

MASARYKOVA UNIVERZITA
FAKULTA INFORMATIKY



Optické měření srdečního pulsu z krátkého kamerového záznamu

BAKALÁŘSKÁ PRÁCE

Michal Šindelář

Brno, jaro 2016

MASARYKOVA UNIVERZITA
FAKULTA INFORMATIKY



Optické měření srdečního pulsu z krátkého kamerového záznamu

BAKALÁŘSKÁ PRÁCE

Michal Šindelář

Brno, jaro 2016

Místo tohoto listu vložte kopie oficiálního podepsaného zadání práce a prohlášení autora školního díla.

Prohlášení

Prohlašuji, že tato bakalářská práce je mým původním autorským dílem, které jsem vypracoval samostatně. Všechny zdroje, prameny a literaturu, které jsem při vypracování používal nebo z nich čerpal, v práci řádně cituji s uvedením úplného odkazu na příslušný zdroj.

Michal Šindelář

Vedoucí práce: RNDr. Vladimír Ulman Ph.D.

Poděkování

Rád bych především poděkoval RNDr. Vladimíru Ulmanovi Ph.D. za vedení mé bakalářské práce, odborné rady a ochotu.

Dále děkuji rodině, kamarádům a všem, kteří mě při studiu a tvorbě práce podporovali. Především potom rodičům, kteří navíc souhlasili se zveřejněním fotografií demonstrující měření využitá v práci.

Shrnutí

Cílem této práce je vytvořit aplikaci, která na základě krátkého optického záznamu obličeje určí tepovou frekvenci snímané osoby a následně těp poutavou formou zobrazí uživateli. Pro realizaci práce je nutné vytvořit samotný algoritmus pro určení tepu. Zobrazení tepové frekvence je založené na *Eulerovském zvýraznění časových frekvencí* ve videu.

Práce nejdříve zkoumá fyziologické důsledky proudění krve v po-kožce, pomocí kterých by bylo možné tepovou frekvenci určit. Po srovnání možností je pro určení tepu vybrána analýza fluktuace barevného odstínu kůže, která je způsobena prouděním krve. Posléze je teoreticky popsáno, jak je možné ze zaznamenaných změn barevného odstínu v závislosti na čase získat výsledný tep. Postup využívá časovou frekvenční doménu vytvořenou pomocí Fourierovy transformace. Výsledný tep je určen hledáním nejsilnější přítomné frekvence v přípustném intervalu získaném ze statistických údajů o lidském tepu.

V praktické části je popsána implementovaná aplikace v prostředí *OpenCV* a demonstrována její funkčnost na vstupním záznamu. V této části jsou diskutovány požadavky na vstupní záznam. Ten musí mít minimální délku 500 snímků při vzorkovací frekvenci nejméně dvakrát vyšší než skutečný tep. Důležité je zaručení stabilní polohy obličeje. Na závěr jsou popsána měření tepu ze záznamů splňujících vstupní požadavky. Prezentovaný postup vykazuje maximální chybu okolo 5% tepů/min.

Klíčová slova

Digitální zpracování obrazu, určení tepu z optického záznamu, Eulerovské zesílení, fotopletysmografie

Obsah

1	Úvod	1
1.1	<i>Digitální zpracování obrazu</i>	1
1.2	<i>Tepová frekvence</i>	1
1.3	<i>Cíle aplikace</i>	2
2	Teorie	5
2.1	<i>Diskrétní obraz</i>	5
2.1.1	<i>Diskrétní šedotónový obraz</i>	5
2.1.2	<i>Diskrétní barevný obraz (se třemi kanály)</i>	5
2.1.3	<i>Obrazový záznam</i>	5
2.2	<i>Fourierova transformace</i>	6
3	Aparáty pro realizaci práce	7
3.1	<i>Určení tepu</i>	7
3.2	<i>Vizualizace tepu</i>	9
4	Použité vývojové a výzkumné nástroje	11
4.1	<i>Použitý hardware</i>	11
5	Postup	13
5.1	<i>Objektivní pravda</i>	13
5.2	<i>Postup určení tepu</i>	13
6	Implementace	21
6.1	<i>Datové struktury</i>	21
6.2	<i>Popis aplikace</i>	21
6.2.1	<i>Hlavní cyklus</i>	21
6.2.2	<i>Analýza video záznamů</i>	24
6.2.3	<i>Určení tepu</i>	25
6.2.4	<i>Vizualizace tepu</i>	26
6.2.5	<i>Rozdíl mezi režimy</i>	28
6.3	<i>Měření a požadavky na vstupní záznam</i>	29
7	Závěr	35
	Rejstřík	37

Bibliografie	37
A Dodatek	39

Seznam tabulek

- 6.1 Výsledky ovlivnění výsledků při určení masky pro výpočet intenzit snímků. Princip byl popsán v Kapitole 6.2.1. Hodnoty ilustrují situace, kdy je pro výpočet intenzit využit nejdříve celý snímek, potom oblast detekovaného obličeje a nakonec čela. 31
- 6.2 Výsledky testování rozdílných charakteristik intenzit snímků popsaných v Kapitole 5.2 pro určení tepu. Ve sloupci *Výsl. tep* je vždy vypočítaný tep aplikací pomocí příslušného přístupu v hlavičce. Ve sloupci *Chyba* je uvedena chyba měření v případě, že výsledný tep spadá mimo povolený interval. Při tomto měření ještě nebyl kladen důraz na výkon aplikace, takže nebylo nutné použít co nejmenší počet snímků pro určení. 32
- 6.3 Výsledky měření pro určení minimálního počtu snímků pro určení tepu. Pro každý záznam je několikrát určen tep pokaždé s jiným počtem snímků. Při použití více než 500 snímků už nebylo naměřeno zpřesnění výsledků. 33

Seznam obrázků

- 1.1 Ilustrace ideálního elektrokardiogramu. 2
- 3.1 Příklad vlastních vektorů při detekci tepu pomocí balistokardiografických pohybů obličeje. Každá bílá šipka odpovídá amplitudě a směru kontribuce příslušného registrovaného bodu k vlastnímu vektoru. Převzato z [5]. 8
- 3.2 Zvýrazněné barevné změny způsobené fluktuací krve. V horní řadě jsou snímky videa v daných časových okamžicích. Ve spodní řadě je potom ilustrace tepu. Tepové údery by mely korespondovat se snímky, ve kterých je pokožka odlišně zabarvená. Tyto barevné změny jsou okem nepostřehnutelné. 9
- 5.1 Záznam čela a manuálně určená maska pro zpracování. 14
- 5.2 Originální hodnoty mediánů intenzit zeleného kanálu v jednotlivých snímcích záznamu. 15
- 5.3 Frekvenční doména naměřených intenzit spolu s maskou filtru pásmová propust. 16
- 5.4 Hodnoty mediánů intenzit zeleného kanálu v jednotlivých snímcích záznamu po filtrace. 17
- 5.5 Frekvenční doména naměřených intenzit po filtrace. 17
- 5.6 Příklady různorodých záznamů pro ověření detekčních schopností postupu. 18
- 5.7 Simulace hodnot mediánů intenzit zeleného kanálu v jednotlivých snímcích záznamu po filtrace. 19
- 5.8 Frekvenční doména simulace zrychlujícího se pulzu s detailem zaměřeným na relevantní frekvence. 20
- 6.1 Uživatelské rozhraní pro volbu režimu aplikace. 22
- 6.2 Vhodná místa pro analýzu globálních světelných změn. 25
- 6.3 Ilustrace principu vytvoření jednoho snímku časové řady z detekovaného obličeje. 27

- 6.4 Vlevo je zobrazen originální snímek z časové řady. Vpravo je součet originálního snímku a snímku po filtrování a zvýraznění frekvencí. Proměnná K odpovídá počtu snímků v původní frontě. 28
- 6.5 Uživatelské rozhraní po dokončení výpočtu tepu i jeho vizualizace. Vlevo zobrazen aktuální vstupní snímek se zobrazením detekovaného obličeje a čela. Vpravo je zobrazen výsledný snímek s vizualizací tepu a s hodnotou vypočítaného tepu. 29
- 6.6 Na levém snímku je zobrazen snímek z původního videa. Pro testování je uměle přidáván Gaussovský šum a tep počítán znovu. Na každém snímku je zobrazena standardní odchylka šumu σ a výsledný tep bpm . I z velmi silně zašumělého videa, ilustrovaného pravým snímkem, se podařilo vypočítat tep spadající do korektního intervalu. 30

1 Úvod

1.1 Digitální zpracování obrazu

Digitální zpracování obrazu je vědní obor informatiky, který nabízí techniky pro zkoumání digitálního obrazu za účelem zisku informací, či vyprodukovaní *vylepšeného* obrazu. Typickým odvětvím, ve kterém nachází digitální zpracování obrazu uplatnění, je medicína.

V medicíně se využívá např. pro zobrazení výsledků rentgenu, magnetické rezonance a dalších. Tato práce se bude zabývat dalším potenciálním využitím digitálního zpracování obrazu konkrétně pro zkoumání tepové frekvence lidského těla z optického záznamu.

1.2 Tepová frekvence

Každý člověk má tep, který je vyvolán vypuzením krve ze srdeční komory do aorty a následně dále do těla. Tep odpovídá počtu úderů srdce za minutu. Jeho hodnota se liší člověk od člověka.

