VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

DIPLOMOVÁ PRÁCE

Brno, 2022 Bc. Dominik Řičánek

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV AUTOMATIZACE A MĚŘICÍ TECHNIKY

DEPARTMENT OF CONTROL AND INSTRUMENTATION

EULEROVSKÁ VIDEO-MAGNIFIKACE

EULERIAN VIDEO MAGNIFICATION

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE — Bc. Dominik Řičánek

AUTHOR

VEDOUCÍ PRÁCE — Ing. Peter Honec, Ph.D.

SUPERVISOR

BRNO 2022

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

Diplomová práce

magisterský navazující studijní program **Kybernetika, automatizace a měření**

Ústav automatizace a měřicí techniky

Student: Bc. Dominik Řičánek ***ID:*** 203334

Ročník: 2 ***Akademický rok:*** 2021/22

NÁZEV TÉMATU:

Eulerovská video-magnifikace

POKYNY PRO VYPRACOVÁNÍ:

Vytvořte systém pro vyhodnocení životních funkcí s využitím eulerovské video-magnifikace (Eulerian Video Magnification - EVM) nebo jiného algoritmu.

1. Proveďte rešerši EMV a alternativních metod pro detekci malých prostorových, jasových nebo barevných změn obrazu v čase.
2. Vyberte vhodný use-case (detekce dechu, pulsu, spánkové fáze, emocí apod.).
3. Na základě okrajových podmínek (nutná bitová hloubka, rozlišení, framerate…) navrhněte vhodný hardware pro snímání.
4. Sestrojte přípravek a nasnímejte dataset.
5. Nasnímaná data odfiltrujte (low pass, high pass) a pomocí EMV extrahujte relevantní informace
6. Otestujte a zhodnoťte.

DOPORUČENÁ LITERATURA:

HLAVAC V., SONKA M., BOYLE R.: Image Processing, Analysis, and Machine Vision, ISBN 978-0495082521

Termín zadání: 7.2.2022 ***Termín odevzdání:*** 18.5.2022

Vedoucí práce: Ing. Peter Honec, Ph.D.

doc. Ing. Petr Fiedler, Ph.D.
předseda rady studijního programu

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

V této práci si přiblížíme problematiku eulerovské video-magnifikace (EVM), její počátky a inovace, které vedly až k možnosti jejího využití pro detekci srdečního rytmu v reálném čase. Pomocí Gaussovy pyramidy jsou zvýrazněny rozdíly v prostorových změnách výřezů ze snímaného videa, získaných pomocí MediaPipe Face Mesh algoritmu, využívajícího strojové učení. Výsledkem je dostatečně výrazný časový signál, ze kterého se následně Fourierovou transformací získává tepová frekvence detekovaného subjektu. Je zde navržena metoda měření pro vyhodnocení použitelnosti EVM ve sledovacích systémech. Na konec jsou probrány nedostatky takového systému a možné způsoby jejich nápravy.

KLÍČOVÁ SLOVA

Gaussova pyramida, lineární eulerovská video-magnifikace, bezkontaktní měření pulsu, detekce pózy obličeje

ABSTRACT

In this paper we go over the research behind eulerian video magnification, its beginnings and inovations which lead to its potential use in real time heart rate measurement. The spacial differences are exaggerated using the Gaussian pyramid in conjuncture with face detection using the MediaPipe Face Mesh solution, which leverages machine learning for ultra fast performance. The resulting signal is clear enough to allow pulse extraction using the Fourier transform. Evaluation is conducted upon the final prototype to see how suitable is EVM in combination MediaPipe for use in surrveilance systems. At the end we discuss the limitations and how to possibly overcome them.

KEYWORDS

Gaussian pyramid, linear eulerian video magnification, non-contact pulse measurement, face pose detection

ŘIČÁNEK, Dominik. *Eulerovská video-magnifikace*. Brno, 2022, 77 s. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav automatizace a měřicí techniky. Vedoucí práce: Ing. Peter Honec, Ph.D.

Vysázeno pomocí balíčku `thesis` verze 4.00; `http://latex.feec.vutbr.cz`

PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma „Eulerovská video-magnifikace“ jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno ....

podpis autora

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu diplomové práce panu Ing. Peteru Honcovi, Ph.D. a také panu Ing. Šimonu Bilíkovi za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

Speciálně pak děkuji své rodině, která mi byla oporou po celou dobu mého dosavadního studia.

Obsah

Úvod **15**

1 Měření životních funkcí **17**
1.1 Srdeční rytmus . . . 17
1.1.1 Měření srdečního rytmu . . . 17
1.2 Dechová frekvence . . . 20
1.2.1 Měření dechové frekvence . . . 21
1.3 Shrnutí . . . 22

2 Zvýraznění nepatrných změn pomocí EVM **25**
2.1 Lineární EVM . . . 26
2.1.1 Zvýraznění pohybu [1] . . . 26
2.1.2 Výběr faktoru α [1] . . . 27
2.2 Problém nežádoucích pohybů . . . 28
2.3 Fázová EVM . . . 28
2.3.1 Zvýraznění pohybu pomocí fáze [2] . . . 29
2.3.2 Výběr faktoru α [2] . . . 30
2.3.3 Fázová EVM pomocí Rieszovy pyramidy [3] . . . 32
2.4 Ukázka využití . . . 33
2.5 Shrnutí . . . 35

3 Detekce subjektu **37**
3.1 Viola-Jones detektor obličeje . . . 37
3.1.1 Haarovy příznaky . . . 37
3.1.2 Integrální obraz . . . 38
3.1.3 AdaBoost . . . 39
3.1.4 Kaskádové zapojení klasifikátorů . . . 40
3.1.5 Ukázka využití . . . 40
3.2 MediaPipe Face Mesh . . . 41
3.2.1 Prokústésova analýza . . . 41
3.2.2 Modely MediaPipe . . . 42
3.2.3 Detekce obličeje . . . 43
3.2.4 Ukázka využití . . . 43
3.3 Shrnutí . . . 44

4 Řešení **45**
4.1 Výpočetní jednotka . . . 45
4.2 Kamera . . . 46

4.3 Uživatelské rozhraní 47
4.4 Zdroj 47
4.5 Snímač tepové frekvence 48
4.6 Software 49
4.6.1 Předspracování v Matlabu 49
4.6.2 Předzpracování v Pythonu 50
4.6.3 Implementace EVM v Pythonu 51
4.7 Měření 51
4.7.1 Podmínky scény 52
4.7.2 Sběr dat 52
4.8 Shrnutí 54

5 Výsledky **57**
5.1 Volba parametrů algoritmu 57
5.2 Výsledky měření pomocí Viola-Jones 57
5.3 Výsledky měření pomocí MediaPipe Face Mesh 58
5.3.1 Vzdálenost 59
5.3.2 Úhel natočení tváře 60
5.3.3 Barevný tón kůže 61
5.3.4 Osvětlení 62
5.3.5 Pohyb 63
5.3.6 Klid 64
5.4 Shrnutí 65

Závěr **69**

Literatura **71**

Seznam příloh **75**

6 Obsah elektronické přílohy **77**

Seznam obrázků

1.1 EKG svody . . . 18
1.2 Umístění hrudních elektrod . . . 19
1.3 Signál EKG [28] . . . 19
1.4 Beurer Pulzní oxymetr PO 80 . . . 20
1.5 Dechové křivky . . . 22
2.1 Eulerovská video-magnifikace pro zvýraznění srdečního pulsu v obličeji 25
2.2 Princip fázové EVM . . . 30
2.3 EVM pomocí Rieszových pyramid . . . 33
2.4 Příklad zvýraznění pulsu na tváři . . . 34
2.5 Monitoring dýchání dítěte . . . 34
3.1 Haarovy příznaky . . . 38
3.2 Integrální obraz . . . 39
3.3 AdaBoost . . . 40
3.4 Obličej detekovaný pomocí Viola-Jones . . . 41
3.5 Průběh generického ICP algoritmu . . . 42
3.6 Kabshův algoritmus . . . 42
3.7 Obličej detekovaný pomocí masky ze dvou různých úhlů. . . . 44
4.1 Výpočetní jednotka . . . 46
4.2 Kamera . . . 47
4.3 Kamera . . . 48
4.4 Vyříznutá oblast . . . 50
4.5 Detekce obličeje . . . 51
4.6 Výřez měřené oblasti . . . 52
4.7 EVM flowchart . . . 53
4.8 Hlavní funkce . . . 54
4.9 Záznam tepu . . . 55
5.1 Časový průběh srdečního rytmu změřený EVM . . . 66
5.2 Časový průběh zeleného kanálu . . . 67

Seznam tabulek

1.1 Tab 1.1: Fyziologické hodnoty tepové frekvence u různých věkových skupin [22] . . . 17
1.2 Fyziologické hodnoty dechové frekvence u různých věkových skupin. [22] . . . 21
5.1 Výsledky raného řešení . . . 58
5.2 30cm od kamery . . . 59
5.3 1m od kamery . . . 59
5.4 2m od kamery . . . 60
5.5 Horizontální úhel natočení tváře . . . 60
5.6 Vertikální úhel natočení tváře . . . 61
5.7 Obecný úhel natočení tváře . . . 61
5.8 Bílý tón kůže . . . 61
5.9 Světle hnědý tón kůže . . . 62
5.10 Tmavě hnědý tón kůže . . . 62
5.11 LED osvětlení . . . 62
5.12 Sluneční světlo . . . 63
5.13 Pomalý pohyb . . . 63
5.14 Rychlý pohyb . . . 63
5.15 Velmi rychlý pohyb . . . 64
5.16 Tepová frekvence v klidu . . . 64

Úvod

Mít k dispozici síť milionů mikroskopických senzorů, schopnou pokrýt povrch snímaného objektu zní jako senzace, ačkoliv přesně k takovému zařízení bychom mohli přirovnat dnešní digitální kamery.

Lidské oko není dobře adaptované na rozpoznávání malých prostorově-časových změn. Přitom kolem nás existuje celý neviditelný svět, plný užitečných informací. Například lidská hlava nepatrně mění svůj tvar i barvu v reakci na srdeční rytmus. Zvýrazněním těchto nepatrných změn se zabývá rodina algoritmů nazývaná eulerovská video-magnifikace (eulerian video magnification), se kterou přišli Wu et al. [1] v roce 2012, vycházejíc z dřívějších prací využívajících Lagrangeovu perspektivu projekce trajektorie jednotlivých částic kapaliny, resp. pixelů v obraze (optical flow).

Eulerovská perspektiva závisí na postupné změně vlastností média v čase - jako například rychlosti částic v kapalině. Pouze tedy zvýrazňuje malé změny, kdežto Lagrangeova extrapoluje. Díky tomu dokáže být EVM nejen výpočetně jednodušší, a tudíž použitelná v aplikacích běžících v reálném čase, ale také méně náchylná na artefakty. [1]

Přesto je EVM původně navržená na MIT náchylná na šum, jehož amplitudu, stejně jako amplitudu požadovaného signálu, lineárně zvyšuje. Kromě toho, využití Laplacových pyramid, ač výpočetně méně náročné než předcházející metody, je stále relativně náročný algoritmus. Pokroky v této oblasti, kterých bylo dosaženo od roku 2012, navíc mají svá vlastní úskalí a kompromisy, o kterých si povíme v patřičné kapitole.

Zvýraznění barevných a tvarových změn nachází své využití například v bezkontaktním měření tepové frekvence v experimentální biologii, kdy není praktické aplikovat elektrody nebo by to přímo znemožňovalo průběh experimentu. O to se snažili Lauridsen et al. [12] při současném snímání několika axolotlů mexických (relativně pasivních obojživelníků) nebo při snímání tepové frekvence kuřecích embryí in vivo, přes skořápku. Mimo to byli schopni změřit rychlost pulzní vlny nebo také infrazvukovou komunikaci slonů.

Další využití lze nalézt například ve zlepšení rozpoznávání výrazů v obličeji, jak zkoumají Ngo et al. [8], kde v případě nedostatečně výrazných grimas mají moderní algoritmy nízkou úspěšnost. EVM je zde využito právě pro zvýraznění nepatrných výrazů v obličeji.

Nejsme ovšem omezeni pouze na oblast biologie a medicíny. Zheng et al. [9] se v laboratorních podmínkách, zabývají detekcí vad ve struktuře mostu. A zdaleka nejsou jediní. Qiwen Qiu [18] zkoumá využití EVM, v kombinaci se vzduchovou tryskou, pro automatickou detekci defektů vláknových kompozit. Obě metody přitom mohou být realizovány pomocí levně dostupných kamer.

Je samo sebou, že za uplynulých deset let se touto problematikou zabývaly spousty dalších studií, mimo jiné i práce inženýra Bernátka z fakulty biomedicíny VUT [5], a autor sám se o EVM poprvé dozvěděl z populárně naučného videa Steava Moulda [33], které za tu dobu naakumulovalo bezmála 1,7 milionu zhlédnutí. Přesto zůstává EVM žhavou oblastí výzkumu s potřebou inovativních řešení problémů, které, v době psaní této práce, limitují její jinak široký potenciál.

1 Měření životních funkcí

Ať už u okamžitého vyhodnocování nebo dlouhodobého sběru dat, často se dostáváme do situací, kdy potřebujeme nějakým způsobem vyhodnocovat životní funkce. A v takových případech se neobejdeme bez jejich měření. Jako životní funkce zde považujeme jakékoliv procesy zajišťující pokračování existence nějakého organismu, konkrétně se zde budeme zabývat organismem lidským a omezíme se na životní funkce jednoduše měřitelné pomocí EVM, tedy srdečního rytmu a dechové frekvence, které měří například také lůžkový monitor. Později se zaměříme speciálně na snímání srdečního rytmu a praktičnost jeho využití v monitorovacích systémech.

1.1 Srdeční rytmus

Nebo také "tepová frekvence"je periodické stahování a uvolňování srdečního svalu. Jde o nevědomou funkci samotného srdce, vyvolávanou spontánními elektrickými vzruchy ve speciálních buňkách umístěných přímo v srdci. Fáze stahování a uvolňování se odborně nazývají systola, kdy krev pod tlakem proudí ze síní do komor a z komor do aorty nebo plicnice, a diastola, což je období srdečního klidu mezi systolami. [27] Tento proces se cyklicky opakuje a má periodu asi 0,8 sekund, neboli frekvenci 1.25 Hz. Konvencí je uvádět srdeční rytmus v počtu úderů za minutu (anglicky: beats per minute - BPM) a odtud pochází průměrná tepová frekvence dospělého člověka, jako: $1.25 * 60 = 75$ BPM.

Věk	Bdělost [BPM]	Spánek [BPM]
novorozeně	100-180	80-160
kojenec	100-160	75-160
dítě	70-110	60-90
dospělý	60-90	50-90

Tab. 1.1: Tab 1.1: Fyziologické hodnoty tepové frekvence u různých věkových skupin [22]

Hodnoty v tabulce 1.1 odpovídají fyziologickým normám a můžeme podle odchylek od těchto hodnot diagnostikovat patologie.

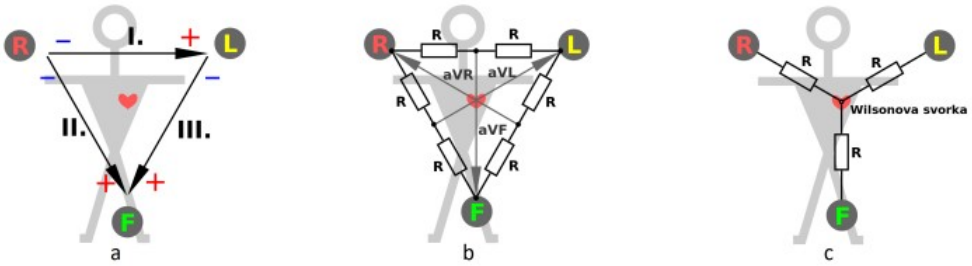
1.1.1 Měření srdečního rytmu

Dnes používané metody měření srdečního rytmu všechny vyžadují blízký kontakt s tělem, a snímají efekty způsobené tlukotem srdce jako jsou například elektrické vzruchy, změna okysličení krve, zvukové efekty nebo změna krevního tlaku.

Jednou z těchto metod je elektrokardiografie (EKG), která využívá měření změn elektrického potenciálu tlukoucího srdce, pomocí elektrod na povrchu těla. Ze znalosti tvaru jednoho celého cyklu jsme pak podle jeho opakování v čase schopni vypočítat tepovou frekvenci.

