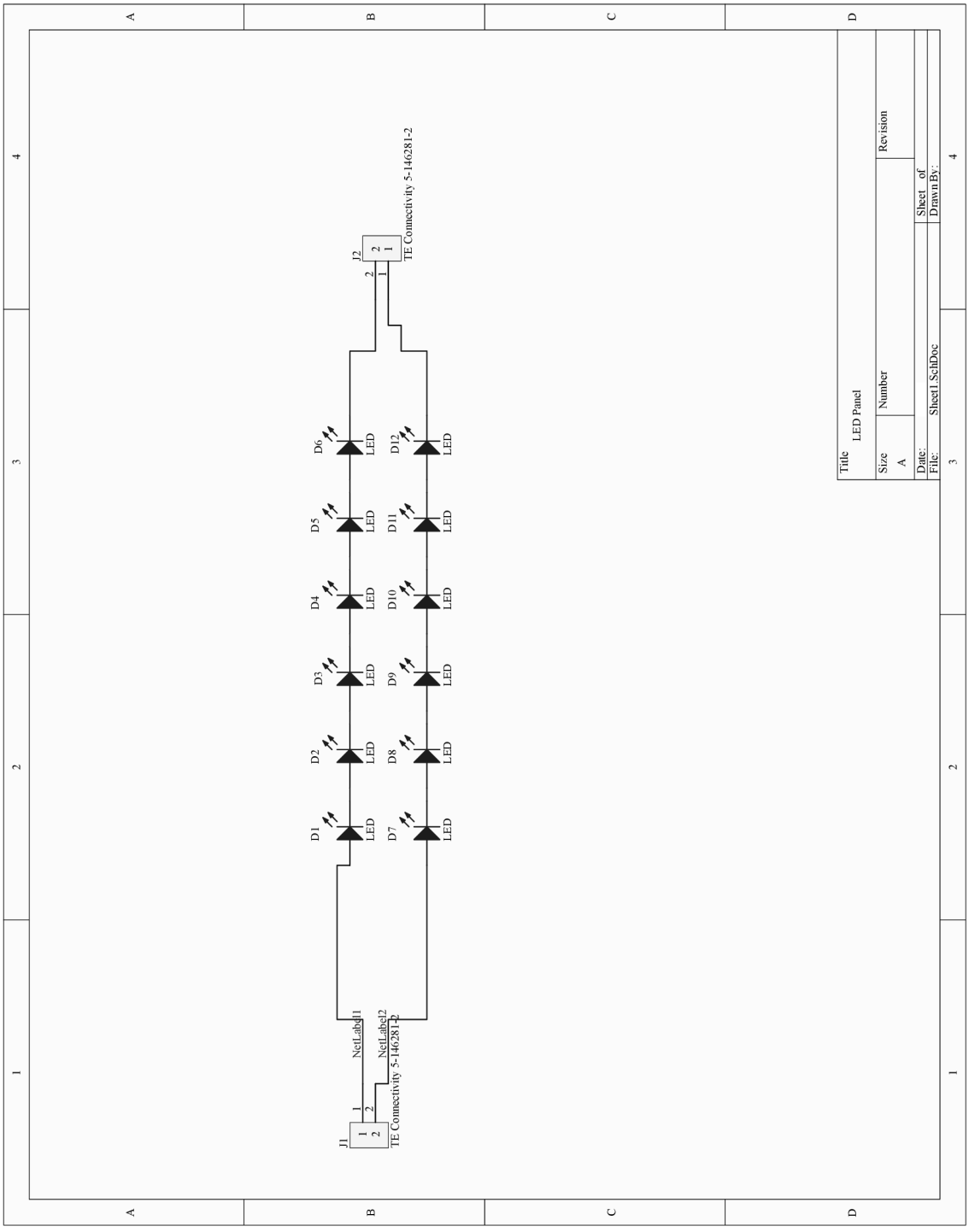

Figure A.7: Circuit of the green LED panel.

Table A.1: Feature selection for the hand geometry recognition classifier

fn	**pn**	**pm**	**fn**	**pn**	**pm**	**fn**	**pn**	**pm**
0	1	2	45	1	35	89	16	29
1	2	3	46	1	36	90	16	30
2	3	4	47	1	37	91	17	18
3	4	5	48	2	4	92	17	19
4	5	6	49	2	5	93	17	20
5	6	7	50	2	6	94	17	21
6	7	8	51	2	7	95	17	22
7	8	9	52	2	8	96	18	20
8	9	10	53	2	29	97	18	21
9	10	11	54	2	30	98	18	22
10	11	12	55	3	4	99	19	21
11	12	13	56	3	5	100	19	22
12	13	14	57	3	6	101	20	22
13	14	15	58	3	7	102	23	25
14	15	16	59	3	8	103	23	26
15	16	17	60	4	6	104	23	27
16	17	18	61	4	7	105	23	28
18	18	19	62	4	8	106	23	29
19	19	20	63	5	7	107	23	29
20	20	21	64	5	8	108	23	30
21	21	22	65	6	8	109	24	25
22	22	23	66	9	11	110	24	26
23	23	24	67	9	12	111	24	27
24	24	25	68	9	13	112	24	28
25	25	26	69	9	14	113	24	29
26	26	27	70	9	15	114	25	27
27	27	28	71	9	29	115	25	28
28	28	29	72	9	30	116	25	29
29	29	30	73	10	11	117	26	28
30	30	31	74	10	12	118	26	29
31	31	32	75	10	13	119	27	29
32	32	33	76	10	14	120	31	33
33	33	34	77	10	15	121	31	34
34	34	35	78	11	13	122	31	35
35	35	36	79	11	14	123	31	36
36	36	37	80	11	15	124	31	37
37	1	3	81	12	14	125	32	34
38	1	4	82	12	15	126	32	35
39	1	5	83	13	15	127	32	36
40	1	8	84	16	18	128	33	35
41	1	15	85	16	19	129	33	36
42	1	22	86	16	20	130	34	36
43	1	29	87	16	21			
44	1	30	88	16	22			

Appendix B

Biometry enabled scanner

Figure B.1: Database (cropped) from the linear scanner page 1. Left hand, dorsal side.

Figure B.2: Database (cropped) from the linear scanner page 1. Palmar side.