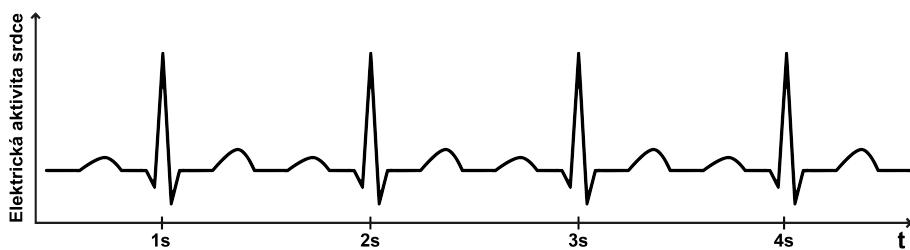
Klidová tepová frekvence je jedním ze základních rutin měření v medicínské péči. Tato frekvence odpovídá situaci, kdy srdce vypuzuje minimální množství krve. Pro správné měření klidové tepové frekvence musí být člověk zdravý a ve fyzickém i psychickém klidu.

Pro určení pracovního intervalu tepu, který má smysl zkoumat, práce využívá *Národní zdravotní statistické záznamy* [1] měřené roku 2011 v USA. Klidový tep ovlivňuje více faktorů jako fyzická kondice, zdraví a především věk. Pro novorozence do jednoho roku je průměrná frekvence 129 tepů/min zatímco pro dospělé už jen 72 tepů/min.

Srdeční frekvence nižší než 60 tepů se považuje za abnormálně pomalou a označuje se jako bradykardie. Hraniční hodnota se ovšem vždy vztahuje i k věkové a výkonnostní kategorii, např. u sportovců je běžné mít sníženou klidovou frekvenci i pod 60. Opakem bradykardie je tachykardie¹, která je u dospělých obvykle diagnostikována nad 100 tepů/min.

1. Blíže viz <http://www.heart.org/HEARTORG/Conditions/Arrhythmia/AboutArrhythmia/Tachycardia-Fast-Heart-Rate_UCM_302018_Article.jsp>

1. Úvod



Obrázek 1.1: Ilustrace ideálního elektrokardiogramu.

Práce vylučuje abnormální případy a zaměří se na interval 55-180 tepů/min, kterým bude pokryto vše kromě extrémně nízkých tepů způsobených bradykardií, či profesionálním tréninkem vrcholových sportovců. Pracovní interval bude stěžejní při odhalování výsledné frekvence z videa. Rozšíření intervalu o ještě nižší frekvenci je problematické a většinou vede k nekorektním výsledkům.

Srdeční tep můžeme znázornit pomocí přístroje EKG. EKG je zkratka pro vyšetřovací metodu zvanou elektrokardiografie, kdy přístroj snímá elektrické aktivity srdce. Zanáší je do grafu zvaného elektrokardiogram (viz Obrázek 1.1) a umožňuje hodnocení aktivity srdce². V elektrokardiogramu odpovídají intervaly mezi jednotlivými lokálními maximy tepové frekvenci srdce.

1.3 Cíle aplikace

Cílem této aplikace je využití fyziologických vlastností proudění krve v těle za účelem určení tepové frekvence pouze ze záznamu obličeje. Po určení bude tep vhodným způsobem vyobrazen uživateli.

Aplikace bude pracovat buď s vloženým záznamem, nebo pro záznam bude využívat přímo webkameru zařízení, na kterém poběží. Při snímání pomocí levnějších (běžných) webkamer se dají očekávat problémy s proměnlivou vzorkovací frekvencí a také s automatickou změnou expozice.

Aplikace bude mít celkem 3 různé režimy:

2. Blíže viz <<http://emedicine.medscape.com/article/1894014-overview>>

1. *Statický* režim při dodaném záznamu

V tomto režimu uživatel dodá aplikaci hotový záznam. Ta bude zpracovávat záznam celý a výpočet skončí až po zpracování celého videa. Tento režim by měl (v závislosti na délce videa) vést k nejlepším výsledkům, zároveň je ale časově nejnáročnější a prodeleva bez interakce s uživatelem je v celé délce doby zpracování. Dále v práci bude diskutován dostatečný počet zpracovávaných snímků pro stabilní detekci tepu a jeho zobrazení. Potom bude možné zpracování videa rozdělit do více vláken pro urychlení výpočtu.

2. *Reálný*³ režim webkamera

V tomto režimu bude aplikace snímat obraz webkamerou. Snímaný obraz bude aplikace přímo zobrazovat a po výpočtu bude vhodně zobrazovat originální a zpracované video se zvýrazněnou fluktuací krve vedle sebe.

3. *Reálný* režim při dodaném záznamu

Tento režim bude velmi podobný předchozímu režimu se snímáním pomocí webkamery. Je vytvořen především kvůli jednoduššímu testování a pro simulaci chování při zpracování online záznamu. Umožní se oprostit od případných problémů se snímáním webkamerou.

Všechny režimy budou mít možnost uložení výstupu. V případě statického režimu to bude uložení celého zpracovaného záznamu. Ve zbývajících to potom bude ukládání výstupu, dokud uživatel program neukončí.

3. Slovem *reálný* je myšlena situace, kdy všechny výpočty probíhají paralelně se zobrazeným výstupem uživateli.

2 Teorie

Protože v práci budou využívány pojmy v kontextu digitálního obrazu, jsou v dalším textu uvedeny potřebné definice.

2.1 Diskrétní obraz

2.1.1 Diskrétní šedotónový obraz

Práce bude pracovat s definicí šedotónového obrazu jako intenzitní funkce E , která pro každý obrazový pixel $p = (x, y)$ přiřazuje příslušnou skalární hodnotu $E(x, y)$, tedy

$$E(p) = E(x, y), \quad (2.1)$$

kde x, y označíme jako celočíselné souřadnice obrazu na pravoúhlé mřížce, pro které platí $0 \leq x \leq M - 1$, $0 \leq y \leq N - 1$ s tím, že M, N odpovídá rozdílu velikosti obrazu. Šedotónové intenzity $E(p)$ nabývají hodnot $0 \leq E(p) \leq G_{max}$. Práce pracuje se standardním 8-bitovým obrazem, tedy $G_{max} = 2^8 - 1 = 255$.

2.1.2 Diskrétní barevný obraz (se třemi kanály)

Práce bude dále pracovat s definicí barevného obrazu se třemi kanály jako C .

$$C(x, y) = (E_1(x, y), E_2(x, y), E_3(x, y))^T, \quad (2.2)$$

kde souřadnice x, y jsou definovány stejně jako v případě šedotónového obrazu. Každý vektor $(E_1, E_2, E_3)^T$, kde $0 \leq (E_1, E_2, E_3) \leq G_{max}$ odpovídá konkrétní barvě v barevném prostoru. V práci práci bude nejvíce použit barevný prostor RGB , kde můžeme pro lepší orientaci označit $(E_1, E_2, E_3) = (r, g, b)$. Minoritně bude použit barevný prostor $NTSC$, kde označení zůstane původní.

2.1.3 Obrazový záznam

Za záznam (video) se dále považuje vektor V diskrétních barevných obrazů v barevném prostoru RGB . Potom i -tý snímek odpovídá:

2. TEORIE

$$C_i = V(i), \quad (2.3)$$

kde $i, 0 \leq i \leq K$, je index snímku ve vektoru V , K délka vektoru V . Dále fps je vzorkovací frekvence záznamu udávána v jednotce Hz.

2.2 Fourierova transformace

Pro analýzu frekvence bude využita jednodimenzionální diskrétní Fourierova transformaci. Ta je definovaná jako:

$$\Phi(\omega_x) = \sum_{x=0}^{N-1} f(x) \exp\left(-i\left(\frac{2\pi}{N}\omega_x x\right)\right), \quad (2.4)$$

kde se proměnná ω_x nazývá *prostorová / časová frekvence*, x a ω_x stejně jako hodnoty funkce $f(x)$ jsou reálná čísla, $i = \sqrt{-1}$ je komplexní jednotka, hodnoty funkce $\Phi(\omega_x)$ jsou komplexní čísla, hodnota funkce $\Phi(\omega_x)$ pro dané ω_x je závislá na všech hodnotách $f(x)$ pro všechna x .

3 Aparáty pro realizaci práce

Lidské vidění má omezenou citlivost na prostoro-časové změny [2]. Mnoho změn, které spadají mimo rozeznávací schopnosti, nesou informace, které se dají pomocí digitálního zpracování obrazu využívat pro různé účely. Práce bude využívat změny, které jsou způsobené tepem (resp. prouděním krve). Pro lidské oko jsou nepostřehnutelné, ale pomocí záznamu kvalitní optickou soustavou a následným digitálním zpracováním obrazu se dají odhalit.

3.1 Určení tepu

Pro potřeby zkoumání pomocí optického záznamu se budeme zabývat tzv. neinvazními metodami. Obecně v medicíně spojením *neinvazivní metody* myslíme taková vyšetření, při kterých není nutný zásah do lidského těla jako např. odběr krve. My v kontextu práce samozřejmě eliminujeme i požadavek na připevnění jakéhokoli snímače na tělo snímané osoby.

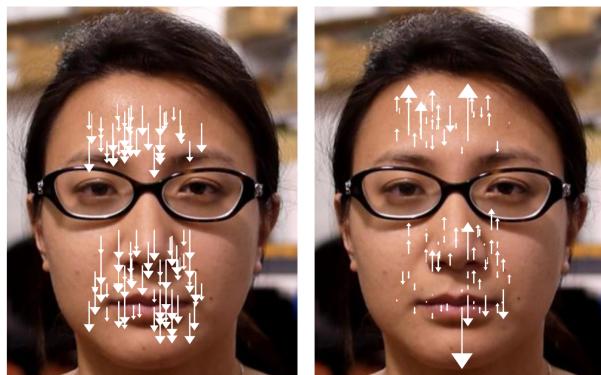
Pro účely této práce se nabízí dvě neinvazivní vyšetřovací metody a to:

- balistokardiografie, [3]
- pletysmografie [4].