Citlivost EKG umožňuje určovat jednotlivé srdeční fáze. Velkou nevýhodou je ale nutnost použití elektrod, které se připevňují na kůži. Tyto elektrody se standardně umisťují podle třech svodů:

- končetinové bipolární svody
- hrudní unipolární svody
- končetinové unipolární svody

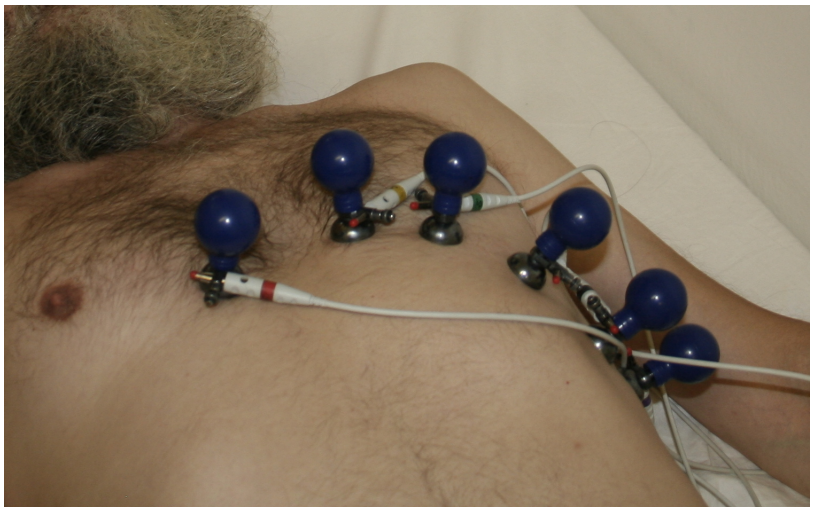


Obr. 1.1: EKG: a - bipolární (Einhovenův trojúhelník), b - unipolární (Goldbergovy svody), c - Wilsonova svorka [22]

Končetinové bipolární svody tvoří tzv. Einthovenův trojúhelník, jehož vrcholy (R, L, N/F) leží na končetinách. Napětí se měří bipolárně mezi dvěma vrcholy a to podle šipek znázorněných na 1.1 a.

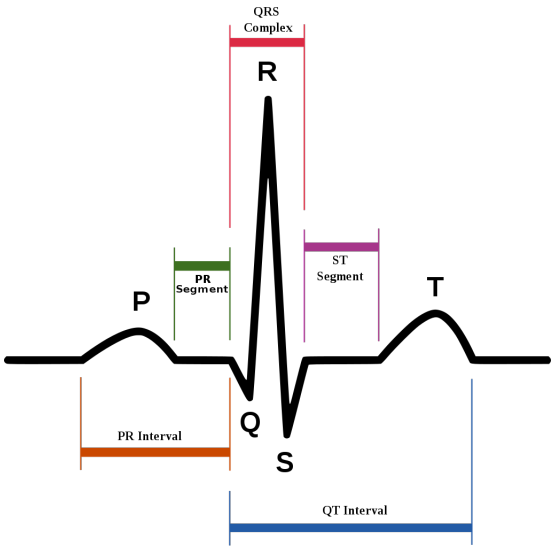
Končetinové, nebo také Goldbergovy unipolární svody umožňují lepší rozlišení. Referenční body jsou umístěné na protilehlých stranách vrcholů trojúhelníku a každý svod reprezentuje potenciál jedné příslušné elektrody 1.1 b.

Hrudní unipolární svody se značí měřením jednotlivých elektrod k referenčnímu bodu, tzv. Wilsonově svorce 1.1 c. Elektrody se pak umisťují většinou na levou stranu hrudního koše, přibližně sledujíc šesté žebro jak je znázorněno na 1.2.



Obr. 1.2: Umístění hrudních elektrod - Autor: Kychot – Vlastní dílo, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=7883095

Výstupem snímání je EKG křivka 1.3, kterou lze rozdělit do tří hlavních částí: P vlna, T vlna a QRS komplex. Patologické jevy se pak dají vypozorovat z tvarů těchto částí. R-R interval (vzdálenost špiček R vlny dvou po sobě jdoucích EKG křivek) se nabízí pro jednoduché určení periody a tedy i tepové frekvence.



Obr. 1.3: Signál EKG [28]

Další metodou je pulsní oxymetrie, která využívá kolísání objemu arteriální krve v tkáni v průběhu pulzní vlny.[29] Mimo tepovou frekvenci měří také obsah kyslíku v krvi pomocí rozdílu v absorpci spekter mezi okysličenou a neokysličenou krví. Měření se provádí pulzním oxymetrem, což je jednoduché zařízení složené ze zdroje světla emitujícího 660 nm a 940 nm a detektoru 1.4

Obr. 1.4: Beurer Pulzní oxymetr PO 80

A existují mnohé další metody, jako například fonografická nebo oscilační, u nichž se měří zvuková frekvence srdce a frekvence změn krevního tlaku respektive. Jak ovšem zaznělo na začátku, všechny tyto metody měření vyžadují blízky styk s kůží měřeného subjektu.

1.2 Dechová frekvence

Dechová frekvence (DF) je měřitelná veličina dýchání, při němž, za pomoci chemických a fyzikálních procesů jako například difuze plynů a osmotického tlaku, dochází k okysličování krve. Kromě dechové frekvence se dále měří třeba kapacita plic, rozdělená do několika kategorií, koncentrace vydechovaných plynů nebo odpor dýchacích cest. Funkci dýchání zajišťuje bránice, což je svalová membrána fungující podobně jako například membránová pumpa. Nádech pak probíhá stažením bránice, čímž se zvýší objem plic a nepřímo úměrně tlak v plicích. Tím pádem atmosférický tlak žene okolní vzduch do plic, dokud nedojde k vyrovnání tlaků. Po krátkém zamyšlení se to může jevit trochu zvláštní, přeci jen zadržíme-li dech, má vzduch v plicích tendenci uniknout. Vskutku, po uvolnění bránice dochází ke smrštění plic, což vede naopak k vyššímu tlaku v jejich nitru a důsledkem je výdech, který je proto normálně pasivním procesem. S trochou soustředění však můžete zkusit zadržet dech bez uzavření dýchacích cest, udržováním své bránice ve statické poloze (jen si to zkuste, je to celkem zvláštní pocit). Dýchání je tedy stejně jako tlukot srdce cyklický proces. Jak už jsme zmínili, výdech je většinou pasivní proces a tedy nedobrovolný - ztratíme-li vědomí, naše plíce se automaticky vyprázdní, na rozdíl od například kytovců! Lze ovšem vydechnout i aktivně pohybem břišních svalů. Nádech je vědomý, ale, stejně jako

výdech, za normálních okolností nedobrovolný proces. To proto, že lidské tělo využívá skupiny buněk silně citlivých na pokles pH krve v důsledku zvýšení koncentrace CO_2. A ačkoliv je vědomé řízení dýchání mozkem nadřazené řízení podvědomému, a je proto možné dýchání, na rozdíl od srdečního rytmu, značně ovlivňovat, při ztrátě vědomí začne dýchání automaticky po překročení krajní hranice koncentrace CO_2 - opět na rozdíl od kytovců! To se nám může stát osudným, ztratíme-li vědomí pod vodou.

DF je nedílnou součástí základního vyšetření pacienta. Při měření se hledí na rychlost nádechu, kvalitu a pravidelnost intervalů. Fyziologické hodnoty DF pro různé věkové kategorie jsou vypsány v tabulce 1.2 [22]

Věk	DF $[min^{-1}]$
novorozeně	50-60
kojenec	35-40
dítě	20-35
dospělý	16-20

Tab. 1.2: Fyziologické hodnoty dechové frekvence u různých věkových skupin. [22]

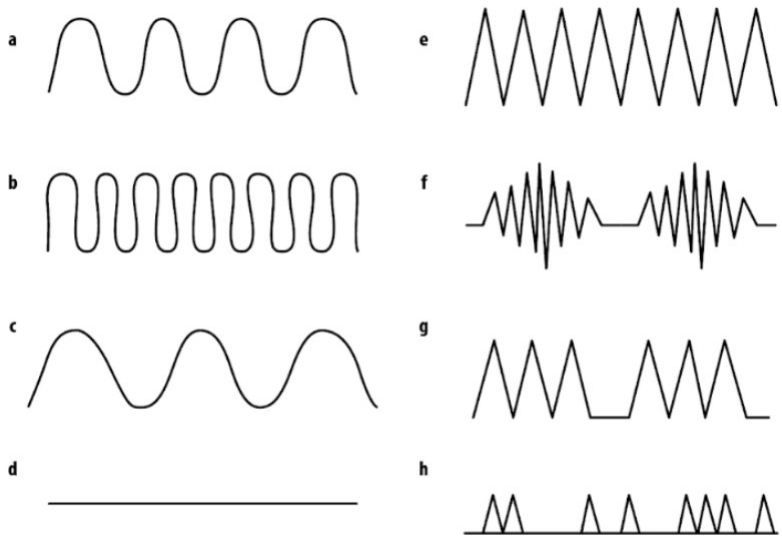
Podobně jako u tepové frekvence lze dechovou frekvenci využít ke sledování patologických jevů, jelikož při zvýšené zátěži organismu, třeba v důsledku nedostatku kyslíku v okolním vzduchu, dochází ke zvýšení DF. Lze tak potencionálně využít EVM pro včasnou detekci CO.

Pokud má DF nenormálně zvýšenou hodnotu, mluvíme o tachypnoe. Naopak u příliš nízkých hodnot se jedná o bradypnoe. Stav, kdy nedochází k dýchání vůbec se nazývá apnoe. Dále lze v DF pozorovat různá onemocnění jako například Kussmaulovo dýchání, Cheyneovo-Stokesovo dýchání nebo dušnost - odborně nazývanou dypnoe.

1.2.1 Měření dechové frekvence

Společně s tepovou frekvencí je dechová frekvence kritickou pro fungování organismu a proto ji lze využít pro monitorování pacientů a včasnou detekci života ohrožujících stavů. Některé studie dokonce hovoří o předpovídání srdeční zástavy pomocí zvýšené DF [21].

Měření DF opět vyžaduje blízký kontakt s pacientem. Jednou z možných metod je využití pneumografu, což je zařízení obepínající hrudní koš a snímající vlastní deformaci v důsledku dýchání. Využívá se zde odporových, tlakových nebo indukčních senzorů. Velkou nevýhodou je šum způsobený pohybem pacienta nesouvisejícím s dýcháním.



Obr. 1.5: Dechové křivky, a - normální dech, b - zrychlený dech, c - zpomalený dech, d - apnoe, e - Kussmaulovo dýchání, f - Cheyneovo-Stokesovo dýchání, g - Biotovo dýchání, h - agonální dýchání [22]

Dalšími metodami jsou spirometrie a pneumotachografie, které nepřímo měří DF pomocí měření objemu a rychlosti respektive, a směru proudění vzduchu. [22]

Dechovou frekvenci lze také měřit pomocí EKG a to díky fenoménu nazývaném sinusová arytmie. Jde o změnu tepové frekvence v závislosti na nádechu a výdechu - při nádechu se zvyšuje a při výdechu klesá. Prakticky lze takto měřit DF pouze u pacientů, kteří jsou v klidu.[30]

1.3 Shrnutí

V této kapitole jsme si vysvětlili dva fyziologické procesy, které jsou kritické pro správný chod organizmu a proto využitelné při sledování a vyhodnocování životních funkcí subjektů. Jde o srdeční rytmus a dechovou frekvenci.

V obou případech je u dnešních měřících přístrojů nutné být ve fyzickém kontaktu s měřeným subjektem, ať už pomocí epidermální elektrody nebo deformačních pásů.

Představili jsme si také hodnoty typické pro normální zdravý organizmus a jaké patologické jevy můžou výchylky od těchto hodnot předpovídat.

Dále jsme si vysvětlili princip různých metod, využívaných pro měření srdečního rytmu a dechové frekvence.

Potřeba fyzického propojení subjektu a měřícího přístroje je značně limitující faktor, který nás omezuje nejen ve vzdálenosti měřícího přístroje od subjektu, ale

také v jednoduchosti měření a ve zvýšené zátěži pro individuální péči. V následující kapitole si vysvětlíme jak funguje algoritmus, který by tyto problémy dokázal vyřešit a umožnit tak například jednoduché monitorování nemocničních pacientů, které dnes není praktické připojovat k lůžkovým monitorům.

2 Zvýraznění nepatrných změn pomocí EVM

Jelikož EVM může zároveň zvýrazňovat jak nepatrné změny v barvě, tak i ve tvaru, naskýtají se dvě možnosti jak extrahovat srdeční puls. Obě se opírají o detekci snímané osoby a mají své výhody i nevýhody k jejichž odstranění je nutné video sekvenci předzpracovat.

Při sledování změny tvaru hlavy, který se mění s tlukotem srdce, je například nutné 'pouze' najít obličej snímaného subjektu. Tento přístup testují Zhang et al. [6] přičemž jako hlavní přednost vynášejí, že jde o metodu robustní vůči dynamickým změnám osvětlení scény a - s pomocí akceleračního filtru - vůči výrazným pohybům subjektu, které jsou zásadní nevýhodou druhého přístupu.

Ten pracuje se zvýrazněním změn barvy kůže, taktéž přímo spjaté s tlukotem srdce, ovšem navíc k nalezení obličeje subjektu, respektive odhalené kůže, také vyžaduje nalezení optimálního bodu na snímané kůži (např. střed čela). Touto metodou se zabývá většina literatury, počínaje nejranější studií z MIT [1] a budeme se jí zabývat i my.

Tato dvě uplatnění jsou samozřejmě jen malý úvod do nadcházející kapitoly a praktických využití EVM v této oblasti je nespočet. Jako příklad může být uveden Abnousi et al. [7], jež se pokouší o lepší prevenci selhání srdce pomocí zvýraznění pohybu krční tepny, čímž zvyšují přesnost diagnóz ze strany profesionálních kardiologů nebo Maurya et al., kteří zkoumají využití EVM pro kontrolu dýchání novorozenců [17].

Obr. 2.1: Eulerovská video-magnifikace pro zvýraznění srdečního pulsu v obličeji. [1]

2.1 Lineární EVM

Je stále nejrychlejší typ eulerovské video-magnifikace, který vyvinuli Wu et al. [1] na MIT. Základní myšlenka vzešla z tehdy několik let starého algoritmu využívajícího Lagrangeovu metodu sledování lokálních pohybů v obraze. Na rozdíl od výpočetně náročné Lagrangeovy metody využívá Wu zpracovávání globálních prostorových změn v čase.

2.1.1 Zvýraznění pohybu [1]

Představme si 1D signál vykonávající translační pohyb.

Nechť $I(x,t)$ značí jas obrazu na pozici x v čase t. Pozorované jasové změny lze vyjádřit jako funkci $\delta(t)$ tak, že $I(x,t) = f(x + \delta(t))$ a $I(x,0) = f(x)$. Cílem zvýraznění pohybu je vytvořit syntetický signál

$$\hat{\mathrm{I}}(x,t) = f(x + 1(1+\alpha)\delta(t)). \tag{2.1}$$

pro faktor zesílení α.

