

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 projektu je vytvořit aplikaci, která nabídne vytvoření krátkého kamerového záznamu obličeje, přehrání záznamu a současně vyobrazení srdečního tepu (například publikování ikonky v detekovaném rytmu). Srdeční tep bude detekován pouze na základě analýzy obrazu z kamery. Metoda je založena na zdůraznění slabé, okem těžko pozorovatelné fluktuace barevného odstínu kůže vlivem přítomnosti nebo nepřítomnosti krve.

Těžiště práce je nalezení a zrealizovaní procedury zachycení a zpracování snímků videa. Konkrétně bude potřeba programově zajistit stabilní snímací podmínky, optimalizovat snímané video pro další zpracování, spustit existující program na zesílení barevných fluktuací (nebo ho vytvořit znova), detektovat pulsování barev ve zpracované sekvenci a určit tepovou rychlosť. Vše se potom zabalí do GUI aplikace. Ideální výsledek práce je online zpracování videa, to je však poměrně ambiciózní.

Práce bud' využije hotovou implementaci původních autorů této techniky a vloží ji do hotové aplikace, nebo celý postup naprogramuje samostatně (s přihlédnutím na kontext aplikace) v libovolném programovacím jazyku.

Literatura: Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Fredo Durand and William T. Freeman. Eulerian Video Magnification for Revealing Subtle Changes in the World. In ACM Transactions on Graphics (Proc. SIGGRAPH 2012), 31, 4. 2012.

## **Klíčová slova**

Digitální zpracování obrazu, určení tepu z optického záznamu, Eulerian magnification



# Obsah

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

<b>Rejstřík</b>	<b>37</b>
<b>A An appendix</b>	<b>37</b>

## **Seznam tabulek**

- 6.1 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. V tomto měření ještě nebyl kladen důraz na výkon aplikace, takže nebyl kladen důraz na co nejmenší počet snímků pro určení. 32
- 6.2 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. 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. 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. 16
- 5.5 Frekvenční doména naměřených intenzit po filtrace. 17
- 5.6 Příklady různorodých záznamů pro ověření detekčních schopností postupu. 17
- 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. 19
- 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. 26

- 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 bufferu. 27
- 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. 28
- 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  $\sigma$  a výsledný tep  $bpm$ . I z velmi silně zašumělého videa, ilustrovaného pravým snímkem, se podařilo vypočítat tep spadající do korektního intervalu. 31

# 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 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.

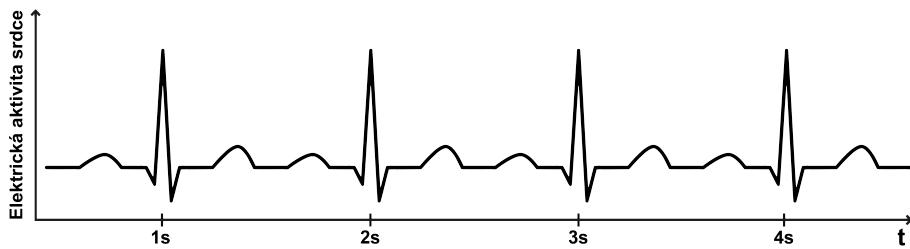
Srdeční frekvence nižší než 60 tepů se považuje za abnormálně pomalou a označuje se jako bradykardie. Hraniční hodnota se ovšem vždy vztahuje i k věkové a výkonnostní kategorii, např. u sportovců je běžné mít sníženou klidovou frekvenci i pod 60. Opakem bradykardie je tachykardie<sup>1</sup>, která je u dospělých obvykle diagnostikována nad 100 tepů/min (u dospělých).

---

1. Blíže viz [http://www.heart.org/HEARTORG/Conditions/Arrhythmia/AboutArrhythmia/Tachycardia-Fast-Heart-Rate\\_UCM\\_302018\\_Article.jsp](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 srdce<sup>2</sup>. V elektrokardiogramu odpovídají intervaly mezi jednotlivými lokálními minimy 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 prodleva bez interakce s uživatelem je po celou dobu 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*<sup>3</sup> 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 fluktuační 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  $\mathbf{p} = (x, y)$  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 rozměru 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. TEORIE

### 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:

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

kde  $i, 0 \leq i \leq K$ , je index snímku v poli  $V$ ,  $K$  délka vektoru  $V$  a  $f_{ps}$  je vzorkovací frekvence záznamu udávána v jednotce  $Hz$ .