Balistokardiografie je neinvazivní metoda, která snímá pohyby částí těla způsobené činností kardiovaskulárního systému. Detekce tepu pomocí balistokardiografie spočívá v analýze pohybů obličeje s omezením na ty pohyby, které odpovídají přílivu krve při každém úderu srdce.

Využitím této metody se již zabývala skupina vědců z MIT v publikaci [5]. Pro experimentální určení tepu z optického záznamu využívají videa obličeje o minimální délce 60–90 sekund. V algoritmu využívají sledování obličeje pomocí registrace bodů v obrazu. Princip sledování je zobrazen na Obrázku 3.1. Následně změny přivedou na 1D signál, ve kterém hledají vrcholy grafu, které odpovídají úderům srdce.

3. APARÁTY PRO REALIZACI PRÁCE



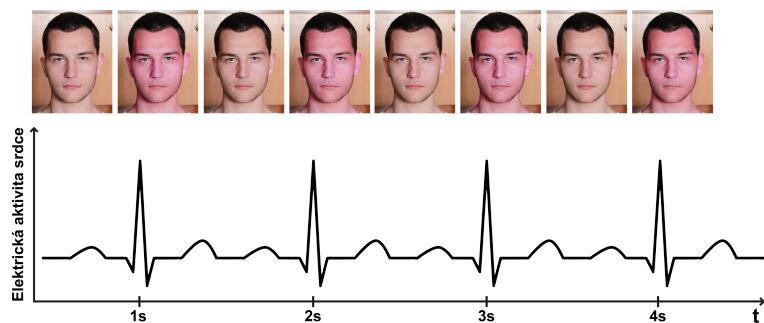
Obrázek 3.1: Příklad vlastních vektorů při detekci tepu pomocí balistokardiografických pohybů obličeje. Každá bílá šipka odpovídá amplitudě a směru kontribuce příslušného registrovaného bodu k vlastnímu vektoru. Převzato z [5].

Velkou výhodou jsou malé, resp. žádné nároky na viditelnost kůže. Např. v publikaci [5] je dosaženo téměř stejných výsledků i v případě, že má snímaný subjekt na obličeji masku. Naopak největším problémem tohoto postupu je příliš velká časová náročnost, už jen samotné snímání podle autorů a jejich měření musí být minimálně minuta. Tímto se postup stává pro účely této práce nevhodným.

Pletysmografie obecně vychází z měření změn velikosti orgánů. Vyšetření je realizováno pomocí pletysmografu¹. Z této metody vychází fotopletysmografie (PPG). Tato disciplína byla představena v roce 1930 a využívá světelnou odrazivost a propustnost krve. Vychází z principu, že krev absorbuje více světla než okolní tkáň [6].

Hemoglobin, metaloprotein červených krvinek, je přenášen prouďním krve a vytváří tak pletysmografické signály. Tyto signály se projevují změnou hodnot intenzit barevných kanálů pokožky v čase. Nejsilnější změny jsou v zeleném kanálu, neboť hemoglobin absorbuje zelenou složku světla více než jiné [7]. Z tohoto důvodu se zpravidla nejvíce zkoumají intenzity právě zeleného kanálu. Princip změny intenzit v pokožce je ilustrován v Obrázku 3.2.

1. Přístroj, který se používá při běžných vyšetřeních, např. při pletysmografii plic.



Obrázek 3.2: Zvýrazněné barevné změny způsobené fluktuací krve. V horní řadě jsou snímky videa v daných časových okamžicích. Ve spodní řadě je potom ilustrace tepu. Tepové údery by měly korespondovat se snímky, ve kterých je pokožka odlišně zbarvená. Tyto barevné změny jsou okem nepostřehnutelné.

Čím větší plochu ve snímcích záznamu zabírá pokožka, tím se dají změny detektovat přesněji. Nejstabilnější využití tohoto přístupu se zdá být v případě, že pokožka zabírá celou plochu záznamu². Aplikace se proto bude snažit najít ve snímcích záznamu obličeje vhodnou masku pokožky, kterou bude dále zpracovávat.

Pro detekci tepu je vybrána tato metoda, protože výsledky experimentů v [4] se zdají být časově méně náročné (délka záznamu i samotné zpracování) a stabilnější než přecházející metoda vycházející z balistokardiografie.

3.2 Vizualizace tepu

Vizualizace tepu je založená na algoritmu nazvaném *Eulerovské zvýraznění* (angl. *Eulerian Magnification*) [2]. Cílem tohoto postupu je odhalovat změny v optických záznamech, které je pouhým okem nemožné (nebo velmi těžké) pozorovat.

Zaměřuje se na dva typy změn:

2. Tento postup využívá např. aplikace *Instant Heart Rate* dostupná pro platformu Android z <<https://play.google.com/store/apps/details?id=modula.android.instantheartrate>>.

3. APARÁTY PRO REALIZACI PRÁCE

- zvýraznění malých pohybů,
- zvýraznění barevných změn.

Postup pro odhalení barevných změn funguje na principu zvýrazňování frekvencí v časové frekvenční doméně (bez omezení na objekt snímání). V publikaci je mj. přímo demonstrované zvýraznění proudění krve v obličeji na Obrázku 3.2 s dopředu známou tepovou frekvencí.

Dostupný je implementovaný algoritmus³ v platformě MATLAB⁴. Implementace funguje příliš pomalu na to, aby byla použitelná v kontextu online zpracování. Zpracování videa o délce 10 sekund (při rozlišení 800x600 pixelů a vzorkovací frekvenci 30 snímků/s) trvá přes 2 minuty. Proto je nutné algoritmus reimplementovat, podrobněji prozkoumat, případně upravit, aby byl použitelný i pro tuto práci.

3. Implementace v MATLABu i spustitelný program je dostupný z webových stránek projektu <<http://people.csail.mit.edu/mrub/evm/>>

4. MATLAB je maticově orientovaná platforma pro řešení vědeckých a inženýrských výpočtů s výbornou schopností vyjadřovat matematické výpočty. Blíže viz <<http://www.mathworks.com/products/matlab/>>.

4 Použité vývojové a výzkumné nástroje

Cíle této práce kladou velké nároky na výkon implementace. Aplikace bude zpracovávat náročný vstup v podobě optického záznamu, v případě použití webkamery dokonce neokomprimovaný.

Po stanovení cílů aplikace je pro implementaci využit kompilovaný jazyk C++. Pro operace zpracování obrazu bude použita knihovna OpenCV¹. Jednoduché uživatelské rozhraní je implementováno pomocí frameworku Qt².

Pro nároky aplikace je nutné vícevláknové programování, které citelně sníží časové nároky na zpracování. Pro tyto účely se nabízí část knihovny Boost³ pro C++.

Vizualizace dat (generování grafů) bude provedeno v MATLABu. Ostatní použité nástroje budou popsány přímo v části popisu aplikace a její demonstrace.

Aplikace bude vyvíjena na operačním systému OS X, ale díky povaze všech použitých technologií bude spustitelná i na ostatních UNIXových distribucích a s drobnými úpravami i na systému Windows.

4.1 Použitý hardware

Aplikaci je vyvíjena a testována na notebooku MacBook Pro v konfiguraci Intel Core i5 2.7 GHz Broadwell, RAM 8GB s integrovanou grafickou kartou Intel Iris Graphics 6100.

Pro pořizování záznamů jsem používal různé přístroje od poloprofesionální zrcadlovky Canon 60D, přes běžný kompakt Nikon, až po integravanou webkameru v notebooku. Dále budou stanoveny požadavky na vstupní video.

-
1. OpenCV (*Open source computer vision*) je volně šířitelná knihovna navržená pro efektivní výpočty se silným důrazem na real-time práci. Dostupné z: <<http://opencv.org/>>.
 2. Qt je multiplatformní framework pro tvorbu uživatelského rozhraní. Dostupné z <<https://www.qt.io/>>
 3. Boost poskytuje volně šířitelné, přenositelné knihovny pro C++. Dostupné z: <<http://www.boost.org/>>

5 Postup

5.1 Objektivní pravda

Pojmem *objektivní pravda* (z angl. *ground truth*) považujeme hodnotu, o jejíž pravdivosti a přesnosti nemáme pochyb a vůči které se porovnávají další hodnoty, získané např. experimentem měření. V rané fázi práce byla jako objektivní pravda použito počítání tepu srdce pomocí prstu přiloženého na zápěstí za časovou jednotku a následné přepočítání na výsledný tep. Ovšem pro zpřesnění výsledků bylo počítání nahrazeno měřením pomocí tlakoměru, který mimo tlak měří i tep. U běžně dostupných tlakoměrů se udává chyba měření tepu $\pm 5\%$. Zvolen byl běžně dostupný tlakoměr, který v uživatelském manuálu [8] udává stejnou hodnotu.

5.2 Postup určení tepu

Nejprve práce testuje, zda metoda pro určení tepu opravdu funguje, zda je pro práci použitelná a co je potřeba zaručit pro stabilitu algoritmu. Bylo třeba začít nahrávat různé záznamy kůže a vyvíjet postup, pomocí kterého je možno extrahovat potřebné informace k určení tepu.

Základní postup bude demonstrován na záznamu čela obličeje, viz Obrázek 5.1. Tento záznam je pořízen v situaci, kdy snímaný leží na zemi a zároveň si měří tep pomocí tlakoměru.