Taylorovým rozvojem prvního řádu funkce $f(x+\delta(t))$ získáme jas obrazu v čase t podle x jako

$$I(x,t) \approx f(x) + \delta(t)\frac{\partial f(x)}{\partial x}. \tag{2.2}$$

Nechť $B(x,t)$ je $I(x,t)$ filtrovaný v čase pomocí pásmové propusti pro každou pozici x (z rovnice 2.2 vše kromě $f(x)$). Pro zatím předpokládejme, že pohybový signál, $\delta(t)$, spadá do neutlumené části pásmové propusti (tento předpoklad později uvolníme). Potom máme

$$B(x,t) = \delta(t)\frac{\partial f(x)}{\partial x}. \tag{2.3}$$

Zvýraznění se provede vynásobením filtrovaného signálu faktorem α a přičtením k původnímu signálu $I(x,t)$, z čehož plyne výsledný signál

$$\tilde{I}(x,t) = I(x,t) + \alpha B(x,t). \tag{2.4}$$

Sloučením rovnic 2.2, 2.3 a 2.4 získáme

$$\tilde{I}(x,t) \approx f(x + (1+\alpha)\delta(t)). \tag{2.5}$$

To znamená, že tento proces zvýrazňuje pohyb - prostorový posun $\delta(t)$ lokálního obrazu $f(x)$ v čase t, byl zvýrazněn na hodnotu $(1+\alpha)$

Taylorův rozvoj prvního řádu je dobrou aproximací pro relativně malý pohyb $\delta(t)$. V případě obecného příkladu, kdy $\delta(t)$ nespadá kompletně do nefiltrované oblasti pásmové propusti. Nechť $\delta_k(t)$ reprezentuje různé komponenty časového spektra $\delta(t)$. Každé $\delta_k(t)$ bude tlumeno s faktorem γ_k. Po aplikaci pásmové propusti získáme signál

$$B(x,t) = \sum_k \gamma_k \delta_k(t) \frac{\partial f(x)}{\partial x} \tag{2.6}$$

(stejné jako rovnice 2.3. Díky násobení v rovnici 2.4, lze útlum časových frekvencí přirovnat k faktoru zvýraznění α, závislém na frekvenci $\alpha_k = \gamma_k \alpha$, čímž získáme zvýrazněný výstup

$$\tilde{I}(x,t) \approx f(x + \sum_k (1+\alpha_k)\delta_k(t)) \tag{2.7}$$

Zvýrazňovací faktor pohybového signálu α se tedy stává modulárním faktorem α_k pro každé subpásmo pohybového signálu δ_k. Jinými slovy, jelikož jde o lineární analýzu, lineárně zvýrazňujeme také nechtěný pohyb, tedy šum.

2.1.2 Výběr faktoru α [1]

V praxi předpoklad z předchozí podkapitoly platí pro hladké obrazy a malé pohyby. V případě rychle se měnícího obrazu (vysokých prostorových frekvencí) $f(x)$, bude aproximace prvním řádem Taylorova rozvoje nepřesná pro vysoké hodnoty zvýraznění, $1 + \alpha\delta(t)$, které jsou výsledkem jak vyššího zvýrazňovacího faktoru α, tak většího pohybu $\delta(t)$. Pohyb lze částečně ošetřit zaměřením se na fázové rozdíly po sobě jdoucích snímků, nebo využitím stabilizačních algoritmů. Faktor α ale můžeme zvolit tak, aby byl optimální.

Pomocí prostorové frekvence, ω, můžeme vyvodit jak velký byl měl faktor α být v závislosti na požadovaném pohybu $\delta(t)$. Aby zvýraznění signál $\tilde{I}(x,t)$ byl přibližně věrný skutečnému zvětšení, $\hat{\mathrm{I}}(x,t)$, musí platit, že

$$\tilde{I}(x,t) \approx \hat{\mathrm{I}}(x,t) \Rightarrow f(x) + (1+\alpha)\delta(t)\frac{\partial f(x)}{\partial x} \approx f(x + (1+\alpha)\delta(t)) \tag{2.8}$$

Nechť $f(x) = cos(\omega x)$ pro prostorovou frekvenci ω a $\beta = 1 + \alpha$.

Požadujeme, aby

$$cos(\omega x) - \beta\omega\delta(t)sin(\omega x) \approx cos(\omega x + \beta\omega\delta(t)). \tag{2.9}$$

Z trigonometrie vyplývá, že

$$cos(\omega x) - \beta\omega\delta(t)sin(\omega x) = cos(\omega x)cos(\beta\omega\delta(t)) - sin(\omega x)sin(\beta\omega\delta(t)). \quad (2.10)$$

Měly by tedy platit následující aproximace

$$cos(\beta\omega\delta(t)) \approx 1 \quad (2.11)$$

$$sin(\beta\omega\delta(t)) \approx \beta\delta(t)\omega \quad (2.12)$$

Pro malé úhly $\beta\omega\delta(t) \leq \frac{\pi}{4}$ budou rovnice 2.11 a 2.12 platit s tolerancí 10% (jelikož $sin(\frac{\pi}{4}) = 0.9\frac{\pi}{4}$). V případě prostorové vlnové délky, $\lambda = \frac{2\pi}{\omega}$, pohybového signálu, získáme

$$(1 + \alpha)\delta(t) < \frac{\lambda}{8} \quad (2.13)$$

Rovnice 2.13 nám tedy udává nejvyšší hodnotu zvýrazňovacího faktoru, při zachování přesného zvýraznění požadovaného pohybu.

Vyšší α je ovšem žádoucí pro požadovaný signál, zatímco bychom chtěli, aby pro ostatní pohyby bylo jen malé. Ve výsledku tedy aplikujeme α na chtěné subpásmo prostorových frekvencí a snížíme jej podle 2.13, nebo potřeb uživatele, u pásem vyšších prostorových frekvencí.

2.2 Problém nežádoucích pohybů

Vzhledem k podstatě EVM je optimální snímaný subjekt takový, který se, kromě nepatrných pohybů, jinak nehýbe (např. při mrknutí bude extrémně zvýrazněn jak pohyb víčka, tak změna barvy v oblasti oka). Spolu s chybou měření a závislostí na osvětlení jde o jeden z hlavních problémů využití EVM v bezkontaktních měřících systémech.

Řešením je soustředit se především na zvýraznění prostorově fázových rozdílů pomocí tzv. Rieszovy pyramidy.

2.3 Fázová EVM

Jak bylo řečeno v úvodu, původní lineární eulerovská video-magnifikace naráží na zásadní problém se zvětšováním šumu zároveň s užitečným signálem a limituje se tak na menší zvětšení (ve formě užšího rozsahu α) a menší SNR (signal to noise ratio) finálního videa.

Stejně jako se rozdíly ve fázi mezi sinusoidy podobají translaci, se fázové rozdíly komplexní řízené pyramidy vztahují k lokálním pohybům v prostorovém subpásmu. [2] Jinými slovy, pokud rozložíme 2D snímek do několika odstupňovaných pásem, navzájem posunutých o 90° (tedy aplikujeme komplexní pyramidu), můžeme podle změn lokální fáze časových změn pixelů nalézt a zvýraznit oblasti s jinak nepostřehnutelným pohybem.

Zpracováváním fáze namísto amplitudy je pak možné dosáhnout až čtyřikrát vyššího zvětšení [2], aniž by se zvýrazňoval i šum. Mimo to je také možné pohyb ve videu utlumit, což je užitečné pro případ zvýrazňování barev. Nevýhodou je ovšem fakt, že je tato metoda navržená na měření periodických pohybů a také dvakrát vyšší výpočetní náročnost, způsobená využitím komplexních pyramid.

2.3.1 Zvýraznění pohybu pomocí fáze [2]

Fázová EVM závisí na komplexní řízené pyramidě (z ang. complex steerable pyramid) pro měření a modifikaci lokálních pohybů. Jedná se o omezení oproti lineární EVM v tom, že fáze nezvýrazňuje lokální změnu barvy. Jak ale uvidíme, dává nám mnohem více možností ve zvýraznění pohybu.

Pro lepší představu si znovu uvedeme 1D příklad, kdy je intenzita obrazu f ovlivněná globální translací, $f(x + \delta(t))$, pro nějakou translační funkci $\delta(t)$. Cílem je vytvořit syntetickou sekvenci se zvýrazněným pohybem, $f(x + (1 + \alpha)\delta(t))$, kde α je faktor zvýraznění.

Pomocí Fourierovi řady lze napsat f jako součet komplexních sinusoidů,

$$f(x + \delta(t)) = \sum_{\omega=-\infty}^{\infty} A_\omega e^{i\omega(x+\delta(t))} \tag{2.14}$$

kde každé pásmo koresponduje s jednou z frekvencí ω.

Z rovnice 2.14, pásmo pro frekvenci ω je komplexní sinusoid

$$S_\omega(x, t) = A_\omega e^{i\omega(x+\delta(t))} \tag{2.15}$$

Protože je S_ω sinusoid, jeho fáze $\omega(x + \delta(t))$ obsahuje informaci o pohybu. Podle Fourierova teorému se zpoždění v časovém spektru projeví jako posun ve fázovém spektru a naopak. Můžeme proto přes fázi manipulovat s pohybem, a tedy pohyb zvýraznit.

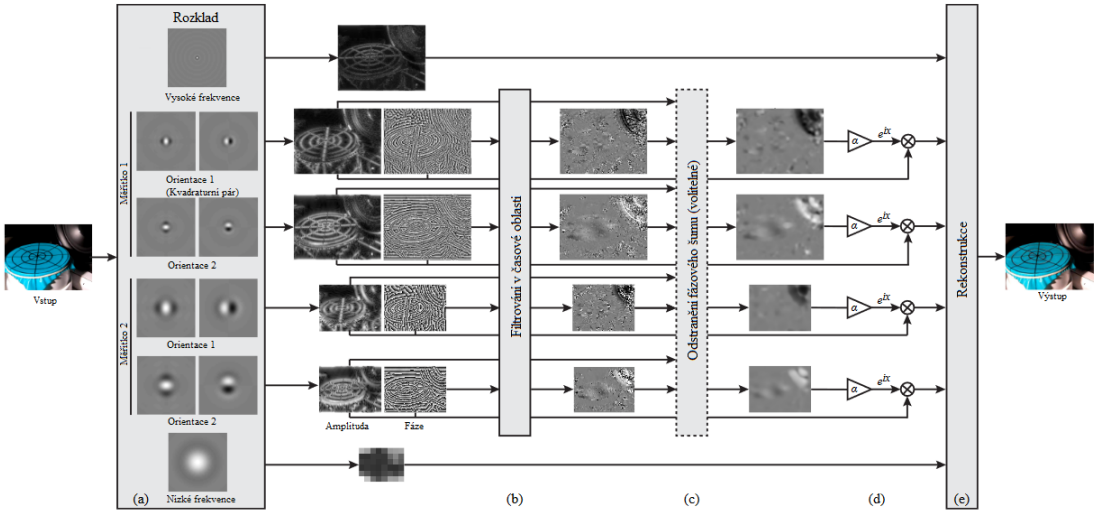
Izolování pohybu určitých frekvencí zajistíme filtrací fáze $\omega(x+\delta(t)$ v čase pásmovou propustí. Pro zjednodušení derivace uvažujeme, že časový filtr pouze odstraní DC složku ωx. Výsledkem je

$$B_\omega(x, t) = \omega\delta(t). \tag{2.16}$$

Vynásobíme vyfiltrovanou fázi $B_\omega(x,t)$ faktorem α a zvýšíme fázi $S_\omega(x,t)$, abychom dostali zvýrazněný pohyb v subpásmu

$$\hat{S}_\omega(x,t) := S_\omega(x,t)e^{i\alpha B\omega} = A_\omega e^{i\omega(x+(1+\alpha)\delta(t))} \tag{2.17}$$

Výsledkem je komplexní sinusoid, který má pohyb zvýrazněný přesně $(1+\alpha)$ krát oproti vstupní sekvenci. Výstupní sekvence se nakonec rekonstruuje kolapsem pyramidy. Po rozšíření do 2D se objevuje problém směrovosti použitých prostorových filtrů, proto musí obrazová pyramida být řiditelná, směrovost musí jít měnit.



Obr. 2.2: Rozkladem se oddělí amplituda lokálních vlnek od jejich fáze (a). Poté se fáze časově filtrují nezávisle na směru, lokaci a měřítku (b). Volitelně se aplikuje prostorové vyhlazení podle amplitudy, aby se zvýšilo SNR (c). Poté se zesílí nebo utlumí fáze filtrované pásmovou propustí (d) a video se rekonstruuje (e). [2]

2.3.2 Výběr faktoru α [2]

Při posunu pohybující se části obrazu, pomocí posuvu fáze, eventuálně narazíme na hranici za kterou tuhle část již posunout nemůžeme. Tohle také souvisí s omezeným uplatněním fázové EVM při řešení velkých pohybů v obraze.

Jako aproximaci posouvané oblasti si představme Diracův impulz provádějící translaci v čase, posouvaný fázovým posuvem Gaborových vlnek - to jsou komplexní sinusoidy modulované Gaussovou funkcí. Jak se Diracův impulz fázově posunuje, zároveň je zeslabován Gaussovou funkcí Gaborových vlnek. Tím pádem, si zvolíme hranici tak, aby docházelo jen k malému zeslabení.

1D Gaborova vlnka má přechodovou funkci ve frekvenční doméně

$$e^{-2\pi(\omega_x-\omega_0)^2\sigma^2}, \tag{2.18}$$

kde ω_0 je filtrovaná frekvence a $\frac{1}{\sqrt{2\sigma}}$ je šířka Gaussovy funkce ve frekvenční doméně. Inverzní Fourierova transformace pak dává impulzní charakteristiku v prostoru

$$S_\omega(x,0) = e^{\frac{-x^2}{2\sigma^2}} e^{2\pi i\omega_0 x}, \tag{2.19}$$

komplexní sinusoid v Gaussově obálce. Pro obecný případ Diracova impulzu posunutého o $\delta(t)$ v čase t platí

$$S_\omega = e^{\frac{-(x-\delta(t))^2}{2\sigma^2}} e^{2\pi i\omega_0(x-\delta(t))} \tag{2.20}$$

Aplikací pásmové propusti na fázi v čase 0 a v čase t, dává

$$B_\omega(x,t) = 2\pi\omega_0\delta(t), \tag{2.21}$$

syntéza fázových rozdílů pro zvýraznění pohybu pomocí α je potom

$$2\pi\omega_0\alpha\delta(t) \tag{2.22}$$

Tento rozdíl fází koresponduje s posunem Diracova impulsu o dalších $\alpha\delta(t)$ pixelů. Musíme tedy ohraničit posun $\alpha\delta(t)$ tak, aby aproximace posuvu odpovídala skutečnému pohybu. Za použití jedné směrodatné odchylky Gaussova rozložení bude mít výsledný signál zhruba 61% původní amplitudy a máme tedy

$$\alpha\delta(t) < \sigma \tag{2.23}$$

V případě "octave-bandwidth"řízene pyramidy je pod Gaussovou obálkou přibližně jedna perioda sinusoidu. Což nám dává hranici $\alpha\delta(t) < \frac{1}{4\omega_0}$. Obrácenou hodnotou je prostorová vlnová délka $\lambda = \frac{1}{\omega_0}$ a dostáváme

$$\alpha\delta(t) < \frac{\lambda}{4} \tag{2.24}$$

Z rovnice 2.24 můžeme vypozorovat, že nízké prostorové frekvence lze zvýrazňovat více než vysoké frekvence.

Toto zjednodušení Diracova impulzu posouvaném Gaborovou vlnkou lze rozšířit na 2D případ a reálná videa. Při překročení hraniční podmínky dané 2.24 začne docházet o vytváření artefaktů.

V případě "half octave-bandwidth"pyramidy jsou pod Gaussovou obálkou dvě periody. A tedy $4\sigma \approx \frac{2}{\omega_0}$, z čehož vyplývá, že

$$\alpha\delta(t) < \frac{\lambda}{2}. \tag{2.25}$$

Což umožňuje zvýšení faktoru zvýraznění α oproti Wu et al. [1] až 4 krát. Výměnou za to je bohužel zpomalení algoritmu z důvodu zvýšeného počtu filtrů v pyramidě.

2.3.3 Fázová EVM pomocí Rieszovy pyramidy [3]

Výše uvedené nevýhody je možné zmírnit použitím speciálního typu komplexní pyramidy, využívající kvaternionů: Rieszovy pyramidy (podle maďarského matematika Frigyese Riesze).

Ta se poskládá z prostorových subpásem, získaných použitím invertibilní náhrady za Laplacovu pyramidu (která se používá v původní studii z MIT [1]), a následnou aproximací Rieszovi transformace pro všechna pásma.

Jelikož jde o speciální typ Hilbertovy transformace, tak tímto způsobem automaticky získáme kvadraturní pár, navzájem posunutý o 90°, v závislosti na dominantní orientaci pixelů.