## 2.2 Fourierova transformace

Pro analýzu frekvencí 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  $\omega_x$  nazývá *prostorová / časová frekvence*,  $x$  a  $\omega_x$  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  $\omega_x$  je závislá na všech hodnotách  $f(x)$  pro všechna  $x$ .

## 3 Aparáty pro realizaci práce

Lidské vidění má omezenou citlivost na prostorovo-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. propíchnutí pleti v kontextu vyšetřovacích metod. 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].

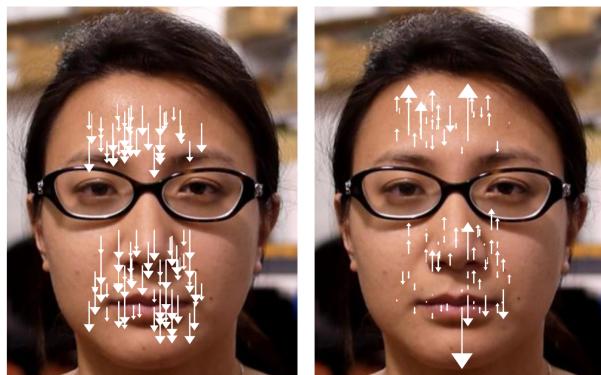
Balistokardiografie je neinvazivní metoda, která snímá pohyby částí těla způsobené činností kardiovaskulárního systému. Detekce tepu pomocí balistokardiografie spočívá v analýze pohybů obličeje s omezením na ty pohyby, které odpovídají přílivu krve při každém úderu srdce.

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

Velkou výhodou jsou malé, resp. žádné nároky na viditelnost kůže. Např. v publikaci je dosaženo téměř stejných výsledků i v případě, že

### 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].

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 pletysmografu<sup>1</sup>.

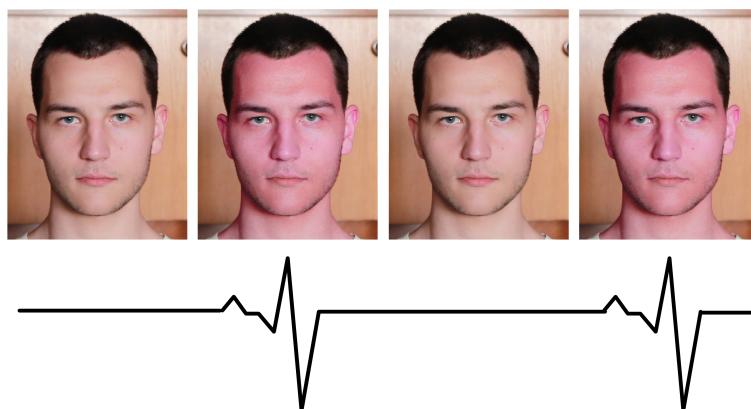
Z této metody vychází fotopletysmografie (PPG). Tato disciplína byla představena v roce 1930 a využívá světelnovou odrazivost a propustnost krve. Vychází z principu, že krev absorbuje více světla než okolní tkáň [6].

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.

Čím větší plochu ve snímcích záznamu zabírá pokožka, tím se dají změny detektovat přesněji. Nejstabilnější využití tohoto přístupu se zdá

---

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 odpovídat se snímky, ve kterých je pokožka odlišně zabarvená. Tyto barevné změny jsou okem nepostřehnutelné.

být v případě, že pokožka zabírá celou plochu záznamu<sup>2</sup>. Aplikace se proto bude snažit najít ve snímcích záznamu obličeje vhodnou masku pokožky, kterou bude dále zpracovávat.

Pro detekci tepu je vybrána tato metodu, 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 *Eulerian Video Magnification* [2]. Cílem tohoto postupu je odhalovat změny v optických záznamech, které je pouhým okem nemožné (nebo velmi těžce) 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 3.2 s dopředu známou tepovou frekvencí.

Dostupný je implementovaný algoritmus<sup>3</sup> v platformě MATLAB<sup>4</sup>. 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, který musí splňovat požadavky na minimálním rozlišení, vzorkovací frekvenci a kvalitu (z pohledu šumu).