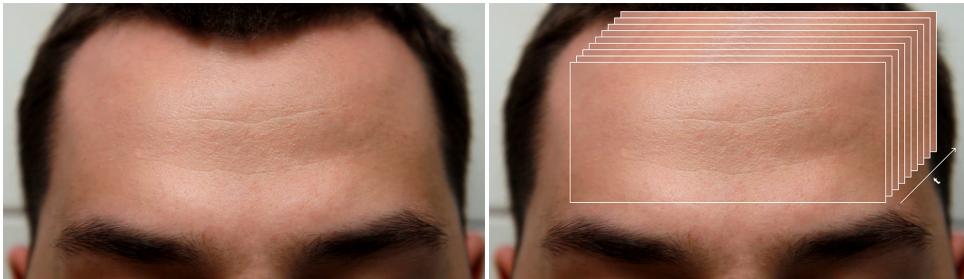
Pro ověření byla maska kůže pro analýzu určena manuálně. Masku je zvolena tak, aby kůže pokryvala celou plochu, viz Obrázek 5.1 vpravo. Právě na této ploše jsou nejsnáze dostupné informace, které chceme použít, v podobě změn barevné intenzity v čase.

Do vektoru s video záznamem $V(i)$ budou vloženy obrazy, kde i -tá hodnota pole bude charakterizovat i -tý snímek ze záznamu. Máme více možností na co se v každém snímku zaměřit:

1. Průměr intenzit všech kanálů

$$\frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} r(x,y) + g(x,y) + b(x,y)}{M * N}, \quad (5.1)$$

5. Postup



Obrázek 5.1: Záznam čela a manuálně určená maska pro zpracování.

2. Průměr intenzit zeleného kanálu

$$\frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} g(x, y)}{M * N}, \quad (5.2)$$

3. Medián intenzit zeleného kanálu

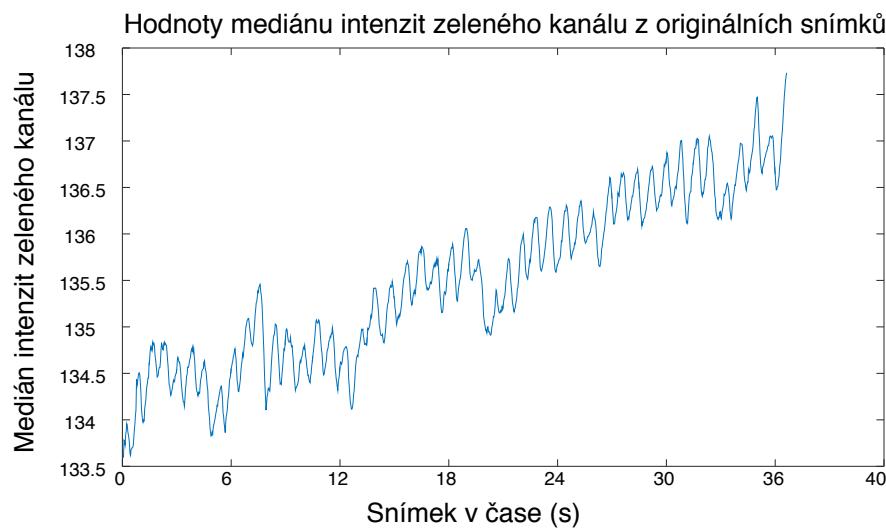
$$\text{median}(g(x, y)), x \in [0, M - 1], y \in [0, N - 1]. \quad (5.3)$$

V tomto případě jsou rozdíly mezi různými přístupy minimální. Zkoumaný záznam je téměř ideální pro extrakci informací. Kůže je stabilizovaná (snímaný má opřenou hlavu o zem) a zabírá většinu záběru, záznam je ve vysokém rozlišení (1920x1080) a ve vzorkovací frekvenci 30 snímků/s. Dále budou rozdíly ilustrovány na méně kvalitních vstupech, kdy se výsledky budou lišit.

Pro ilustrační účely je zvolena analýza pomocí *mediánu intenzit zeleného kanálu* měřena v určené masce čela. Hodnoty mediánů zobrazme pomocí grafu na Obrázku 5.2. Už v tomto neupraveném grafu jsou viditelné periodické frekvence. Za funkci barevných změn odpovídajících tepu je považována cosinusoida. V analyzovaném poli $V(i)$ délky K se nachází časové frekvence F

$$F \in [0, \frac{K - 1}{K} * f_{ps}]. \quad (5.4)$$

Nejdříve je pole převedeno pomocí Fourierovy transformace do frekvenční domény. Pro omezení frekvencí na ty, které jsou pro tento účel informačně přínosné, využijeme dříve zmíněný obvyklý interval tepu pro většinu lidí. Na základě této informace je použit lineární filtr

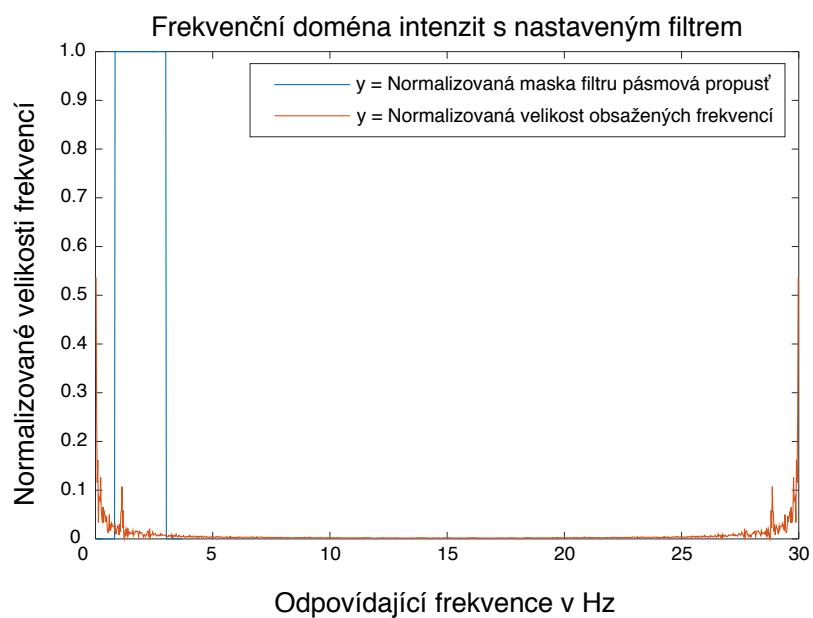


Obrázek 5.2: Originální hodnoty mediánů intenzit zeleného kanálu v jednotlivých snímcích záznamu.

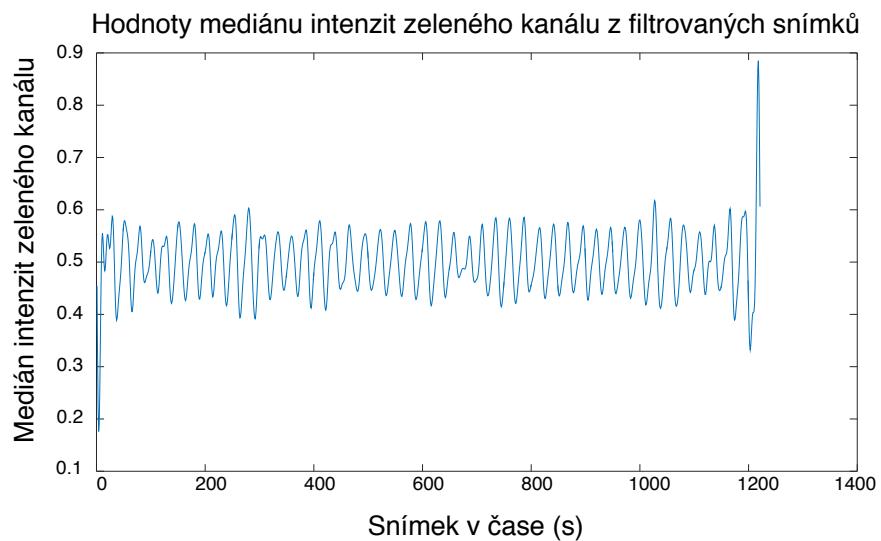
pásmovej propusť ve frekvenční doméně. Jako rozsah tohoto filtru jsou nastaveny hodnoty 55 resp. 180 tepů/min, které jsou převedeny do jednotky Hz, tedy 55/60 Hz a 180/60 Hz. V grafu na Obrázku 5.3 je zobrazená reálná složka frekvenční domény dohromady s filtrem pásmová propusť. Na ose x jsou označeny přítomné frekvence (v jednotce Hz) a na ose y jejich normalizované hodnoty (v jaké míře je daná frekvence přítomna).

Graf změn mediánu v závislosti na čase je vidět na Obrázku 5.4. V tomto případě byl v původním grafu na Obrázku 5.2 vidět celkový nárůst hodnot mediánu v čase stejně jako příliš rychlé změny způsobené šumem, nebo malými posuny obličeje v záběru. Nyní po filtraci zbyly už jen frekvence, které potenciálně mohou odpovídat tepu, ostatní byly odfiltrovány. Nejsilnější ze zbylých frekvencí nich bude potom odpovídat výslednému tepu.

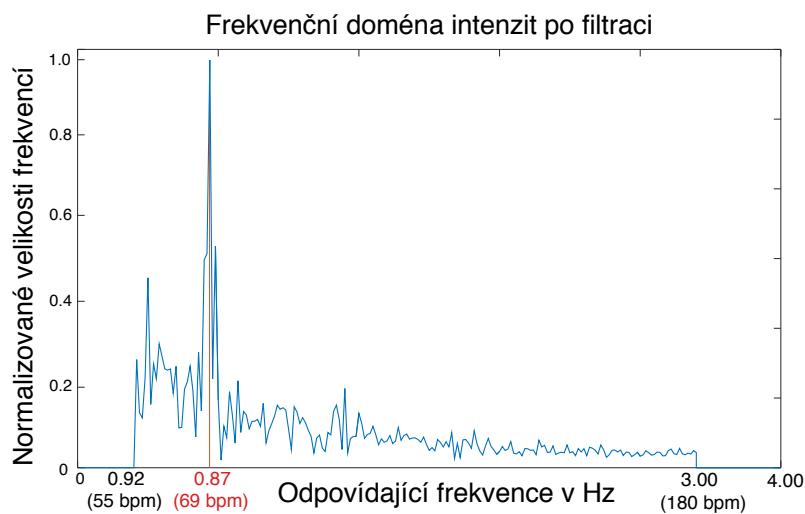
Na grafu frekvenční domény po filtraci pásmovou propustí na Obrázku 5.5, je zřetelně vidět nejsilnější obsažená frekvence 0,87 Hz (po přepočtení 69 tepů/min). Skutečná hodnota tepu pro tento záznam byla 67 tepů/min, při udávané chybě měření 5% jsou tedy hodnoty



Obrázek 5.3: Frekvenční doména naměřených intenzit spolu s maskou filtru pásmová propusť.