Aproximace Rieszovi transformace pro dvě dimenze se dosáhne pomocí přechodových funkcí:

$$-i\frac{\omega_x}{||\vec{\omega}||}, -i\frac{\omega_y}{||\vec{\omega}||} \tag{2.26}$$

kde $\vec{\omega} = [\omega_x, \omega_y]$ jsou subpásma ve frekvenční doméně. Po aplikaci přechodových funkcí z 2.26, získáme výstup Rieszovi transformace $(R_1; R_2)$, který společně s intenzitou vstupního obrázku subpásma, vytváří sférické souřadnice:

$$I = Acos(\Phi); R_1 = Asin(\Phi)cos(\Theta); R_2 = Asin(\Phi)sin(\Theta) \tag{2.27}$$

Z těchto koeficientů můžeme vytvořit kvaternion $\mathbf{r}$.

$$r = I + iR_1 + jR_2 \tag{2.28}$$

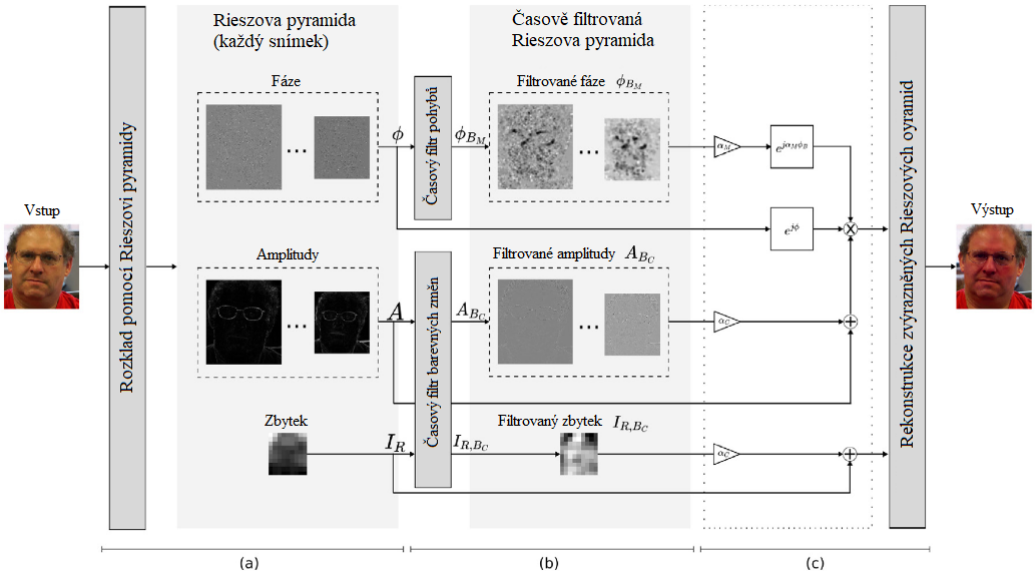
$$r = Acos(\Phi) + iAsin(\Phi)cos(\Theta) + jAsin(\Phi)sin(\Theta) \tag{2.29}$$

Po odstranění znaménkové nejasnosti získáme amplitudu a fázi:

$$A = ||r|| \tag{2.30}$$

$$i\Phi cos(\Theta) + j\Phi sin(\Theta) = log(r/||r||) \tag{2.31}$$

Následně jsou fáze každého pixelu rozbaleny a filtrovány pomocí LTI (linear time-invariant) filtru. Fázové rozdíly se vynásobí a přičtou k původním fázím a pyramida se rekonstruuje do finálního videa.



Obr. 2.3: (a) Rozklad jednotlivých snímků pomocí Rieszových pyramid: fáze, amplituda a zbytek. (b) Časové filtrování fáze, amplitudy a zbytku každého pixelu. (c) Rekonstrukce po zvýraznění faktorem α_M pro pohyb a α_C pro intenzitu/barvu. [4]

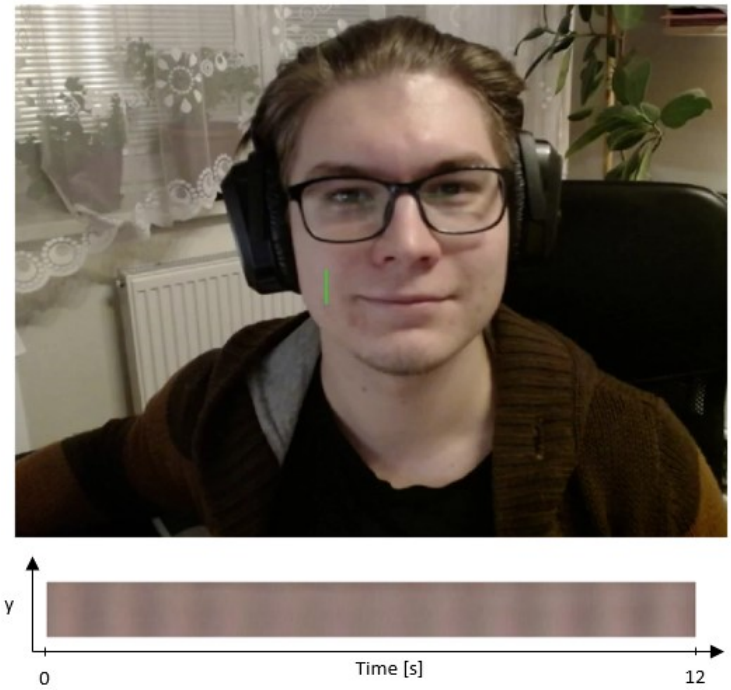
2.4 Ukázka využití

Jako příklad je zde uvedeno video zpracované původní lineární EVM z MIT. Video bylo natočené webkamerou Logitech C920 s rozlišením 600x480 pixelů, v nekvalitně osvětlené místnosti, s frekvencí 30 snímků za vteřinu po dobu 10 sekund.

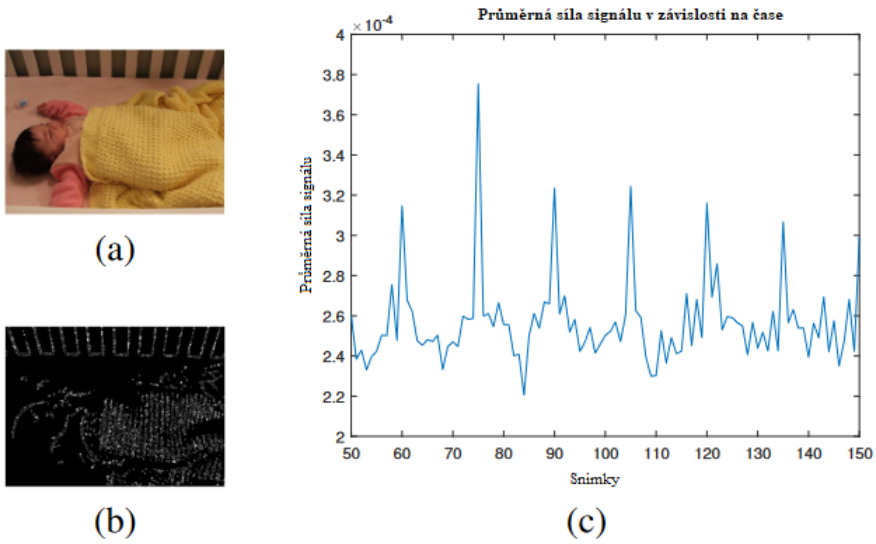
Doba zpracování se pohybovala okolo jednotek minut pro 2.4, video v HD kvalitě se pak zpracovávalo celou hodinu. Je tedy hezky vidět, že původní algoritmus nebyl navržený na aplikace běžící v reálném čase.

Z časového průběhu barevně zvýrazněné části jde odhadnout srdeční rytmus zhruba 75 bpm, což odpovídá průměrným klidovým hodnotám. Pozdější měření budou porovnávána s údaji z pulzního oxymetru.

Kromě měření změn barev, lze také měřit nepatrné změny pohybu, k čemuž dochází v 2.5 k měření dýchací frekvence spícího dítěte.



Obr. 2.4: Příklad natočený webkamerou Logitech C920. Dole je časový průběh změn barvy kůže. Zeleně je detekovaná oblast na pravé tváři.



Obr. 2.5: Měření DF dítěte. (a) Video záznam spícího dítěte. (b) Pohyb zvýrazněný fází. (c) Získaný signál. [3]

2.5 Shrnutí

Jak jsme si ukázali, eulerovská video-magnifikace má v oblasti biomedicíny široké uplatnění a zároveň, jako každá nová technologie, naráží na spousty problémů. Mezi tyto problémy patří například zesílení šumu, což je možné vyřešit přechodem na zvětšování fázových rozdílů. Kromě zlepšení SNR se tímto způsobem také podařilo udělat menší pokrok i v problematice nežádoucích pohybů, které je možno ve fázové oblasti utlumovat.

Na druhou stranu došlo k výraznému zvýšení výpočetních nároků. Za účelem vyřešení tohoto problému byla vymyšlena metoda s využitím Rieszových pyramid a kvaternionů, které ovšem vyžadují alespoň přibližně konstantní trajektorii snímaných objektů a navíc jde o poměrně komplikovaný matematický aparát. Fázová EVM také hůře zvládá pracovat s neperiodickými signály.

Navzdory všem nevýhodám je ale bezkontaktní snímání životních funkcí, a jiných nepostřehnutelných změn v obraze, pomocí levných kamer příliš lákavé. Navíc již existují privátní subjekty nabízející služby založené na právě této technologii, takže je o inovace v oblasti eulerovské video-magnifikace potencionálně i komerční zájem.

3 Detekce subjektu

Aplikace EVM v reálném čase jde uskutečnit, pokud jsme schopni získat výřez měřeného subjektu. Kromě toho jde také o jednu z možných metod odstranění velkých pohybů. Otázkou je, zda bude zrychlení EVM dostatečné na to, aby ospravedlnilo použití dalšího výpočetně náročného algoritmu detekce osob.

Po úspěšné detekci osoby a stabilizaci prostorového výřezu v čase, lze hledat odhalenou kůži například podle typických barevných tónů nebo přímo aplikovat EVM a zkoušet Fourierovu transformaci na oblastech s nejpravděpodobnějším výskytem signálu. Alternativně je možné detekovat přímo obličej a měřit signál z oblasti čela nebo tváří.

Nabízí se různé algoritmy, které jsou podporované v python knihovnami jako například OpenCV nebo MediaPipe. Jedním takovým je algoritmus Viola-Jones, který využívá Haarových příznaků a který s úspěchem použil inženýr Bernátek [5]. Dnes už ale existují daleko efektivnější metody a tak se v následujících kapitolách podíváme jak na zastaralý Viola-Jones, který nevyužívá naplno možností strojového učení, tak na algoritmus, který právě díky strojovému učení nevědomky používají miliony, ne-li miliardy lidí, díky jeho oblibě na sociálních médiích.

3.1 Viola-Jones detektor obličeje

Je objektový detektor, se kterým v roce 2001 přišli Paul Viola a Michael Jones. [31] Viola-Jones pracuje nejlépe s obrazem, ve kterém je obličej natočený přímo ke kameře a ne příliš šikmý.

Přesto jde o poměrně robustní algoritmus, který je možné sbíhat v reálném čase. Skládá se ze čtyř kroků:

1. Detekce Haarových příznaků
2. Výpočet integrálního obrazu
3. Adaboost trénování
4. Kaskádové zapojení klasifikátorů

3.1.1 Haarovy příznaky

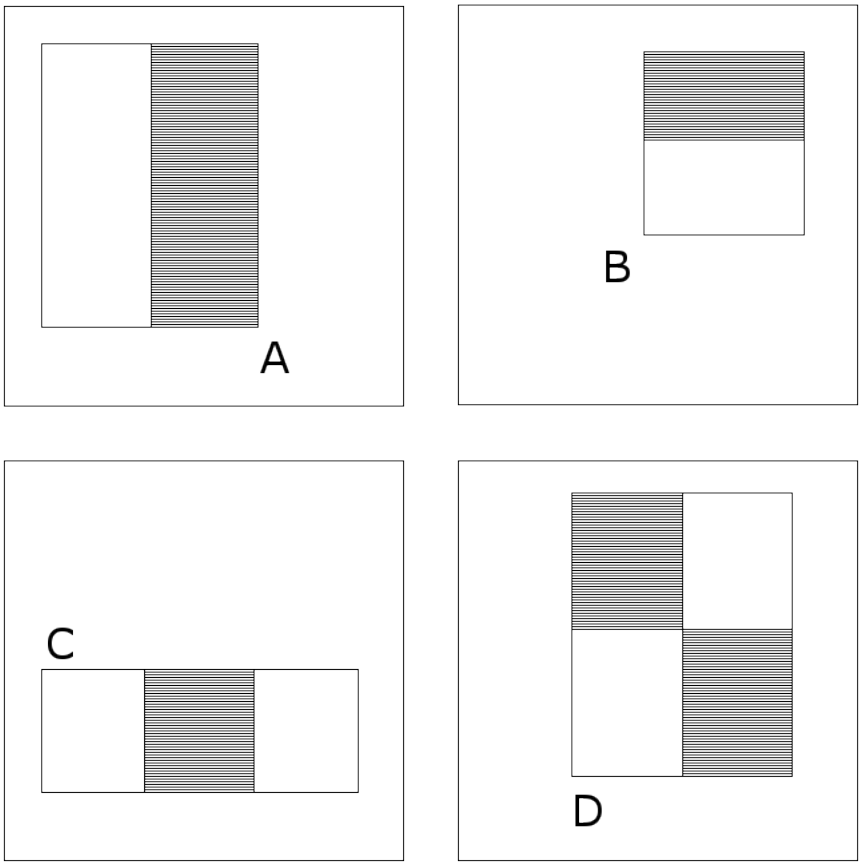
Znázorněny na obrázku 3.1, jde o masky složené z bílých a černých obdélníků, které se přikládají na zpracovávaný obraz. Za své jméno vděčí podobnosti s Haarovými vlnkami.

Využívají se pro detekci charakteristických rysů lidského obličeje (například toho, že no je většinou světlejší než oblast pod očima). Hodnota příznaku se vypočítá jako

rozdíl součtů jasů obrazu pod černou částí příznaku a součtu jasů pod bílou částí příznaku

$$f(x) = \sum_{w=1}^{i} x(w) - \sum_{b=1}^{j} x(b), \tag{3.1}$$

kde $\sum_{w=1}^{i} x(w)$ je suma pod bílou částí a $\sum_{b=1}^{j} x(b)$ suma pod černou.



Obr. 3.1: Příklad Haarových příznaků. B - příznak podobný jasovému rozdílu mezi očima a lícem, C - příznak podobný nosu. [31]

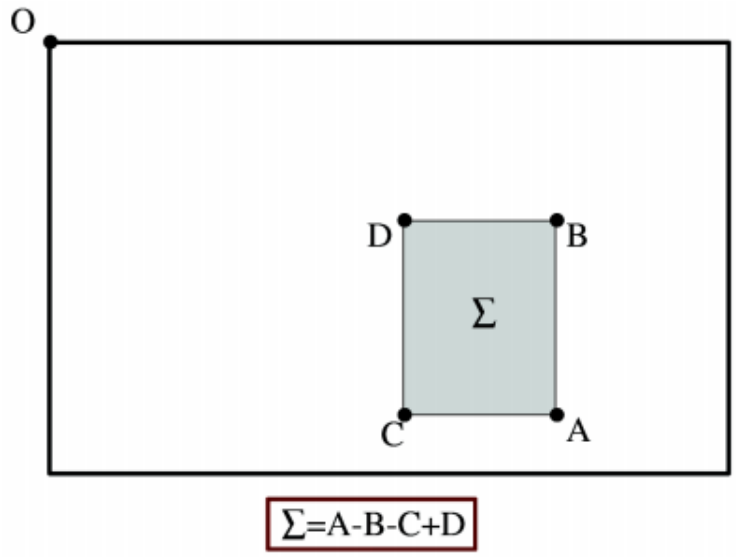
3.1.2 Integrální obraz

Integrální obraz je forma reprezentace obrazu. Každý bod integrálního obrazu je součet hodnot obrazových bodů nad a vlevo od odpovídajícího bodu v původním obrazu. Tato reprezentace umožňuje rychlé (výpočetně nenáročné) generování hodnot obrazových bodů integrálního obrazu. Hodnota obrazových bodů v každém obdélníku integrálního obrazu lze vypočítat pouze se čtyřmi hodnotami integrálního

obrazu. Hodnotu každého bodu integrálního obrazu lze vypočítat pomocí rovnice [5]

$$I_s(x) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i,j), \tag{3.2}$$

kde $x = (x, y)$ je pozice v integrálním obraze $I_s(x)$ a $I(i, j)$ je vstupní obraz.[32].



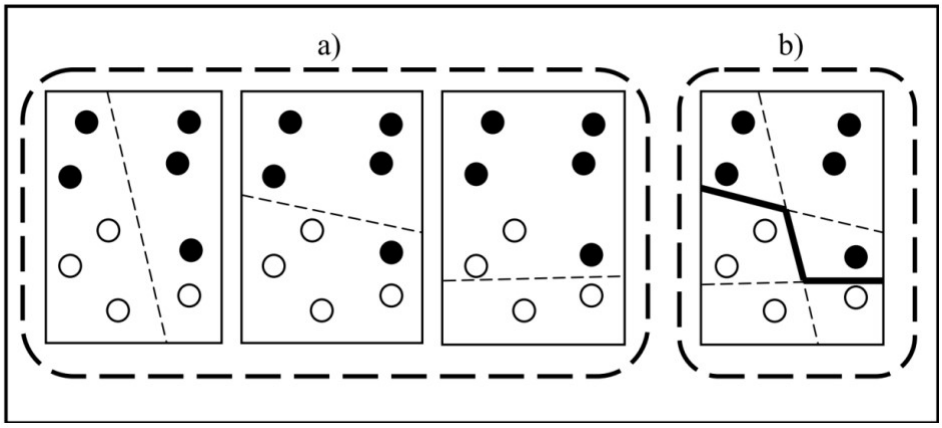
Obr. 3.2: Výpočet hodnoty plochy $\sum$ vymezenou vrcholy A, B, C a D v integrálním obraze. [32]

3.1.3 AdaBoost

Algoritmus AdaBoost je v detektoru Viola Jones využit jako klasifikační algoritmus. Účelem algoritmu je co největší zesílení (boosting) klasifikace, tedy výběr nejvhodnějších příznaků pro vytvoření klasifikátoru. Jako vstup pro algoritmus jsou využity jednotlivé příznaky z trénovací množiny spárované s informací o třídě příznaku. Poté jsou všem příznakům rovnoměrně nastaveny váhy. Takto vznikají tzv. slabé klasifikátory, jejichž přesnost přesahuje hodnotu 50%, což je přesnost odhadu. V každé následující iteraci je vybrán klasifikátor s nejmenší chybou, nepřesahující hodnotu 0,5. Podle něj jsou pak nastaveny váhy ostatních klasifikátorů tak, aby klasifikátor s nejmenší chybou měl největší váhu. Tímto je dosaženo toho, že v dalším kroku bude hledán takový slabý klasifikátor, který bude lépe klasifikovat správná měření. Požadované efektivity je dosaženo následnou kombinací těchto slabých klasifikátorů, z čehož vznikne silný klasifikátor s požadovanou přesností, viz obrázek 8. Hodnota

silného klasifikátoru je rovna součtu všech slabých klasifikátorů, které zahrnuje, vynásobených jejich vahami, viz rovnice 3.3 [5]

$$H(x) = \sum_{i=1}^{k} h(x_i)w(x_i) \tag{3.3}$$



Obr. 3.3: a) slabé klasifikátory; b) silný klasifikátor [5]

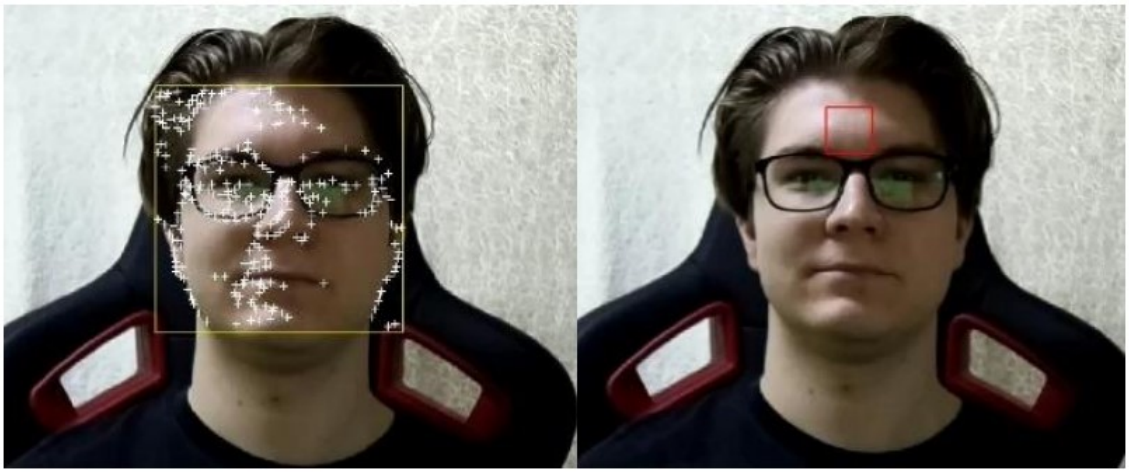
3.1.4 Kaskádové zapojení klasifikátorů

Pomocí kaskádového zapojení klasifikátorů v detektoru Viola Jones se docílí podstatného zefektivnění práce. Protože se obraz zpracovává po dílčích obdélníkových oblastech a ne jako celek, je zbytečné pokračovat v práci v oblastech, ve kterých nebyl hledaný objekt nalezen.Jednodušší a rychlejší klasifikátory jsou umístěny na začátku kaskády a na konci bývají klasifikátory složitější a pomalejší. Šetří se tak časová náročnost klasifikace, protože se náročnější klasifikátory již nemusí zbytečně zabývat oblastmi bez hledaného objektu. Každý z kaskádních klasifikátorů je trénován samostatně pomocí výše popsaného algoritmu AdaBoost. [5]

3.1.5 Ukázka využití

Původně bylo zamýšleno využít právě Viola-Jones algoritmus pro detekci a sledování pózy obličeje a tím pádem vyřešení problému s velkým pohybem při zpracovávání pomocí EVM. Výsledky 3.4 nebyly špatné a budeme se jimi krátce zabývat později. Viola-Jones nicméně existuje už přes 20 let a za tu dobu učinilo strojové vidění velký krok vpřed.

V následující kapitole se proto podíváme na detektor obličeje, který byl nakonec zvolen pro finální projekt, díky jeho rychlosti, jednoduchosti využití a robustnosti.



Obr. 3.4: Vlevo: Obličej detekovaný pomocí Viola-Jones a významné body využité ke sledování pózy obličeje. Napravo: Obdélníková oblast na čele významná pro pozdější zpracování EVM algoritmem. [23]

3.2 MediaPipe Face Mesh

Řeč je o Face Mesh algoritmu od MediaPipe, který je schopen odhadnout 468 3D bodů, charakteristických pro obličej. A to i na mobilních platformách, díky čemuž se s touto technologií už pravděpodobně setkali i ti z vás, kteří se o detekci obličeje nezajímají.

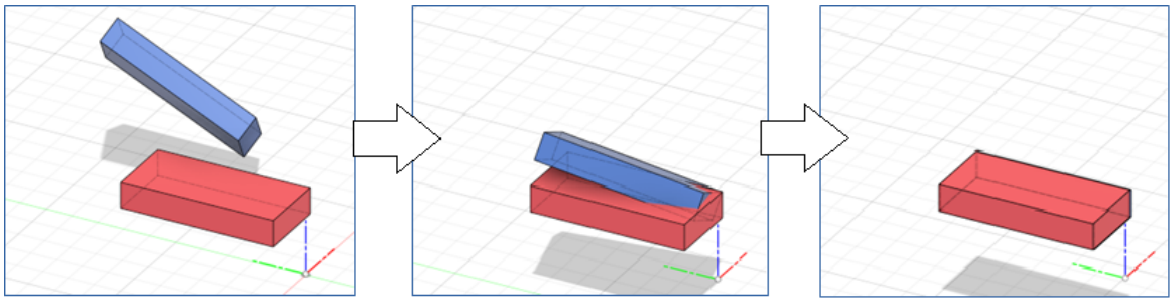
MediaPipe využívá strojového učení - využívá dvou modelů, o kterých si povíme zanedlouho - pro odhad 3D obrysu obličeje, za využití jedné kamery, bez senzoru hloubky. Využívá kompaktního modelu za použití GPU akcelerace a je tak možné snadno běžet detekci obličeje a jeho pózy i v reálném čase.

Jejich Python knihovna navíc obsahuje "Face Transform"modul, který nám umožňuje jednoduše pracovat s detekovanými body a vypočítat jejich pózu za pomoci analytické metody nazývané Prokústésova analýza.[25]

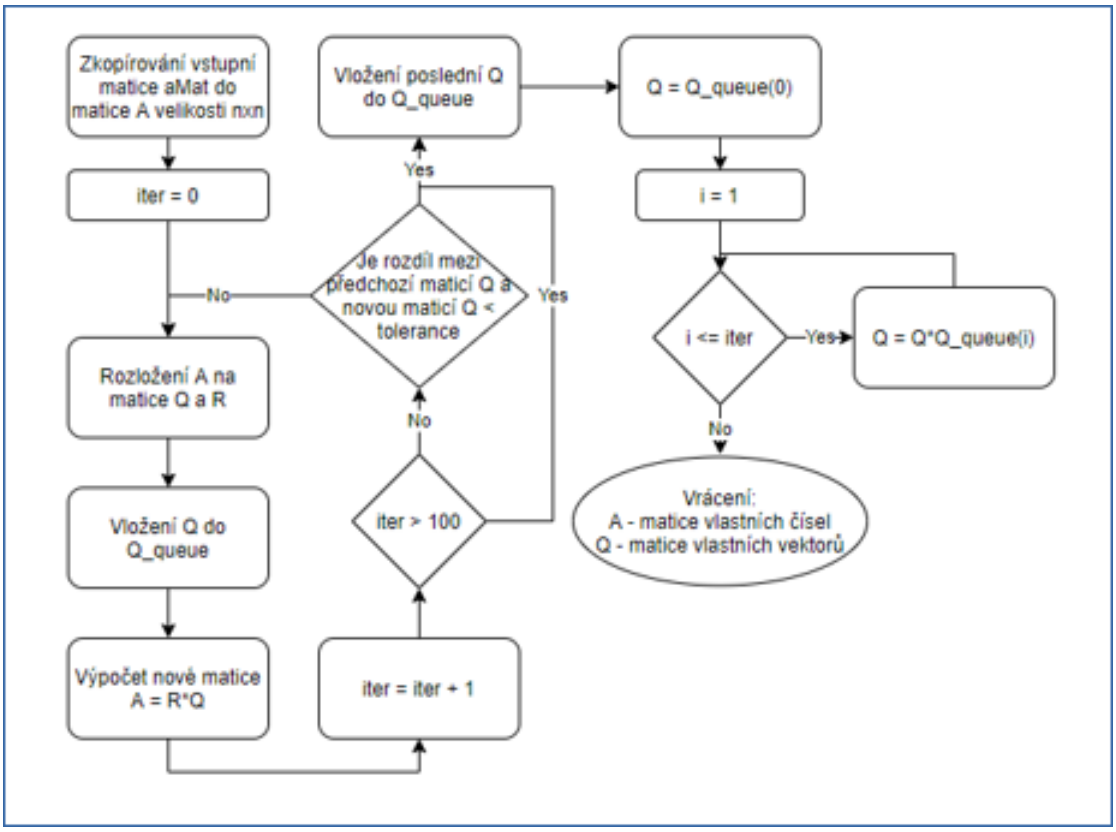
3.2.1 Prokústésova analýza

Prokústésova analýza úzce souvisí s ICP (iterative closest point) algoritmy a jejím cílem je nalezení lineární ortogonální matice transformace, která namapuje jeden objekt (respektive jeho bodové mračno) na druhý. Její název pochází z legendy o řeckém banditovi, který své oběti osekával nebo natahoval tak, aby se přesně vlezly na jeho postel.

Speciálním případem Prokústésovy analýzy je Kabshův algoritmus, který dovoluje pouze rigidní transformaci 3.6



Obr. 3.5: Průběh generického ICP algoritmu, řešící Prokústésův problém. Druhý krok se opakuje dokud nenastane tolerance blízká třetímu kroku.



Obr. 3.6: Jednotlivé kroky Kabshova algoritmu, specifického případu Prokústésovy analýzy. [24]

3.2.2 Modely MediaPipe

MediaPipe využívá dvou modelů hlubokých neuronových sítí. První detekuje obličeje v celém obraze a předává jejich přibližnou polohu druhému modelu, který ve výřezu hledá 3D obrys lidského obličeje.

Jakmile je obličej nalezen prvním modelem a ohraničen druhým, je dále možné

snížit výpočetní náročnost sledováním změny polohy nalezených charakteristických bodů. První model není nutné volat dokud je druhý schopen identifikovat obličej. MediaPipe dále umožňuje grafickou vizualizaci detekovaného obličeje.

3.2.3 Detekce obličeje

MediaPipe využívá BlazeFace [26], detektor obličeje navržený pro hladký běh na GPU mobilních zařízení. Ten je schopný detekovat 6 významných bodů na obličeji (středy očí, uši, střed úst a špičku nosu) pomocí kterých odhaduje rotaci obličeje.

Pro detekci charakteristických 3D bodů na obličeji, MediaPipe naučili neuronovou síť na syntetických 3D datech a 2D, předem určených, skutečných obrazech. Výsledkem je model schopný s dobrou přesností odhadovat podstatné 3D body na skutečném obličeji.

Tento model má na vstupu výřez obličeje z obrazu bez informace o hloubce. Výstupem jsou pak pozice 3D bodů a pravděpodobnost výskytu obličeje ve výřezu. Pro zvýšení přesnosti a robustnosti pak MediaPipe používá bootstrapping.

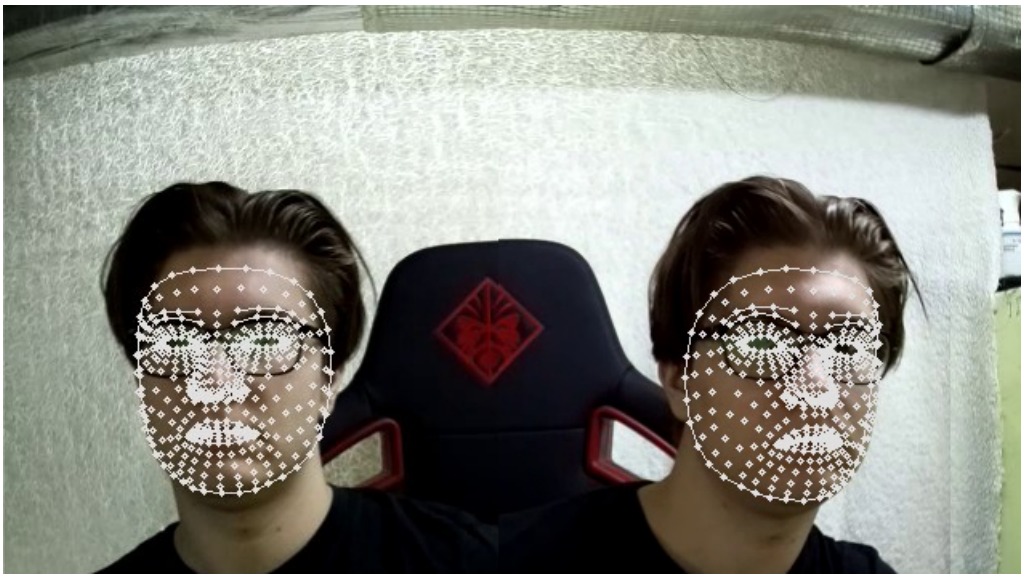
Za ceny vyšší výpočetní náročnosti je navíc možné ještě přesněji odhadovat oblasti okolo rtů, očí a očních duhovek.[25]

3.2.4 Ukázka využití

Jako test bylo zvoleno video zpracovávané původní metodou Viola-Jones + EVM. Oproti původní metodě došlo ke značnému zrychlení detekce obličeje a jako bonus nadále není nutné odhadovat polohu čela, jelikož MediaPipe Face Mesh čelu přiřazuje vlastní bod. Díky automatickému výpočtu pózy odpadá také tento problém spojený s Viola-Jonesem.

Mimo výše zmíněné je také MediaPipe, na rozdíl od OpenCV, mnohem přívětivější k rapidnímu vývoji, v němž Python exceluje.