Obrázek 5.4: Hodnoty mediánů intenzit zeleného kanálu v jednotlivých snímcích záznamu po filtraci.



Obrázek 5.5: Frekvenční doména naměřených intenzit po filtraci.

5. POSTUP



Obrázek 5.6: Příklady různorodých záznamů pro ověření detekčních schopností postupu.

v rozmezí 64–70 tepů/min považovány za korektní. Vypočítaný tep potom do rozmezí spadá.

Minimální délka

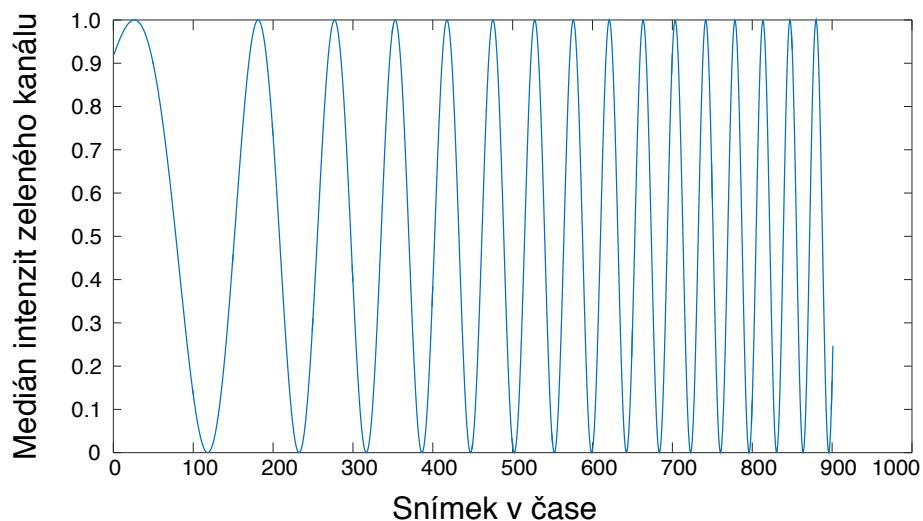
Pro stabilní určování tepu pro různé případy je nutné správně určit minimální délku záznamu. Tento problém se řešil po celou dobu implementace. Během testování se ukázalo, že nejstabilnější délkou záznamu je 500 snímků (při vzorkovací frekvenci 30 snímků/s cca 16 sekund). Tato informace bude podrobněji rozvedena v sekci testování.

Iniciální ověření na záznamech

Mimo předchozího záznamu bylo potřeba pro iniciální ověření analyzovat různorodá videa. Jako netradiční případ se nabízí tep dítěte, který je podstatně vyšší než tep dospělého. Další atypický záznam může být obličej snědého člověka. Jako poslední záznam pro iniciální testování je nasnímaná ruka na podlaze. Na příkladu ruky i snědého obličeje určení fungovalo správně. U záznamu dítěte nebyl k dispozici přesný tep, v tomto případě šlo především o to, zda postup správně detekuje i vyšší tepy. Díky znalosti o tepu dětí se dal předpokládat tep vyšší než 100 a to se potvrdilo, neboť výsledný tep byl 113 tepů/min. Náhledy jsou zobrazeny v Obrázku 5.6.

Limitace postupu

Filtrování ve frekvenční doméně má nesmírnou výhodu v podobě zaměření se na potřebné frekvence. Analýza frekvencí je imunní vůči

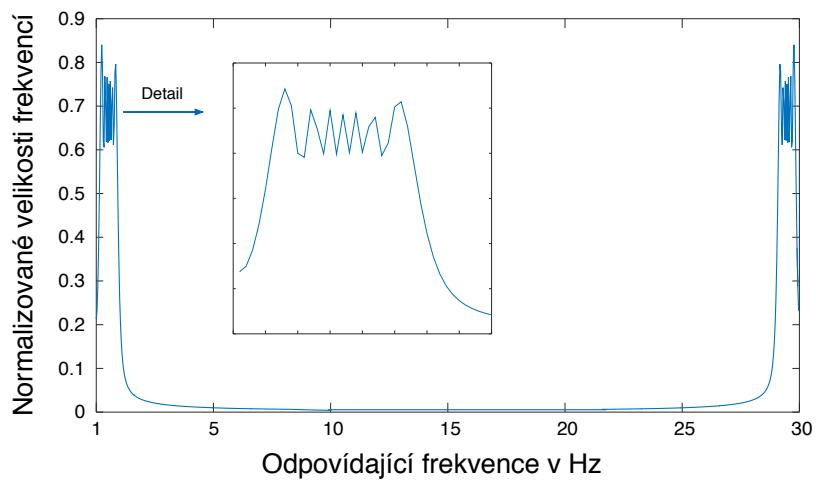


Obrázek 5.7: Simulace hodnot mediánů intenzit zeleného kanálu v jednotlivých snímcích záznamu po filtraci.

globálním změnám světla. Globální změna světla je změna intenzit všech pixelů najednou v závislosti na čase — důležité je, že tato změna se musí týkat všech pixelů v obrazu. V případě, že by na analyzovanou pozici ve snímcích v pravidelné frekvenci blikal dostatečně silný zdroj světla, algoritmus skončí s nekorektními výsledky, kdy výsledná hodnota domnělého tepu bude odpovídat právě frekvenci blikajícího světla. Postup je také do jisté míry robustní i vůči šumu, který se při analýze projevuje jako velmi vysoké frekvence.

Hledání nejsilnější frekvence ve frekvenční doméně selhává v případě nepravidelné frekvence, např. hodně se zrychlující, nebo zpomalující tep, což je zobrazeno na Obrázku 5.8. V tomto případě při hledání nejsilnější frekvence není žádná frekvence dominantní, zobrazeno na obrázku 5.7. Na druhou stranu v takovéto situaci není příliš jasné jaká hodnota tepu je správná. Povaha navrhovaného měření, stejně jako u ostatních konvenčních metod měření tepu, je založena na očekávání pravidelného tepu. Pro určení nepravidelného tepu by musel být využit jiný algoritmus pro detekci lokálních maxim, ale i tak by se dalo jen přepočítat maxima na tepy/min, což by odpovídalo pouze průměrování.

5. POSTUP



Obrázek 5.8: Frekvenční doména simulace zrychlujícího se pulzu s detailem zaměřeným na relevantní frekvence.

Vzhledem k tomu, že se práce zaměřuje na klidovou tepovou frekvenci, tak tento problém nebude vadit. Z testovacích videí vyplynulo, že se během nahrávání po dobu cca 15 vteřin nezmění tep natolik, aby změna způsobila selhání.

6 Implementace

Aplikace nejdříve zobrazí jednoduché uživatelské rozhraní pro volbu různých režimů, načtení videa a volbu pro uložení výstupu viz Obrázek 6.1. Po spuštění se fungování aplikace liší pro *reálné* režimy (zobrazení výstupu spolu s výpočty) a *statický* režim (zpracování celého videa a následné zobrazení výstupu).

6.1 Datové struktury

Pro reprezentaci snímku je využit datový typ *OpenCV::cv::Mat*. Tento datový typ podporuje více typů obrazů, takže je použitelný pro šedotónové i barevné obrazy. Také umožňuje reprezentaci obrazových hodnot v datovém typu *float*.

Záznam bude uložen pomocí datového typu *std::vector*, konkrétně potom jako *vector <Mat>* (vektor matic). Typ *vector* je v *OpenCV* doporučován jako standardní typ a měl by v algoritmech pracovat nejrychleji.

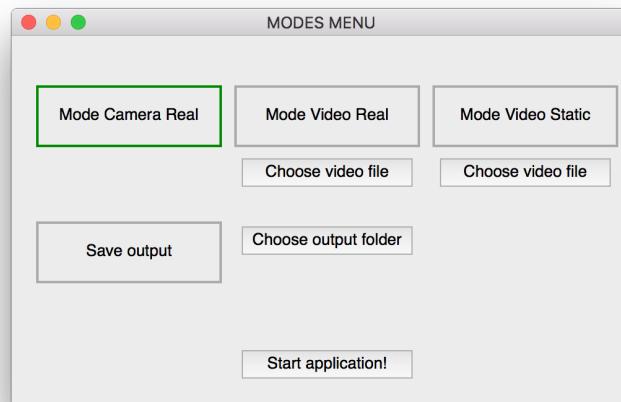
6.2 Popis aplikace

Popis aplikace odpovídá použití jednoho z *reálných* režimů běhu aplikace. Po vybrání módu se zobrazí okno, kde jsou zobrazeny 2 snímky vedle sebe. Nejdříve budou totožné a po dokončení iniciálního výpočtu jeden slouží pro zobrazení vstupního záznamu a druhý naopak zobrazuje zpracované snímky. Rozdíl oproti statickému režimu bude popsán po vysvětlení potřebných částí aplikace.

6.2.1 Hlavní cyklus

Snímky se zobrazují v hlavním cyklu, kdy se při každé iteraci načítá další snímek ze zvoleného vstupu. Samotná časová náročnost tohoto cyklu přímo ovlivňuje plynulost aplikace. V případě externího záznamu se jedná pouze o *trhání* záznamu, které neovlivňuje následné zpracování a výpočty. Při použití webkamery se ale jedná přímo o změnu vzorkovací frekvence, protože další snímek ze vstupního proudu webkamery se vždy načte až po uplynutí jednoho cyklu.