Na obrázku 3.7 je vidět detekovaný obličej ze dvou různých úhlů. Bod na čele zůstává relativně statický. Video bylo natočeno za kontrolovaných podmínek kamerou Raspberry Pi HQ v rozlišení 600x480 pixelů.



Obr. 3.7: Obličej detekovaný pomocí masky ze dvou různých úhlů.

3.3 Shrnutí

Byly představeny dva algoritmy pro detekci tváře ve videozáznamu. Prvním z nich je Viola-Jones, který využívá Haarových příznaků pro detekci charakteristických rysů obličeje.

Detailně jsme si vysvětlili jak Viola-Jones funguje, jaké jsou jeho nedostatky - potřeba dívat se přímo do kamery a malý náklon hlavy - a proč od něj bylo nakonec upuštěno.

Druhým algoritmem je Face Mesh detektor podporovaný MediaPipe. Zde jsme si vysvětlili fungování algoritmů běžících v pozadí, modely strojového učení, které MediaPipe používá a jakým způsobem dosahuje tak rychlé detekce a sledování.

MediaPipe Face Mesh se nepřekvapivě ukázal jako jednoznačný vítěz a proto byl taky zvolen.

4 Řešení

V této kapitole se podíváme na průběh řešení, od raného návrhu po finální produkt. Nejprve se zaměříme na hardwarovou část skládající se z výpočetní jednotky, kamery, uživatelského rozhraní a zdroje energie. Poté se podíváme na softwarovou část, která se dělí na předzpracování obrazu a aplikaci EVM. Nakonec se budeme věnovat realizaci měření, problému se sběrem dat a podmínkami, za kterých byla většina dat sbírána.

4.1 Výpočetní jednotka

Vzhledem k předchozím zkušenostem s implementací zjednodušené verze eulerovské video-magnifikace bylo od začátku jasné, že řešení bude vyžadovat výpočetně výkonnější hardware, než který lze najít na komerčně dostupných mikro-kontrolerech jako jsou čipy od Atmel nebo Espressif.

V předchozím prototypu bylo dostačující využít Raspberry Pi Zero W. Ale, protože se tato práce zabývá kompletní implementací sledování obličeje a EVM algoritmu a jejich využitím v monitorovacích systémech, bylo zvoleno Raspberry Pi 4 Model B s 8 GB RAM, které zároveň naplno využije potenciál Python knihovny od MediaPipe.

Použití malých, lehce přenosných počítačů jde v duchu tzv. 'edge computing', tedy decentralizaci zpracovávání dat. V praxi to znamená, že podobný monitorovací systém by většinu výpočtů zvládl sám a měl by možnost posílat extrahovaná data do centrálního systému k libovolnému dalšímu zpracovávání.

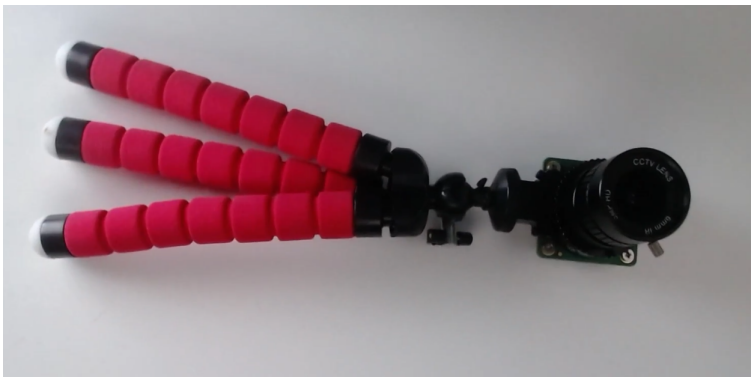
Vývoj proníhal na stolním PC s procesorem AMD Ryzen 5 a 16 GB RAM v prostředí Windows 10 a výsledný program byl poté jednoduše exportován na Raspberry Pi.

Obr. 4.1: Raspberry Pi 4 Model B

4.2 Kamera

Navzdory závěru, který shledali Ramya a Anudev [11], pro praktické využití ve sledovacích systémech není možné zvolit jakoukoliv kameru. To proto, že Ramya a Anudev ve své studii přímo nahrávají záznam kůže a vyhýbají se tím problému kůži nejprve detekovat. Kromě toho, v jejich specifickém případě vyplňuje oblast zájmu většinu záběru, kdežto v praktickém využití může jít jen o procentuálně malou část z celé scény. Jako zajímavý poznatek z výše uvedené práce stojí za zmínku využití infračervených kamer pro monitorování dýchání v noci.

Byla zvolena Raspberry Pi HQ kamera se senzorem Sony IMX477, který má rozlišení 12 Mpx. Je připojen k Rasberry Pi pomocí FPC konektoru. Zvolená kamera má vyměnitelný objektiv a je také možné ji nevratně modifikovat pro snímání IR záření.



Obr. 4.2: Raspberry Pi HQ se senzorem Sony IMX477

4.3 Uživatelské rozhraní

Jelikož je žádoucí, aby celý systém byl přenosný a interaktivní, vyžaduje se použít nějaké formy rozhraní mezi člověkem a počítačem (HMI - human machine interface). Zde se nabízelo vícero možností.

Nejjednodušším řešením by bylo použít levný 16x2 LCD display, stejně jako tomu bylo u prvního prototypu. Zde by ovšem došlo k plýtvání potenciálem, jelikož bychom se limitovali na snímání výlučně jedné osoby, bez možnosti výběru. Výhodou, oproti konečnému řešení, nicméně zůstává cena. A Pokud by bylo účelem vyvinout levné zařízení podobné bezkontaktnímu teploměru, je možné takový display zkombinovat s velmi levnou kamerou.

V této práci jsme se nicméně vydali cestou dotykového displaye. Důvod je především to, že použitá kamera 4.2 umožňuje detekci a snímání více subjektů najednou a dotykový display nám dává možnost selektivně sledovat nebo ignorovat kteroukoliv osobu v záběru.

Výhodou použití přenosného displaye je navíc možnost sběru dat na odlehlých místech, což se ukázalo důležitým především při získávání kamerových záznamů různých tónů pleti.

4.4 Zdroj

S volbou přenosného displaye se stalo nepraktickým napájet náš systém zdrojem přímo ze sítě. Dnes jsou naštěstí hojně k dostání přenosné zdroje energie s dostatečně velkou kapacitou.

Raspberry Pi 4B při zatížení všech čtyř jader konzumuje až $6W$ při $5V$, to je $1.2A$, dotykový display potřebuje maximálně $0.5A$ a HD kamera $0.25A$. To zaokrouhleně dává $2A$ maximálního odběru proudu.

Obr. 4.3: Waveshare 7"IPS LCD (H) display

Nejlevnější powerbanky mají kapacitu $20Ah$, to nám dává 10 hodin funkčního času za nejhorších podmínek, což je dostačující pro demonstraci i testování.

Jednou z nástrah moderních powerbank je, že všechny levné modely používají stejný čip vyráběný v Číně a tento čip powerbanku automaticky vypíná po $30s$ pokud odběr nepřekročí $40\mu A$. To v našem případě není problém, protože i za volného chodu spotřebovává Raspberry Pi okolo $0.5A$.

4.5 Snímač tepové frekvence

Aby bylo možné s jistotou říct, že signál získaný po aplikaci EVM skutečně odpovídá signálu srdečního rytmu, je zpočátku nutné, zároveň s nahráváním videa, tento puls snímat nějakým dalším způsobem. K tomuto účelu byl použit jednoduchý PPG snímač *Beurer Pulzní oxymetr PO 80* 1.4 fungující samostatně a zaznamenávající čas, pro pozdější porovnání.

Nabízeli se tři možnosti.

1. Využít integrovaný čip, jako například MAX30102, a připojit ho přímo k Raspberry Pi. Přičemž zůstává otevřená možnost měřit skutečný puls odděleně pomocí mikrokontroleru a stejného čipu.

2. Měřit pomocí přenosného oxymetru, ideálně takového, který umožňuje průběžně ukládat data.

3. Využít dnes již poměrně rozšířených 'chytrých' hodinek, které většinou mají vlastní snímač PPG a zvládnou posílat data o srdečním rytmu do spárovaného telefonu.

Pro demonstrativní účely by bylo nejvhodnější mít snímač přímo připojený k hlavní výpočetní jednotce. Stejný efekt má ale i oxymetr měřící odděleně, což navíc

výrazně zjednodušuje jeho implementaci.

Přijdeme tak sice o možnost zobrazovat živě jak moc se měření s pomocí EVM liší od skutečného údaje, ale na druhou stranu taková možnost není příliš podstatná a jen by komplikovala používání konečného produktu.

4.6 Software

Vývoj nejprve probíhal v programovacím jazyce Matlab. Především proto, že původní kód obstarávající lineární EVM [1] byl napsán v Matlabu, ale také proto, že Matlab umožňuje jednoduchou práci s řešením problémů počítačového vidění, hlavně díky k tomu navrženým toolboxům.

Po ověření realizovatelnosti tohoto projektu se přešlo na vývoj v programovacím jazyce Python 3.9. Ačkoliv totiž Matlab lze použít i na Raspberry Pi 4B, Python nabízí jednak neplacenou platformu a zároveň možnost jednoduchého využití knihoven, které používají nejmodernějších metod pro řešení problémů ve strojovém vidění.

Co se týče Matlabu, byl využit Computer Vision Toolbox a PyrTools. V případě Pythonu jsou použité knihovny: cv2 (od OpenCV), numpy, scipy a mediapipe.

4.6.1 Předspracování v Matlabu

Nejprve bylo nutné odstranit pohyb z analyzovaného videa.

Toho je v Matlabu docíleno za použití Viola-Jones detektoru obličeje, který má Matlab již natrénovaný a stačí zavolat funkci:

```
faceDetector = vision.CascadeObjectDetector()
```

Následně jsou detekovány obličeje pomocí funkce:

```
faceLocation = step(faceDetector, obraz)
```

a oblast s výskytem obličeje je předána funkci

```
featurePoints = detectMinEigenFeatures(obraz,...
                "ROI", faceLocation)
```

Je vytvořen point tracker a všechny tyto proměnné jsou použity pro jeho inicializaci.

```
pointTracker = vision.PointTracker(...
                'MaxBidirectionalError', 2)
initialize(pointTracker, featurePoints, obraz)
```

Takto inicializovaný point tracker pak odhaduje posun obličeje v čase a lze z něj získat geometrickou transformaci jednotlivých bodů. Nás zajímají především body

v oblasti čela, a proto jsou k významným bodům na začátku přidány body čtverce 16x16 pixelů se středem uprostřed čela.

Pozice bodů čtverce jsou zároveň ukládané do textového souboru, aby se vyhnulo opětovnému sbíhání tohoto velmi pomalého algoritmu v případě, že chceme zpracovávání dat nějak upravit.

Nakonec je vyříznuta oblast ohraničená tímto čtvercem a je z ní vytvořena nová sekvence, která putuje na zpracování pomocí EVM kódu od Wu et al.

4.6.2 Předzpracování v Pythonu

Hledání a vyřezávání obličeje trvalo v Matlabu příliš dlouho, nemluvě o ostatních nevýhodách zmíněných na začátku kapitoly.

V pythonu se nejprve přichystá detektor obličeje od MediaPipe a následně se zpracovává video do posledního snímku. Přičemž se hledají obličeje a pomocí OpenCV a pózy bodů na čele se počítá transformace pro body tvořící obdélník. Nakonec je vyříznutá sekvence uložena do nového videa a poslána ke zpracování modulu EVM.py.

Obr. 4.4: Vyříznutá oblast čela, detekovaná pomocí MediaPipe

Obr. 4.5: Flow chart pro detekci obličeje

4.6.3 Implementace EVM v Pythonu

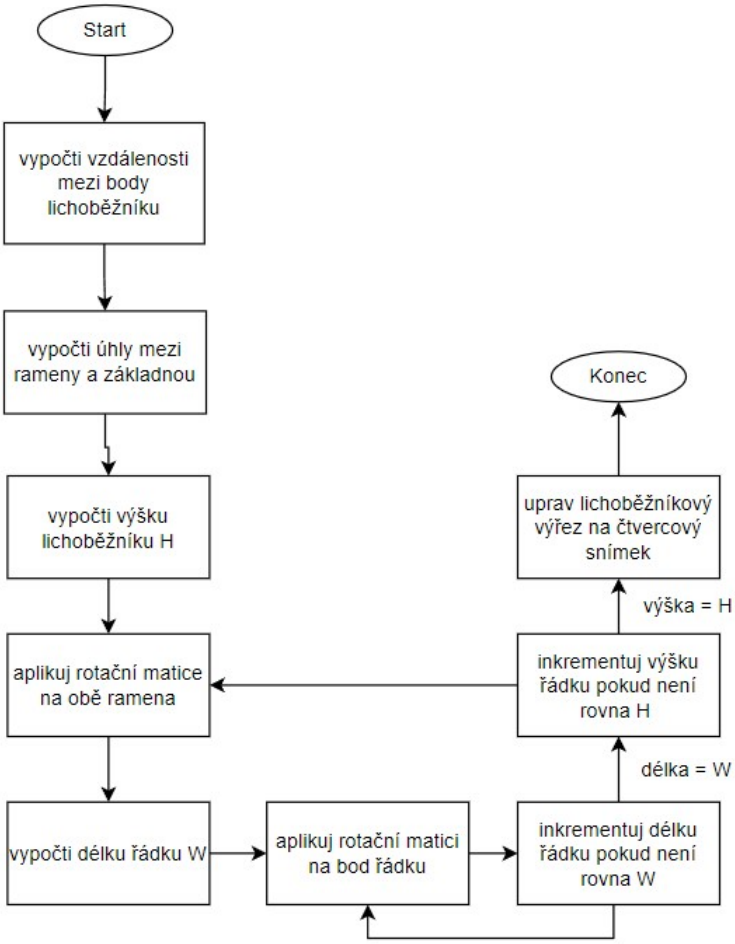
Bylo využito pouze lineární EVM, jelikož jde o nejrychlejší typ tohoto algoritmu a zajímá nás především barva, kterou fázová EVM neovlivňuje, a běh v reálném čase, který fázová EVM omezuje. Principiálně se algoritmus moc neliší od provedení Wu et al., pouze byl přepsán do Pythonu.

Všechny funkce jsou nakonec volány z *main.py*, pro jednoduchost vyhodnocování.

4.7 Měření

Měření probíhalo tak, že se nejprve nahrály videozáznamy za různých podmínek, přičemž zároveň se ukládal údaj o tepové frekvenci pomocí přenosného pulzního oxymetru.

Později se na nahraných videozáznamech provedlo předzpracování a EVM. Na výsledný, barevně zvýrazněný, signál se aplikovala Fourierova transformace a nejvýznamnější frekvenční složka se porovnala se záznamem z pulzního oxymetru, který se získá pomocí programu SpO2 Assistant jak je vidět na obrázku 4.7.



Obr. 4.6: Flow chart pro výřez měřené oblasti

4.7.1 Podmínky scény

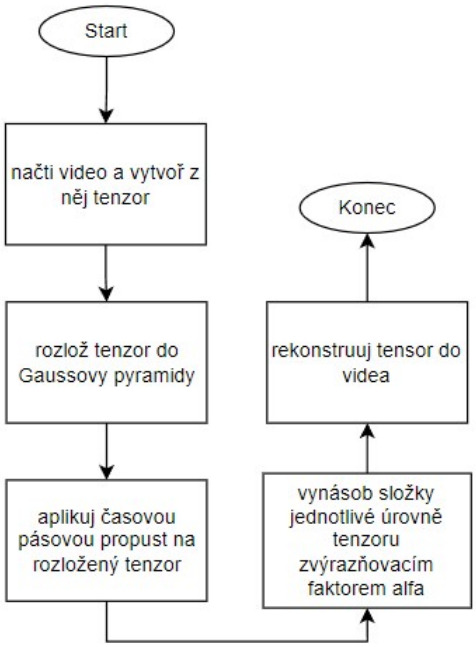
Scéna byla vybrána v laboratoři tak, aby osvětlení přicházelo od stejnosměrného zdroje, jelikož blikání žárovky by se mohlo projevit ve výsledném spektru.

4.7.2 Sběr dat

Většina dat byla nahrávána za kontrolovaných podmínek v laboratoři v různých vzdálenostech a za různého druhu pohybu. Některá data vyžadovala nahrávání v nekontrolovaných venkovních podmínkách a některá takto byla natočena záměrně.