6. IMPLEMENTACE



Obrázek 6.1: Uživatelské rozhraní pro volbu režimu aplikace.

V každé iteraci cyklu je originální vstupní obraz vložen do datové fronty¹, který bude zpracován až po naplnění najednou. Velikost fronty je přímo určena potřebným počtem hodnot pro určení tepu a je nastavena na hodnotu 500 snímků. Tato volba je odůvodněna v Kapitole 6.3.

Vzorkovací frekvence

Už samotné zobrazování snímků při výběru vstupu z webkamery je náročné, protože z proudu webkamery získáváme bezztrátové snímkы bez komprese. Pravidelná maximální vzorkovací frekvence (podporovaná webkamerou) v prostředí *OpenCV* není zaručena ani při pouhém zobrazování v cyklu bez jakýchkoli úprav snímků. Bohužel není k dispozici ani žádná funkce, která by vracela hodnotu aktuální vzorkovací frekvence. Z tohoto důvodu aplikace pro určení využívá alternativní přístup, kdy počítá dobu každé iterace cyklu a aktualizuje si propočet aktuální vzorkovací frekvenci. Pro výpočet je využita knihovna

1. Za datovou frontu je považován záznam dle standardní definice 2.3, který je používán pro shromažďování vstupních snímků.

*chrono*², která nabízí funkce pro měření uplynulého času a je přenositelná napříč různými operačními systémy. Narozdíl od některých ostatních měřících funkcí vrací opravdu uplynulý čas a ne jenom čas spotřebovaný procesorem. Přesnost funkce se pohybuje v řádu milisekund.

Pro zvýšení plynulosti uživatelského rozhraní, resp. pro vyšší vzorkovací frekvenci webkamery, jsou zobrazované snímky podvzorkovány (pouze pro uživatelské rozhraní, interně se do paměti ukládají v plné originální velikosti). Podvzorkování musí probíhat bez interpolace, proto vždy zmenšujeme velikost v násobcích mocniny čísla 2. Interpolace je příliš výpočetně náročná, aby mohla být prováděna v hlavním cyklu. Mocnina čísla 2 je zvolena tak, aby šířka, respektive výška v případě vstupu orientovaného na výšku, byla co nejblíže zdefinované hodnotě *480px*, která na použité konfiguraci zaručovala relativně plynulý obraz.

Inicializace aplikace

Po startu nejdříve probíhá proces inicializace, kdy s prvními 20 snímky aplikace nic nedělá a pouze je zobrazí. Z pravidla se na začátku záznamu nachází artefakty, které se takto ignorují. Jedná se především o úpravu expozice u levnějších (web)kamer, často také začíná záznam stisknutím spouště snímacího zařízení a s tím spojené záhvěvy scény.

Detekce

Pro fungování algoritmu pro určení tepu je třeba izolovat dostatečně velkou oblast, kterou bude vyplňovat pokožka, jak bylo popsáno v Kapitole 3.1. V obličeji se jako největší plocha celistvé kůže nabízí oblast čela, proto je třeba provádět co nejpřesnější detekci obličeje. Pro lepší určení pozice čela je vhodné použít také detekci očí, díky ní se pozice čela určuje podstatně přesněji než jen ze znalosti pozice obličeje.

Pro detekci obličeje a očí využívá práce algoritmus *Haar Feature-based cascade classifiers*. Tuto metodu představili Paul Viola a Michael Jones v roce 2001 v publikaci [9]. Detekce je založena na porovnávání jasových rozdílů pomocí různých konvolučních jader. Nejdříve je

2. Bližší informace viz <<http://www.cplusplus.com/reference/chrono/>>

6. IMPLEMENTACE

potřeba klasifikátor *trénovat* pomocí strojového učení z velkého množství vzorků (pozitivních i negativních). *OpenCV* pro použití nabízí již předtrénované klasifikátory. Mimo stabilitu algoritmu je nesmírnou výhodou i časová efektivita této detekce.

Detekce pomocí kaskádových klasifikátorů může skončit více výsledky, z nichž jsou některé nekorektní. Práce využívá funkci, která v případě více detekovaných obličejů vrací ten, který ve snímku zábírá největší plochu. To v kontextu zpracovávaných záznamů zaručuje správný výsledek. V případě více výsledků při detekci očí, jsou vybrány objekty, které mezi sebou mají nejmenší rozdíl v y-ové složce polohy.

Pro detekce je v programu vytvořená zvláštní třída *Detector*, která pracuje v samostatném vlákně, aby zbytečně nezdržovala hlavní cyklus. Třída se snaží detektovat jak obličej jak ve snímku s originální velikostí, tak i ve zmenšeném snímku.

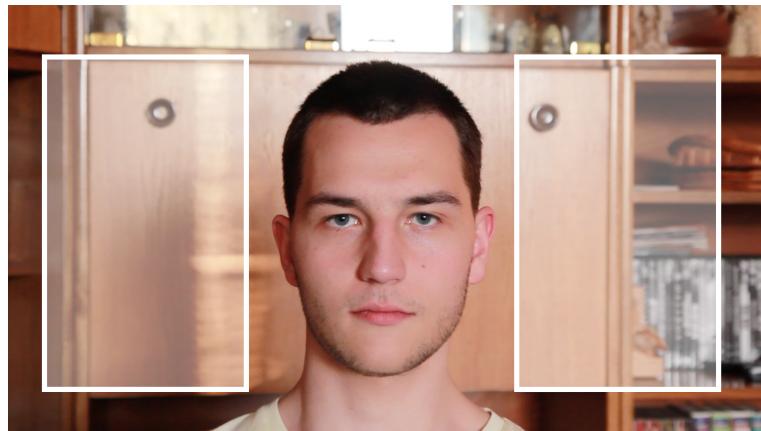
Potom, co je detekován obličej i oči, je na aktuální snímek v uživatelském rozhraní vykresleno ohrazení obličeje i čela, které se při každé další detekci aktualizuje. Uživatel potom vidí sledování obličeje. Toto vykreslování je prováděno především pro zvýšení atraktivity (časové nároky na vykreslení jsou zanedbatelné).

6.2.2 Analýza video záznamů

Fronta se po naplnění v samostatném vlákně předává třídě *Middleware*, která se chová jako mezivrstva uživatelského rozhraní a výpočetní třídy *BpmVideoProcessor*. Ve třídě *BpmVideoProcessor* se odehrávají všechny výpočty potřebné pro určení tepu a jeho vizualizaci.

Při inicializaci třídy *BpmVideoProcessor* se provádí důležitá část určení plochy čela. Do třídy předáváme pozici a velikost již detekovaného obličeje při detekcích v uživatelském rozhraní. Detekce očí v některých případech nefunguje zcela přesně a určení plochy čela pro následné zpracování je natolik důležité, že je jeho určování komplikovanější.

Čelo se při přímém pohledu člověka do kamery nachází (horizontálně) ve středu obličeje. Detekce probíhá postupně pro každý snímek v záznamu dokud nejsou splněny definované podmínky. V každém snímku je provedena detekce očí pomocí kaskádových klasifikátorů. Střed čela je potom určen posunutí bodu, který se nachází uprostřed



Obrázek 6.2: Vhodná místa pro analýzu globálních světelných změn.

oblasti mezi očima a jeho posunutím přibližně o jednu desetinu výšky obličeje nahoru.

Pozice čela je považována za korektní, je-li rozdíl mezi horizontálním středem x_{face} obličeje a středem určeného čela $x_{forehead}$ menší než 15% šířky obličeje w_{face} , neboli:

$$|x_{face} - x_{forehead}| \leq 0,15 * w_{face}. \quad (6.1)$$

Nekorektní pozice čela je použita pouze v případě, že podmínky nejsou splněny v žádném snímku. Tento postup určení čela a definované podmínky se během testování ukázaly jako vhodné.

6.2.3 Určení tepu

Postup pro určení tepu je popsán v Kapitole 3.1. Během práce byly uvažovány také další rozšíření pro zpřesnění určení tepu.

Bylo zkoumáno experimentální potlačení globálních světelných změn, jejichž frekvence spadají do rozsahu uvažovaného jako interval možného tepu. Tyto změny jsou imunní vůči filtrování ve frekvenční doméně. Pro potlačení je tentokrát třeba detektovat prostor mimo obličeji, který je po celou dobu záznamu pozici fixovaný a měly by se v něm projevovat pouze globální změny. Vhodná místa jsou zobrazena na Obrázku 6.2.

6. IMPLEMENTACE

V prostoru je vytvořeno pole s charakteristikou intenzit snímků shodným způsobem jako v masce čela. Vhodným odečtením polí od sebe a následným normalizováním by se měly globální frekvence potlačit a kýžené frekvence odpovídající tepu naopak zvýraznit. Během měření a výsledků ovšem nebyly zaznamenány pozitivní výsledky po aplikaci tohoto postupu, proto nakonec není využit.

Další — také nakonec nevyužitý pokus — spočíval ve stabilizaci záznamu pomocí sledování optického toku³. Před analýzou intenzit byl nejdřív záznam detekovaného čela stabilizován. Zpřesnění výsledků nebylo zaznamenáno, naopak v některých případech, došlo ke zhoršení. K narušení výsledků došlo pravděpodobně kvůli afinním transformacím prováděným pro každý snímek během stabilizace videa a s nimi spojenými interpolacemi.