Testovali se konkrétně tyto parametry:

1. Vzdálenost od kamery
2. Úhel natočení obličeje
3. Rozdílné barevné tóny kůže



Obr. 4.7: Flow chart pro EVM

4. Různé osvětlení
5. Různé druhy pohybu
6. Klidový stav

Pro vzdálenost byly zvoleny tři pozice: 30cm od kamery, 1m od kamery a 2m od kamery. Je totiž nutné vědět jak daleko by se monitorovací systém mohl například v nemocnici instalovat.

Úhel natočení byl zkoušen zvlášť pro horizontální a vertikální natočení a jejich obecnou kombinaci, simulujíc natočení pacienta v lůžku.

Barevný tón dobrovolníků se dělí na bílou, světle hnědou a tmavě hnědou. Všichni dobrovolníci byli muži.

Osvětlení je testováno pod bílou LED lampou napájenou stejnosměrným proudem, a na denním světle, pro porovnání využití ve vnitřních a venkovních prostorách.

Pohyb se dělí na pomalý, rychlý a velmi rychlý, při kterém se záměrně snažíme algoritmus rozbít.

A nakonec klidový stav, pro porovnání s EVM bez stabilizace obličeje.

Obr. 4.8: Flow chart pro main.py

4.8 Shrnutí

Bylo řečeno proč se jako hlavní výpočetní jednotka vybralo Raspberry Pi 4 B a jak probíhal vývoj EVM monitorovacího systému nejprve v Matlabu a později v Pythonu. Zmínila se Raspberry Pi HQ, pomocí které je, zároveň s pulzním oxymetrem, měřen srdeční rytmus. Rozebrala se nutnost použít dotykový display a s tím související přenosnost zařízení, která vyžadovala výběr mobilního zdroje. Nakonec byla vysvětlena strategie měření, jaké parametry v ní účinkovali a jak probíhal sběr dat.

Obr. 4.9: Příklad zaznamenané tepové frekvence z SpO2 Assistant

5 Výsledky

V této kapitole se podíváme jak se EVM v kombinaci s detekcí obličeje pomocí MediaPipe FaceMesh osvědčil vůči skutečnému údaji, měřeném pulzním oxymetrem, za různých situací.

Pro každou situaci bylo natočeno několik různých videí, přičemž se ponechávala jedna konstantní proměnná, jako například úhel natočení obličeje. Každá testovaná situace má minimálně jeden záznam v pohybu, jeden záznam v klidu a jeden s obecným natočením obličeje.

Celý dataset byl nasnímám předem pomocí Raspberry Pi 4 B a následně zpracován v laboratorních podmínkách z důvodu pevné kontroly nad celým procesem. Finální algoritmus jde nicméně sbíhat i v reálném čase.

Zpracování probíhalo poloautomaticky. Bylo nutné vkládat do Python skriptu jména souborů, jinak ale vše probíhalo automatizovaně.

Nejprve se podíváme na volbu parametrů a výsledky prvotní verze tohoto systému, který využíval Matlab a Viola-Jones detektor obličejů.

5.1 Volba parametrů algoritmu

Jako nejlepší ukazatel tepové frekvence se po zvýraznění ukázala být zelená složka, u níž se bere frekvence s nejvyšší amplitudou. Jak je vidět na obrázku 5.2, zlepšil se oproti prvotnímu řešení 5.1 i podíl signálu k šumu.

Při zpracování EVM byl signál filtrován pásovou propustí v rozmezí 1 - 2 Hz, což odpovídá rozmezí 60 - 120 BPM. Dále byl použit faktor zvýraznění $\alpha = 80$, jelikož se tato hodnota osvědčila při zvýrazňování barvy už v práci Wu et al. Ve výjimečných případech, za špatného osvětlení, byl použit faktor $\alpha = 150$.

Natáčelo se v 10 sekundových intervalech, což se ukázalo jako dobrá rezerva, jelikož v některých případech trvalo pár sekund vůbec najít obličej. Zároveň jsme se tak vyhnuli chybám měření, které může způsobit přirozená nepravidelnost tepové frekvence a místo toho měříme průměrnou hodnotu za 10 sekund.

5.2 Výsledky měření pomocí Viola-Jones

Vzhledem k výpočetní náročnosti této prvotní metody byl proveden jen malý počet měření pro hrubou představu o přesnosti celého algoritmu v porovnání s pulzním oxymetrem. Časový průběh jednoho takového měření je vidět na obrázku 5.1.

Veškerá data byla nasnímána pod stejnosměrně napájenou LED lampou, z důvodu zamezení znečištění snímaného signálu blikajícím světlem, jaké vykazují klasické žárovky.

Byly snímány různá rozlišení, snímky za sekundu a testovala se také robustnost algoritmu proti pohybu, natočení a vzdálenosti. Většina videí byla natočena s rozlišením 640x480 pixelů při 30 snímcích za sekundu.

Jak se předpokládalo, nebyl žádný rozdíl mezi dvěma často používanými snímacími periodami, tedy mezi 30 a 60 snímky za sekundu, jelikož dalece přesahují potřebnou snímací frekvenci. Algoritmus se ovšem ukázal být slabý v případě vzdáleností nad 50cm, jak je vidět na 6. a 7. řádku tabulky 5.1. Navíc byl konečný algoritmus těžkopádný a jen sotva by jej šlo využít pro běh v reálném čase.

Tab. 5.1: Výsledky pomocí Viola-Jones

Podmínky	Oxymetr [BPM]	EVM [BMP]
1280x720 resolution, v klidu	75.0	75.8
30 fps, v klidu	82.0	82.4
30 fps, v pohybu	81.0	82.7
60 fps, v klidu	80.0	80.7
60 fps, v pohybu	68.0	65.0
Daleko, v klidu	82.0	63.0
Daleko, v pohybu	81.0	65.3
Samotné EVM, v klidu	82	88.7

5.3 Výsledky měření pomocí MediaPipe Face Mesh

Měření pomocí MediaPipe již probíhalo přesně jak je popsané v předchozí kapitole. Bylo k tomu využito tří dobrovolníků. Každý dobrovolník zastával odlišný odstín pleti, pro předejití rasovému biasu. Všichni dobrovolníci byli muži, což bohužel ponechává slepé místo například v situaci, kdy subjektz zakrývají čelo vlasy. Dobrovolníci byli ve věku od 24 do 45 let, což také znamená slepé místo pro případ, kdy by vrásčité čelo mohlo mít efekt na kvalitu měření.

Nicméně cílem projektu bylo vyhodnocení využití EVM ve sledovacích systémech, způsobem navrženým v předchozí kapitole, a zmíněné nedostatky tedy ponecháváme pro potencionální budoucí výzkum, který může být zaměřen na přesnost měření.

Pojďme se tedy podívat na význam, jež měly testované veličiny na funkci systému.

5.3.1 Vzdálenost

Vzdálenost se měřila ve třech pozicích od kamery, a to 0.3, 1 a 2 metrů. Znovu si zopakujme, že každý videozáznam se charakterizuje právě jednou konstantní podmínkou - v tomto případě vzdáleností. Jinak se od sebe liší pohybem, natočením atd.

Důvodem je pokusit se nalézt kombinace podmínek, za kterých náš algoritmus přestane fungovat. Defektní videozáznamy můžeme později detailněji analyzovat.

U příkladu konstantní vzdálenosti se tak ukázala kombinace klidového stavu obličeje a velké vzdálenosti jako fatální, jak je vidět v 5.4, kdy ani nedošlo k detekci obličeje.

Při dvou metrech je detekovaná oblast čela opticky příliš malá, a tak, i když obličej detekujeme, nedaří se nám dosáhnout kvalitních výsledků. Potencionálním řešením tohoto problému je využití zoomu. Tím jsme se ovšem dále nezabývali.

Dalším zajímavým poznatkem byla odchylka čtvrtého řádku tabulky 5.3. Při tomto měření byla snaha o co nejrychlejší pohyb obličeje, a přesto se výsledek zase tolik neliší od skutečnosti.

Tab. 5.2: 30 cm od kamery

Měření	EVM [BPM]	Oxymetr [BMP]
1	86.3	87
2	73.9	74
3	74.0	75
4	80.1	78
5	80.1	84

Tab. 5.3: 1m od kamery

Měření	EVM [BPM]	Oxymetr [BMP]
1	86.3	90
2	73.9	79
3	86.3	87
4	80.1	86
5	80.1	80

Tab. 5.4: 2m od kamery

Měření	EVM [BPM]	Oxymetr [BMP]
1	76.4	88
2	97.6	91
3	-	90
4	80.2	85
5	92.6	80

5.3.2 Úhel natočení tváře

Natočení tváře se měřilo v horizontální, vertikální rovině a obecné náhodné.

Zde narážíme na velké odchylky od hodnot změřených pulzním oxymetrem. Důvodem je použité osvětlení, které svítí ze shora. Otáčením hlavy, tak dochází k prudkému přechodu jasové úrovně na detekované oblasti čela, což za vhodných podmínek, při pohybu ze strany na stranu, vytváří periodický signál, který pak zanáší frekvenční spektrum. To je obzvlášť vidět v tabulce 5.5 v případě druhého měření, kdy druhá největší amplituda patřila právě 79 BPM.

Na druhou stranu v případě 5.6 došlo u čtvrtého měření k maximálnímu záklonu hlavy a dočasné ztráty detekované masky. Což, podobně jako v předchozí podkapitole rychlý pohyb, překvapivě nemělo velký vliv na kvalitu měření.

V tabulce 5.7 bylo v případě čtvrtého měření dosaženo rozbití algoritmu, díky rychlím pohybům v kombinaci s obecným natáčením obličeje.

Tab. 5.5: Horizontální úhel natočení tváře

Měření	EVM [BPM]	Oxymetr [BMP]
1	92.5	91
2	117	78
3	98.6	86
4	92.4	87
5	77.7	77

Tab. 5.6: Vertikální úhel natočení tváře

Měření	EVM [BPM]	Oxymetr [BMP]
1	80.1	80
2	74.0	75
3	80.1	84
4	73.9	75
5	74.0	75

Tab. 5.7: Obecný úhel natočení tváře

Měření	EVM [BPM]	Oxymetr [BMP]
1	74.0	77
2	80.7	81
3	86.3	83
4	95.1	84
5	74.0	82

5.3.3 Barevný tón kůže

Byly testovány tři odstíny kůže a to bílá, světle hnědá a tmavě hnědá. Dobrovolníci byli muži ve věku 24 až 45 let. Můžeme konstatovat, že EVM není zvlášť rasově předpojatý algoritmus.

U měření tmavě hnědého odstínu kůže 5.10 nastala konstantní odchylka nejspíš proto, že snímání probíhalo za nepříznivých podmínek na přímém slunci.

Tab. 5.8: Bílý tón kůže

Měření	EVM [BPM]	Oxymetr [BMP]
1	77.7	78
2	74.0	75
3	80.7	81
4	80.7	82
5	77.7	78

Tab. 5.9: Světle hnědý tón kůže

Měření	EVM [BPM]	Oxymetr [BMP]
1	60.0	62
2	59.3	60
3	63.2	65
4	61.6	63

Tab. 5.10: Tmavě hnědý tón kůže

Měření	EVM [BPM]	Oxymetr [BMP]
1	88.7	88
2	115.7	106
3	102.9	95
4	107.7	98

5.3.4 Osvětlení

Osvětlení bylo testováno u LED lampy napájené stejnosměrným zdrojem a u slunce. Jak je vidět, intenzivní sluneční svit má na robustnost algoritmu fatální následky (řádky 3 až 5 v 5.12). To je nejspíš způsobené ostrými stíny, jelikož první dvě měření byla kvalitní a provedena, když bylo zamračeno.

Tab. 5.11: Stejnosměrné LED osvětlení

Měření	EVM [BPM]	Oxymetr [BMP]
1	86.3	87
2	80.1	79
3	86.3	86
4	92.4	89
5	80.1	82

Tab. 5.12: Sluneční světlo

Měření	EVM [BPM]	Oxymetr [BMP]
1	86.3	88
2	105.1	106
3	80.1	89
4	76.1	88
5	98.0	89

5.3.5 Pohyb

Byl zkoumán pomalý, rychlý a velmi rychlý pohyb. Náš algoritmus obstál překvapivě dobře a začal se rozbíjet až při velmi rychlém pohybu 5.15.

Tab. 5.13: Pomalý pohyb

Měření	EVM [BPM]	Oxymetr [BMP]
1	94.0	95
2	86.3	86
3	86.3	86
4	74.0	76
5	75.0	75

Tab. 5.14: Rychlý pohyb

Měření	EVM [BPM]	Oxymetr [BMP]
1	88.7	86
2	88.1	88
3	78.3	88
4	86.3	86
5	93.8	90

Tab. 5.15: Velmi rychlý pohyb

Měření	EVM [BPM]	Oxymetr [BMP]
1	88.0	90
2	105.1	102
3	72.0	105
4	94.7	101
5	80.1	98

5.3.6 Klid

Nakonec pro porovnání s původní EVM byl naměřen záznam v co největším klidu. Nepřekvapivě se dosáhlo minimálně stejné přesnosti jako při využití jen EVM. Přičemž samotný algoritmus byl, díky menšímu počtu zpracovávaných pixelů, mnohem rychlejší.

Tab. 5.16: Tepová frekvence v klidu

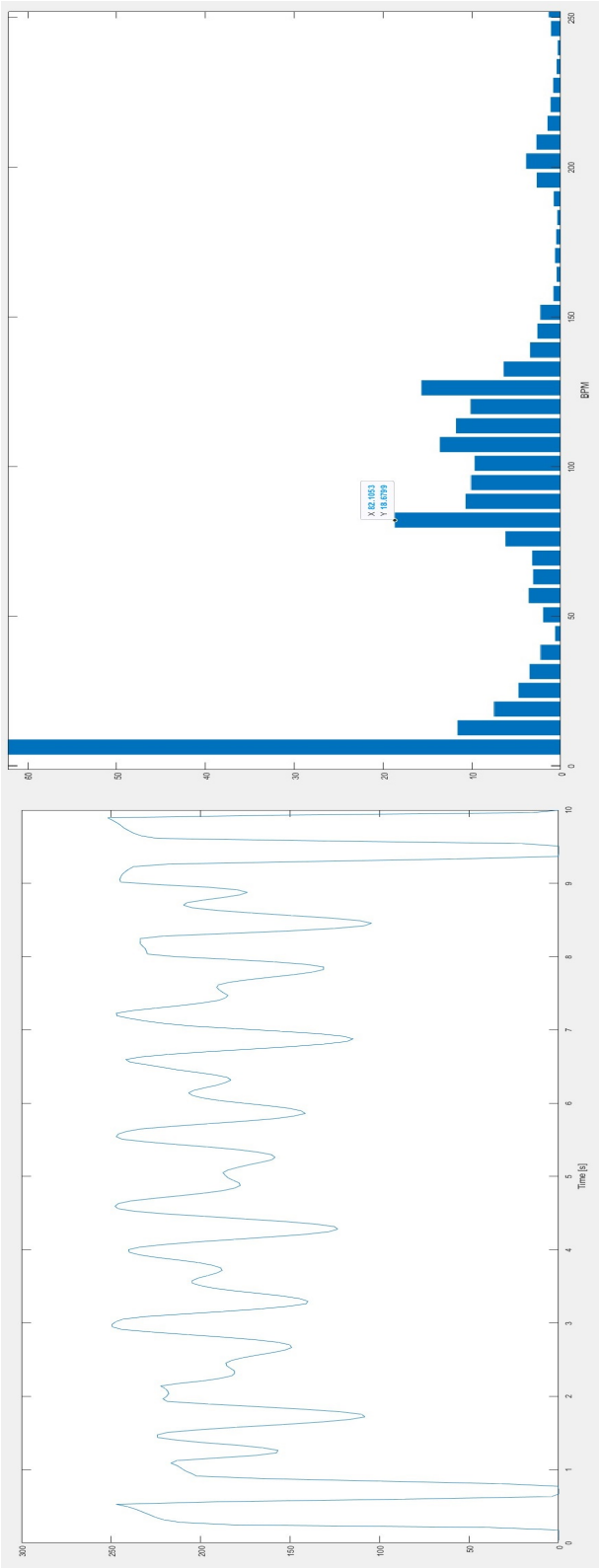
Měření	EVM [BPM]	Oxymetr [BMP]
1	74.0	74
2	74.0	74
3	77.7	79
4	74.0	77
5	86.3	86
6	74.0	75
7	83.1	83
8	83.1	82
9	77.7	79
10	80.1	82