6.2.4 Vizualizace tepu

Pro vizualizaci tepu je nejdříve třeba vytvořit časoprostorovou řadu ze záznamu detekovaného obličeje, která reprezentuje proměny hodnot intenzit jednotlivých pixelů v závislosti na čase. Pro konstrukci je použit pouze zelený kanál a snímky jsou převedeny do barevného prostoru NTSC, který je použit i v algoritmu *Eulerovské zvýraznění*. Časoprostorová řada bude mít M snímků (pro každý jednopixelový sloupec v detekovaném obličeji jeden snímek v časové řadě), z nichž každý bude mít velikost $K \times N$ (označení M, N, K je ze standardní definice záznamu z Rovnice 2.3). Vytvoření jednoho snímku v časové řadě je ilustrováno na Obrázku 6.3.

Pro každý snímek v časoprostorové řadě je sestrojena Gaussova pyramida. Gaussova je pyramida tvoří obraz ve více stupních rozlišení. Každá úroveň je rozostřena Gaussovým filtrem a následně podvzorkována na polovinu rozlišení úrovně o stupeň vyšší. Na jednotlivé úrovně v pyramidě je aplikován filtr pásmová propust. Jako střed intervalu filtru je nastaven již předem vypočítaný tep a rozsah intervalu je nastaven na 10 tepů/min (nastaveno podle subjektivního dojmu nejlepšího výsledku). Aby bylo možné algoritmus použít v kontextu práce, bylo nutné implementovat řešení, které umožní každou úroveň v pyramidě zpracovávat samostatným vlákнем. Po tomto rozdělení

3. Stabilizace byla inspirována algoritmem dostupným z <<http://nghiaho.com/?p=209>>.



Obrázek 6.3: Ilustrace principu vytvoření jednoho snímku časové řady z detekovaného obličeje.

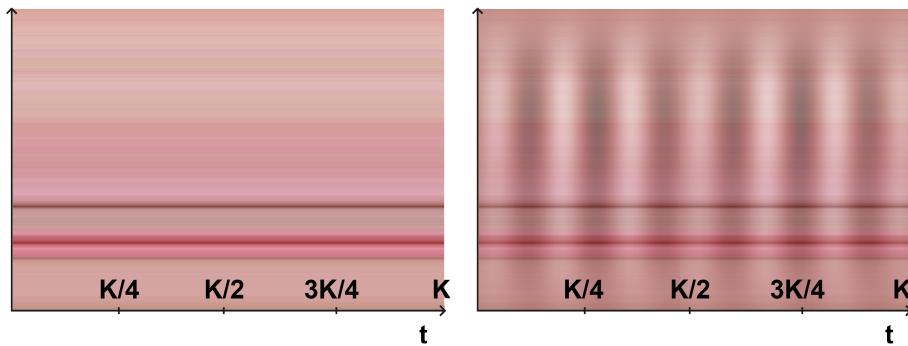
trvá filtrování Gaussovy pyramidy stejnou dobu, jako kdyby byla zpracovávána jen nejvyšší úroveň.

Filtrováním snímků časoprostorové řady je dosaženo potlačení frekvencí, které neodpovídají tepové frekvenci (podobně jako při určování tepu). Po filtraci vždy odpovídající snímek normalizujeme. Z pyramid jsou zpět rekonstruovány snímky časoprostorové řady, která je posléze převedena zpět do formátu standardního záznamu. Ten nazveme jako *vizualizace tepu*. Ukázka zpracování jednoho snímku je zobrazena na Obrázku 6.4.

V hlavním cyklu pořád probíhají detekce obličeje, takže je stále k dispozici aktuální pozice a velikost obličeje v aktuálních vstupních snímcích. Pro kombinaci snímků je nejdříve nutné vytvořit snímek stejné velikosti jako vstupní snímky, který je nulový ve všech barevných kanálech na všech pozicích. Do tohoto snímku je na aktuální pozici obličeje vložen zpracovaný snímek z *vizualizace tepu* viz Obrázek 6.5. Takto je tvořen výstupní snímek při každé iteraci hlavního cyklu.

Ve frontě je po celou dobu maximální počet snímků, při vložení snímku nad kapacitu je nejstarší snímek odstraněn. Po dokončení prvního výpočtu aplikace se ihned předává fronta s novými snímkami k dalšímu zpracování. Opakovaným zpracováním je možné detektovat i pomalu se měnící tep během snímání. Takto aplikace funguje dokud ji uživatel neukončí.

6. IMPLEMENTACE



Obrázek 6.4: Vlevo je zobrazen originální snímek z časové řady. Vpravo je součet originálního snímku a snímku po filtrování a zvýraznění frekvencí. Proměnná K odpovídá počtu snímků v původní frontě.

6.2.5 Rozdíl mezi režimy

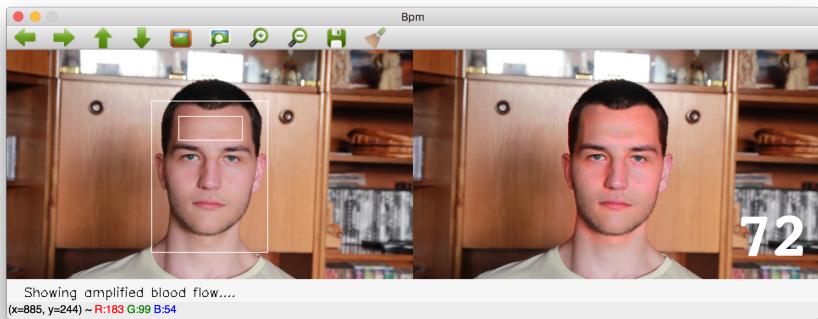
Rozdíl mezi *reálným* režimem při použití webkamery a při načteném externím je opravdu jen v odlišeném vstupu. U webkamery je třeba kontrolovat vzorkovací frekvenci jak bylo zmíněno dříve.

Jedním z rozdílů mezi *reálným* a *statickým* režimem je ve sběru snímků do datového proudu. Zatímco v *reálném* režimu je zobrazen obraz spolu s ukládáním snímků, ve *statickém* režimu jsou snímky ukládány aniž by bylo původní video zobrazené. Pro rychlejší vkládání dat je program vyvíjen s využitím knihovny *OpenCV* s podporou software *FFmpeg*⁴.

Výpočet tepu je proveden shodným způsobem pro všechny režimy. Při výpočtu vizualizace ve *statickém* režimu je záznam rozdělen na části po 250 snímcích a každá je zpracována ve vlastním vlákně. Toto nastavení velmi zrychluje průběh celého výpočtu.

Odlišně je tvořeno i výstupní video. Díky tomu, že ve *statickém* režimu není při výpočtech nic zobrazeno, aplikace po jeho dokončení může nastavit pozici ve videu zpět na začátek. Proto při kombinaci výstupních snímků vždy snímek z *vizualizace tepu* odpovídá vstupnímu snímkmu. Toto není možné při *reálných* režimech, protože by bylo nutné zobrazovat neaktuální snímkы. Tím by byla porušena snaha

4. *FFmpeg* je multiplatformní software pro přehrávání, streamování a konvertování videa a zvuku. Dostupné z <<https://ffmpeg.org/>>



Obrázek 6.5: Uživatelské rozhraní po dokončení výpočtu tepu i jeho vizualizace. Vlevo zobrazen aktuální vstupní snímek se zobrazením detekovaného obličeje a čela. Vpravo je zobrazen výsledný snímek s vizualizací tepu a s hodnotou vypočítaného tepu.

o online zpracování, protože by výstupní snímek časově neodpovídal vstupnímu.

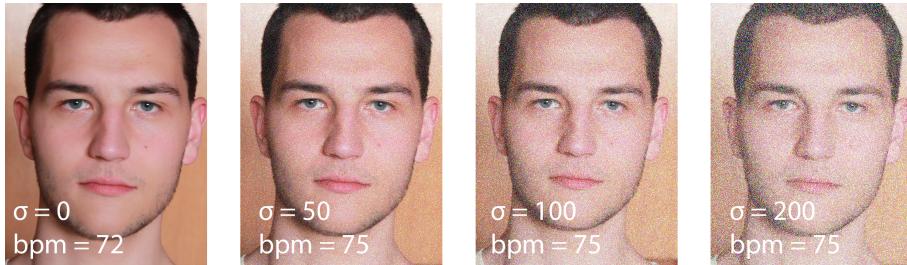
6.3 Měření a požadavky na vstupní záznam

Během implementace bylo nutné určit konstanty, které jsou ve výsledné verzi použity. Měření bylo provedeno na mnoha rozdílných záznamech v rozmezí tepu 58–82 tepů/min. Reálný tep je určen pomocí tlakoměru s již zmíněnou nepřesností $\pm 5\%$, proto jsou vypočítané hodnoty tepu spadající do tohoto intervalu považovány za korektní a chyba je udávána až pro hodnoty mimo interval.

Nejdříve ukážeme, jak je důležité korektně určit masku čela. V Tabulce 6.1 je vidět, že pokus o určení tepu bez provedení detekce čela je nekorektní především pro situace, kdy je reálný tep vyšší. V této situaci postup selže jak při použití pouhé masky obličeje, tak při použití celých snímků.

Další určená hodnota pomocí měření v Tabulce 6.2 je přístup pro tvorbu charakteristických informací o intenzitách v každém snímku (popsáno v kapitole 5.2). Měření opravdu ukázalo, že se barevné

6. IMPLEMENTACE



Obrázek 6.6: Na levém snímku je zobrazen snímek z původního videa. Pro testování je uměle přidáván Gaussovský šum a tep počítán znovu. Na každém snímku je zobrazena standardní odchylka šumu σ a výsledný tep bpm. I z velmi silně zašumělého videa, ilustrovaného pravým snímkem, se podařilo vypočítat tep spadající do korektního intervalu.

změny nejvíce projevují v zeleném kanále. Nejpřesnějšího určení tepu je dosaženo s využitím určování průměrných intenzit hodnot zeleného kanálu, proto bude tento postup nastaven jako výchozí.

Podle provedených měření v Tabulce 6.3 byl určen minimální počet snímků pro určení určení tepu na hodnotu **500** snímků. Tato hodnota odpovídá při vstupní vzorkovací frekvenci 30 snímků/s cca 16 sekundám. Fronta by samozřejmě mohla obsahovat i více snímků, ale pro rychlosť zpracování je nutné použít co nejnižší počet.

Pro možnost určení tepu je třeba určit i minimální vzorkovací frekvenci. Frekvence obsažené v záznamu můžeme díky omezení na interval běžného tepu považovat za pásmově omezený. Podle Nyquistova pravidla je určena minimální vzorkovací frekvenci vstupu $f_{ps_{min}}$ jako dvojnásobek tepové frekvence snímané osoby bpm_{real} , neboli:

$$f_{ps_{min}} \geq 2 * bpm_{real}, \quad (6.2)$$

kde bpm_{real} je reálná tepová frekvence. Např. při tepu 60 tepů/min (1 Hz) je potřeba nasnímat záznam minimální vzorkovací frekvencí 2 snímků/s (2 Hz). Toto omezení platí pro samotné určení tepu, nicméně pro uživatele je video už příliš trhané, navíc vizualizace tepu při takto nízké vzorkovací už nevypadá dobře.

Určení a vizualizace tepu funguje dobře při denním osvětlení. Během testovacím měření zobrazeném na Obrázku 6.6 se ukázalo, že

6. IMPLEMENTACE

Reálný tep záznamu	Celý snímek	Detekovaný obličej	Detekované čelo
70	129	129	78
72	65	65	72
81	58	58	83
82	57	57	87

Tabulka 6.1: Výsledky ovlivnění výsledků při určení masky pro výpočet intenzit snímků. Princip byl popsán v Kapitole 6.2.1. Hodnoty ilustrují situace, kdy je pro výpočet intenzit využit nejdříve celý snímek, potom oblast detekovaného obličeje a nakonec čela.

postup pro určení tepu není příliš náchylný na šum. Při použití určitého umělého osvětlení na bázi zářivek určení tepu nepracuje příliš dobře i při dostatečném nasvícení. Dá se říci, že je pro aplikaci vhodnější zašumělé video za denního světla než záznam s zářivkovým nasvícením — byť kvalitním. Na šum i na nedostatek světla je velmi náchylná samotná vizualizace, která při nekvalitních vstupech nepracuje dobře.

Největším problémem pro určení tepu je záznam s příliš velkými pohyby obličeje. Taková videa zpravidla vedou k nekorektním výsledkům. Bohužel nepomáhá ani stabilizace obrazu před výpočty jak bylo diskutováno v Kapitole 6.2.3. Obecně se dá říci, že příliš velké pohyby obličeje nastávají při snímání bezprostředně po fyzické zátěži vlivem zrychleného dýchání. Na druhou stranu i před měřením tlakoměrem se podle pokynů má člověk vždy 5 min uklidnit a být v klidu, proto tento problém nelze aplikaci vytykat.

6. IMPLEMENTACE

Reálný tep	Všechny kanály		Zelený kanál		Medián zel. kanálu	
	Výsl. tep	Chyba [%]	Výsl. tep	Chyba [%]	Výsl. tep	Chyba [%]
58	68	11	61	0	61	0
60	58	0	58	0	58	0
62	62	0	68	5	68	5
66	66	0	66	0	66	0
68	68	0	68	0	72	0
70	78	3	78	3	78	3
72	71	0	71	0	78	3
73	72	0	75	0	75	0
81	58	25	83	0	115	35
82	57	27	87	1	90	5
Součet chyb		66		9		51

Tabulka 6.2: Výsledky testování rozdílných charakteristik intenzit snímků popsaných v Kapitole 5.2 pro určení tepu. Ve sloupci *Výsl. tep* je vždy vypočítaný tep aplikací pomocí příslušného přístupu v hlavičce. Ve sloupci *Chyba* je uvedena chyba měření v případě, že výsledný tep spadá mimo povolený interval. Při tomto měření ještě nebyl kladen důraz na výkon aplikace, takže nebylo nutné použít co nejmenší počet snímků pro určení.

Reálný tep	200 snímků		300 snímků		400 snímků		450 snímků		475 snímků		500 snímků	
	Hodnota	Chyba [%]										
58	72	18	60	0	63	3	60	0	61	0	61	0
60	63	0	60	0	59	0	60	0	61	0	58	0
62	65	0	65	0	69	6	68	5	68	5	68	5
66	78	13	57	10	74	7	73	6	72	4	66	0
68	78	10	73	3	74	4	73	3	72	1	66	0
70	135	82	130	76	79	7	77	4	76	3	78	5
72	79	4	72	0	72	0	71	0	71	0	72	0
73	86	12	77	0	76	0	74	0	73	0	75	0
81	63	18	60	22	86	1	84	0	83	0	83	0
82	60	23	60	23	55	30	56	28	57	27	87	1
součet dílčích chyb	180		134		58		46		40		11	

Tabulka 6.3: Výsledky měření pro určení minimálního počtu snímků pro určení tepu. Pro každý záznam je několikrát určen tep pokaždé s jiným počtem snímků. Při použití více než 500 snímků už nebylo naměřeno zpřesnění výsledků.

7 Závěr

Podařilo se vytvořit funkční aplikaci pro detekci a následnou vizualizaci tepu, s možností výběru více režimů. Samotný režim s využitím webkamery je velice náročný na hardwarové vybavení stroje, na kterém je používán, proto pravděpodobně nebude fungovat na starších hardwarových sestavách. Postup řešení byl obecně velmi náročný na výkon, proto bylo nutné hojně využívat vícevláknové programování pro ulehčení výpočtu. Aplikace umožňuje ve všech režimech uložit výstupní video ve formátu *.avi*.

Při splnění požadavků diskutovaných v Kapitole 6.3 byl u většiny případů tep určen korektně, s maximální odchylkou 5% tepů/min. Detekce tepu nefunguje korektně pro záznamy, ve kterých snímaná osoba příliš hýbe obličejem, např. při snímání po fyzické zátěži. Pro korektní určení tepu je potřeba nejméně 500 snímků, což při vzorkovací frekvenci odpovídá přibližně 16 sekundám. Po dobu běhu aplikace je nutné, aby snímaná osoba co nejméně hýbala obličejem.

Dalším vylepšení do budoucna se nabízí v řešení problémů příliš velkých pohybů obličeje při snímání, které v současném stavu končí nekorektním určením tepu. Pro vylepšení vizualizace tepu by bylo možné zpřesnit obdélníkovou masku obličeje přímo na masku obličeje např. pomocí detekce hran a vizualizaci zobrazovat jen do této masky. Z hlediska výkonu by bylo možné přesunout část výpočtu na grafickou kartu a tím celou aplikaci zrychlit, především potom při využití webkamery.

Bibliografie

- [1] Y. Ostchega et al. „Resting pulse rate reference data for children, adolescents, and adults“. In: *National Health Statistics Reports* (1999–2008).
- [2] H. Wu et al. „Eulerian Video Magnification for Revealing Subtle Changes in the World“. In: *ACM Trans. Graph. (Proceedings SIGGRAPH 2012)* (2012).
- [3] Andrew D. Wiens a T. Inan. Omer. „A Novel System Identification Technique for Improved Wearable Hemodynamics Assessment“. In: *IEEE transactions on bio-medical engineering* 62.5 (2015), s. 1345–1354.
- [4] W. Verkruyse, L. O. Svaasand a J. S Nelson. „Remote plethysmographic imaging using ambient light“. In: *Ex- press* 16 (2008).
- [5] Guha Balakrishnan, Fredo Durand a John Guttag. „Detecting Pulse from Head Motions in Video“. In: *Proceeding CVPR '13 Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition* (2013), s. 3430–3437.
- [6] A. B. Hertzman a C. R. Spealman. „Observations on the finger volume pulse recorded photoelectrically“. In: *Am. J. Physiol* 119 (1937), s. 334–335.
- [7] E. van Kampen a W. Zijlstra. „Determination of hemoglobin and its derivatives“. In: *Advances in Clinical Chemistry*, Academic Press (1968), s. 158.
- [8] a.s. Fast ČR. „Digitální tonometr S manžetou na paži SENCOR SBPP 690, Návod k obsluze“. In: (2015).
- [9] P. Viola a M. Jones. „Rapid object detection using a boosted cascade of simple features“. In: *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on* (2001), s. 511–518.

A Dodatek

V elektronickém repozitáři <<https://github.com/michalsindelar/bpm>> této práce je k dispozici celý kód implementace. V repozitáři se nachází *readme* soubor s návodem ke zkompilování a odkazem na externí úložiště, kde jsou uložena ukázková testovací videa. Pro každé video je vytvořena jedna složka, ve které se nachází:

- originální video nazvané podle reálného tepu snímané osoby,
- výsledné video s vypočítaným tepem a jeho vizualizací,
- snímek detekovaného obličeje a čela,
- textový soubor s průměrnou hodnotou intenzit zeleného kanálu pro každý snímek použité pro detekci tepu,
- ukázku jak ovlivňuje použitá maska čela výsledek,
- porovnání výsledků při použití různých přístupů z Kapitoly 5.2 pro určení tepu v závislosti na počtu použitých snímků v tabulce ve formátu *.csv*.

Současně je v repozitáři připravena větev *feature/testing*, která tyto výsledky dokáže při nastavení skutečného tepu automaticky generovat.

V elektronické příloze *bpm.zip* se nachází aktuální revize aplikace se zdrojovými kódy a v příloze *testing_samples.zip* je rozšířená verze zmíněných testovacích videí.