5.4 Shrnutí

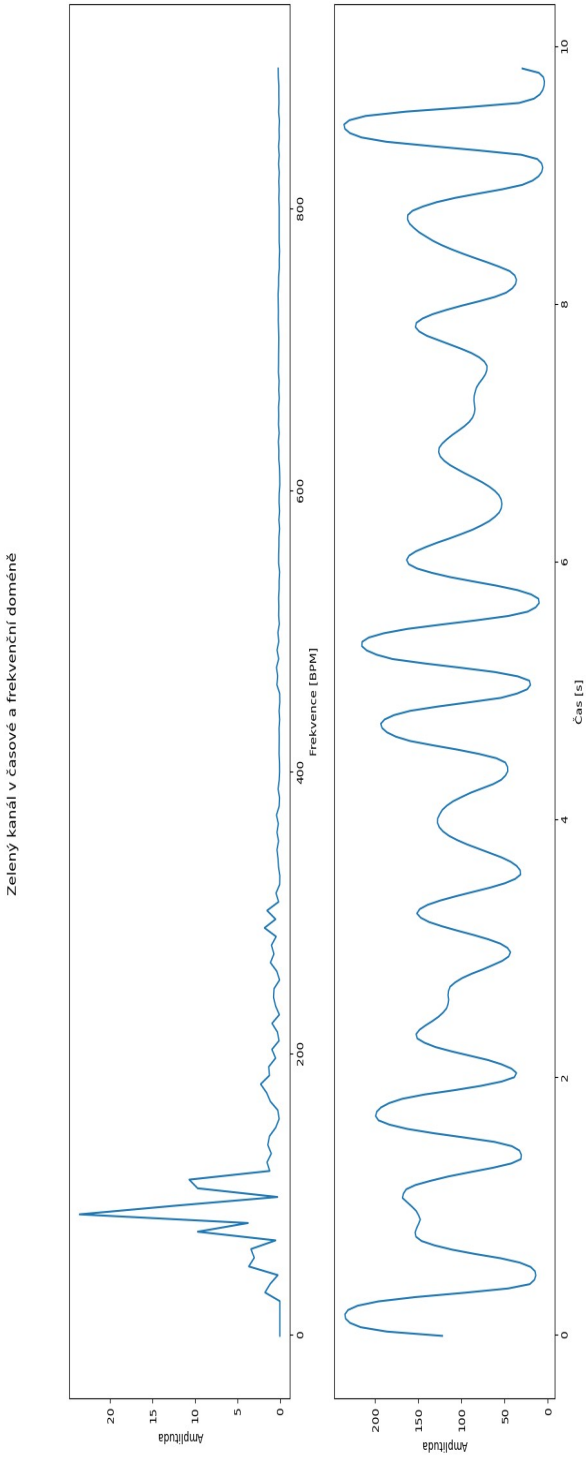
Jak vyšlo najevo, je EVM, i za použití sledování obličeje subjektu, stále náchylný na velké vzdálenosti (více než 2m), při nichž je snímaná plocha příliš malá na to, aby poskytovala jakoukoliv robustnost. To je potencionálně možné vyřešit kamerou z větším rozlišením a softwarovým zoomem. Anebo snímáním celého obličeje.

Dalším vážným nedostatkem je náchylnost na směr světelného zdroje a celkově rychlé jasové změny, například přechod ze stínu do světla. Toto je natolik závažný problém, že by dokonce bylo možné naměřit tepovou frekvenci i u mrtvého! Pokud by na něj svítilo třeba špatně umístěné světlo za větrákem s přesně správnou úhlovou rychlostí. Vzhledem k povaze takového signálu, by ale neměl být problém tuto situaci softwarově ošetřit, jelikož stíny mají většinou poměrně ostrou hranu, kdežto barva čela se mění prakticky najednou, díky rychlosti pulzní vlny.

Dále se v této kapitole probíralo měření různých odstínů kůže, efekt různého osvětlení (kde sluneční světlo dopadlo poměrně špatně, právě díky ostrým stínům), rychlý pohyb, který příjemně překvapil tím, jak náročné bylo algoritmus rozbít. A nakonec se pro porovnání s klasickou EVM provedlo měření v klidu.



Obr. 5.1: Raná verze z Matlabu. Vlevo: Časový průběh amplitudy zeleného kanálu průměrovaných pixelů z oblasti čela po dobu 10s Napravo: Fourierova transformace tohoto signálu s druhou nejvyšší harmonickou značící změřenou tepovou frekvenci (první harmonická je DC složka). [23]



Obr. 5.2: Časový průběh jasu zeleného kanálu pro první měření ve vzdálenosti 30cm.

Závěr

Od publikace průlomové studie z MIT v roce 2012, pokročilo bezkontaktní měření značně kupředu. Od experimentální biologie, přes detekci vad mostů, až po monitorování dýchání spícího dítěte, eulerovská video-magnifikace najde mnohá praktická využití.

Díky pokroku v oblasti fázové EVM bylo možné potlačit vliv šumu a zvýšit tak maximální možné zesílení, ale také utlumit nežádoucí pohyby. Využitím Rieszovy pyramidy pak lze počítat EVM v reálném čase a otevřít tak využití tohoto algoritmu v monitorovacích kamerových systémech; například pro prevenci patologických onemocnění, automatické zavolání zdravotníků nebo monitorování životních funkcí u novorozenců.

Nadále existují problémy s výraznými pohyby v analyzovaném obraze. Cílem této práce tedy bylo vyzkoušet EVM v kombinaci s algoritmy pro detekci obličeje, a možnosti jejich využití v reálném monitorovacím systému. K tomuto účelu byla na platformě Raspberry Pi 4 B realizována EVM v programovacím jazyku Python 3.9, měřící pulz podle zvýraznění změn barvy kůže, a Face Mesh detektor od MediaPipe, zajišťující stabilizaci obličeje. Takto provedené řešení bylo porovnáno s daty získanými z pulzního oxymetru a bylo vyhodnoceno jeho využití v monitorovacích systémech.

Z výsledků měření vyšlo najevo, že EVM bude muset ujít ještě kus cesty, než se stane spolehlivým pro monitorování životních funkcí. Výsledné řešené si obstojně poradilo s problémem velkého pohybu v obraze, a díky MediaPipe se zachovala možnost sbíhání v reálném čase, ale nadále zůstává problém s měnícím se osvětlením, který je o to horším, jelikož může vést k pozitivně negativnímu měření. Budoucí práce se touto problematikou mohou zabývat. Další oblastí, která si zaslouží delší výzkum je porovnání přesnosti EVM pomocí propojeného měřiče TF.

Mimo to, také přetrvávají těžko řešitelné problémy jako například, když není možné jakkoliv detekovat kůži nebo změnu tvaru.

Výsledek je nicméně uspokojivý, už jen proto, že se celkový algoritmus ukázal být překvapivě robustní i vůči záklonům hlavy a rychlým pohybům obličeje. Rozhodně se nejedná o slepou cestu a doufáme, že o EVM ještě v budoucnu uslyšíme.

Literatura

[1] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Frédo Durand, and William Freeman. 2012. Eulerian video magnification for revealing subtle changes in the world. ACM Trans. Graph. 31, 4, Article 65 (July 2012), 8 pages. DOI:https://doi.org/10.1145/2185520.2185561

[2] Wadhwa, Neal, et al. "Phase-based video motion processing."ACM Transactions on Graphics (TOG) 32.4 (2013): 1-10.

[3] Arango, Carlos, et al. "Subtle motion analysis and spotting using the Riesz Pyramid."VISIGRAPP 2018. 2018.

[4] Fontanari, Thomas V., and Manuel M. Oliveira. "Simultaneous magnification of subtle motions and color variations in videos using Riesz pyramids."Computers & Graphics 101 (2021): 35-45.

[5] BERNÁTEK, Pavel. Vizualizace pulzu ve videozáznamu obličeje. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií, 2016, 58 listů : grafy, fotografie + 1 CD-ROM.

[6] ZHANG, Jiahui, Ke ZHANG, Xinhao YANG a Chenrui WEN. Heart rate measurement based on video acceleration magnification. In: 2020 Chinese Control And Decision Conference (CCDC) [online]. IEEE, 2020, s. 1179-1182 [cit. 2021-12-29]. Dostupné z: doi:10.1109/CCDC49329.2020.9164451

[7] ABNOUSI, Freddy, Guson KANG, John GIACOMINI, et al. A novel noninvasive method for remote heart failure monitoring: the EuleriAn video Magnification apPLications In heart Failure studY (AMPLIFY). NPJ digital medicine [online]. BERLIN: NATURE RESEARCH, 2019, 2(1), 80-80 [cit. 2021-12-29]. ISSN 2398-6352. Dostupné z: doi:10.1038/s41746-019-0159-0

[8] LE NGO, Anh Cat, Yee-hui OH, Raphael C.-W PHAN a John SEE. Eulerian emotion magnification for subtle expression recognition. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) [online]. IEEE, 2016, s. 1243-1247 [cit. 2021-12-29]. ISBN 1479999881. ISSN 1520-6149. Dostupné z: doi:10.1109/ICASSP.2016.7471875

[9] ZHENG, Jiayan, Qiumo LI, Ruoyu MAO, Yuzhou WANG a Zhixiang ZHOU. Bridge Vibration Virtual Sensor Based on Eulerian Video Magnification and Gray Mean Difference. IOP conference series. Materials Science and Engineering [online]. Bristol: IOP Publishing, 2020, 964(1), 12026 [cit. 2021-12-29]. ISSN 1757-8981. Dostupné z: doi:10.1088/1757-899X/964/1/012026

[10] ALARIFI, Abdulaziz, Amr TOLBA a Azza S HASSANEIN. Visualization process assisted by the Eulerian video magnification algorithm for a heart rate monitoring system: mobile applications. Multimedia tools and applications [online]. New York: Springer US, 2018, 79(7-8), 5149-5160 [cit. 2021-12-29]. ISSN 1380-7501. Dostupné z: doi:10.1007/s11042-018-6313-x

[11] RAMYA MARIE, S a J ANUDEV. Response Analysis of Eulerian Video Magnification. Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB). Cham: Springer International Publishing, 2019, s. 671-678. ISBN 9783030006648. ISSN 2212-9391. Dostupné z: doi:10.1007/978-3-030-00665-5_66

[12] LAURIDSEN, Henrik, Selina GONZALES, Daniela HEDWIG, et al. Extracting physiological information in experimental biology via Eulerian video magnification. BMC biology [online]. LONDON: BMC, 2019, 17(1), 103-103 [cit. 2021-12-29]. ISSN 1741-7007. Dostupné z: doi:10.1186/s12915-019-0716-7

[13] SHAHADI, Haider Ismael, Zaid Jabbar AL-ALLAQ a Hayder Jawad ALBATTAT. Developed approach for phase-based eulerian video magnification. Telkomnika [online]. 2020, 18(5), 2391-2400 [cit. 2021-12-29]. ISSN 1693-6930. Dostupné z: doi:10.12928/TELKOMNIKA.V18I5.14321

[14] HE, Xiaochuan, Rafik A GOUBRAN a Xiaoping P LIU. Wrist pulse measurement and analysis using Eulerian video magnification. In: 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI) [online]. IEEE, 2016, s. 41-44 [cit. 2021-12-29]. ISBN 1509024557. Dostupné z: doi:10.1109/BHI.2016.7455830

[15] ZAREZADEH, Elham a Mehdi REZAEIAN. Micro Expression Recognition Using the Eulerian Video Magnification Method. Brain. Broad research in artificial intelligence and neuroscience [online]. EduSoft publishing, 2016, 7(3), 43-54 [cit. 2021-12-29]. ISSN 2068-0473.

[16] FONTANARI, Thomas V a Manuel M OLIVEIRA. Simultaneous magnification of subtle motions and color variations in videos using Riesz pyramids. Computers & graphics [online]. Elsevier, 2021, 101, 35-45 [cit. 2021-12-29]. ISSN 0097-8493. Dostupné z: doi:10.1016/j.cag.2021.08.015

[17] MAURYA, Lalit, Pavleen KAUR, Deepak CHAWLA a Prasant MAHAPATRA. Non-contact breathing rate monitoring in newborns: A review. Computers in biology and medicine [online]. United States: Elsevier, 2021, 132, 104321-104321 [cit. 2021-12-29]. ISSN 0010-4825. Dostupné z: doi:10.1016/j.compbiomed.2021.104321

[18] QIU, Qiwen. Automated defect detection in FRP-bonded structures by Eulerian video magnification and adaptive background mixture model. Automation in construction [online]. Elsevier B.V, 2020, 116, 103244 [cit. 2021-12-31]. ISSN 0926-5805. Dostupné z: doi:10.1016/j.autcon.2020.103244

[19] Rahmad, C., et al. "Comparison of Viola-Jones Haar Cascade classifier and histogram of oriented gradients (HOG) for face detection."IOP conference series: materials science and engineering. Vol. 732. No. 1. IOP Publishing, 2020.

[20] Satya Mallick - "Histogram of Oriented Gradients explained using OpenCV"[cit. 2.1.2022] https://learnopencv.com/histogram-of-oriented-gradients/

[21] SMITH, Ian; MACKAY, John; FAHRID, Nahla; KRUCHECK, Don. Respiratory rate measurement: a comparison of methods. British Journal of Healthcare Assistants. 2011, roč. 5, č. 1, s. 18–23.

[22] DÍTĚ, Martin. Multifunkční monitorovací pás. Brno, 2019, 88 s. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce: Ing. Vratislav Harabiš, Ph.D.

[23] Eulerian Video Magnification in Surveillance Systems - Ricanek, Honec - EEICT 2022

[24] ŘIČÁNEK, Dominik. Vícebodová identifikace pozice a orientace objektu. Brno, 2020, 56 s. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav automatizace a měřicí techniky. Vedoucí práce: Ing. František Burian, Ph.D.

[25] MediaPipe Face Mesh dokumentace [cit.18.5.2022] Dostupné z: https://google.github.io/mediapipe/solutions/face_mesh.html

[26] Bazarevsky, Valentin, et al. "Blazeface: Sub-millisecond neural face detection on mobile gpus."arXiv preprint arXiv:1907.05047 (2019).

[27] Systola [cit.18.5.2022] Dostupné z: https://cs.wikipedia.org/wiki/Systola

[28] Elektrokardiogram [cit.18.5.2022] Dostupné z: https://cs.wikipedia.org/wiki/Elektrokardiogram

[29] Pulzní oxymetrie [cit.18.5.2022] Dostupné z: https://www.wikiskripta.eu/w/Pulzní_oxymetrie

[30] Respirační sinusová arytmie [cit.18.5.2022] Dostupné z: https://cs.wikipedia.org/wiki/Respirační_sinusová_arytmie

[31] Viola-Jones detektor [cit.18.5.2022] Dostupné z: https://en.wikipedia.org/wiki/Viola-Jones_object_detection_framework

[32] Integrální obraz [cit.18.5.2022] Dostupné z: https://cs.wikipedia.org/wiki/Integrální_obraz

[33] Steve Mould - "Reveal Invisible Motion With This Clever Video Trick"[cit. 28.12.2021] Dostupné z: https://www.youtube.com/watch?v=rEoc0YoALt0

Seznam příloh

6 Obsah elektronické přílohy

```
/DP-2022 ..... kořenový adresář přiloženého CD
├── git_repo ..... celý git repozitář
│   ├── /Code ..... zdrojové soubory
│   │   ├── /Matlab
│   │   │   ├── cut_out_face.m
│   │   │   ├── extract_pulse.m
│   │   │   ├── main.m
│   │   │   └── show.m
│   │   ├── camera.py ..... skript pro nasnímání videa
│   │   ├── EVM.py
│   │   ├── face_tracking.py
│   │   └── main.py ..... hlavní skript
│   └── 60fps_lowres_HR.mp4 ..... zkušební video
└── text ..... textové soubory
    ├── DP.pdf
    └── DP.zip ..... LaTeX zdrojový kód
```