Po stanovení cílů aplikace je pro implementaci využit kompilovaný jazyk C++, především díky své rychlosti. Pro operace zpracování obrazu bude použita knihovna OpenCV<sup>1</sup>. Jednoduché uživatelské rozhraní je implementováno s pomocí frameworku Qt<sup>2</sup>.

Pro nároky aplikace je nutné vícevláknové programování, které citelně sníží časové nároky na zpracování. Pro tyto účely se nabízí část knihovny Boost<sup>3</sup> pro C++.

Vizualizace dat (generování grafů) bude provedeno v MATLABu. Ostatní použité technologie 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 v části hodnocení

- 
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/>>

#### **4. POUŽITÉ VÝVOJOVÉ A VÝZKUMNÉ NÁSTROJE**

---

dle naměřených hodnot budou diskutovány požadavky na kvalitu a rozlišení vstupního videa.

## 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 byla nahrazena měřením pomocí tlakoměru. Zvolen byl běžně dostupný tlakoměr, který při měření tlaku měří i tep. U běžně dostupných tlakoměrů se udává chyba měření tepu  $\pm 5\%$ .

### 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.

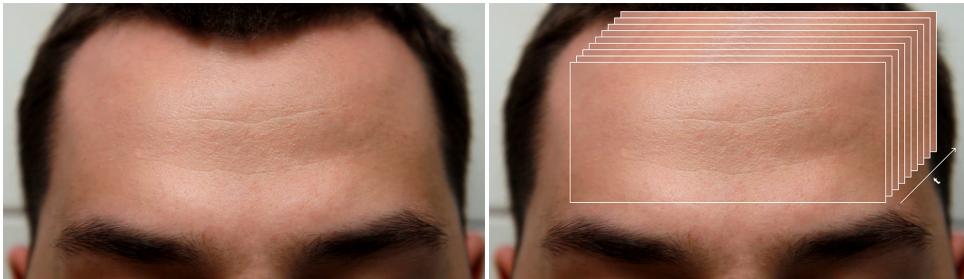
Do vektoru s video záznamem  $V(i)$  budou vloženy obrazy, kde  $i$ -tá hodnota pole bude charakterizovat  $i$ -tý snímek ze záznamu. Máme více možností na co se v každém snímku zaměřit:

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

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

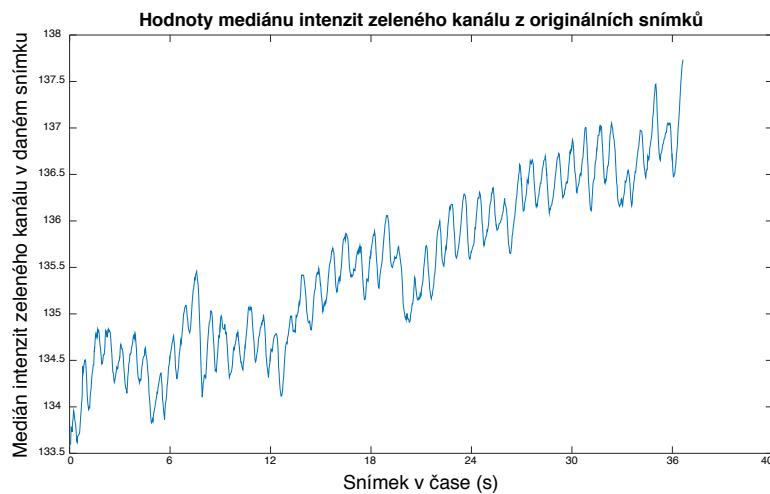
V tomto případě jsou rozdíly mezi různými přístupy minimální. Zkoumaný záznam je téměř ideální pro extrakci informací. Kůže je stabilizovaná (snímaný má opřenou hlavu o zem) a zabírá většinu záběru, záznam je ve vysokém rozlišení (1920x1080) a ve vzorkovací frekvenci 30 snímků/s. Dále budou rozdíly ilustrovány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

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

kde  $K$  je délka pole.

Nejdříve je pole převedeno pomocí Fourierovy transformace do frekvenční domény. Pro omezení frekvencí na ty, které jsou pro tento účel informačně přínosné, využijeme dříve zmíněné obvyklé rozmezí 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.

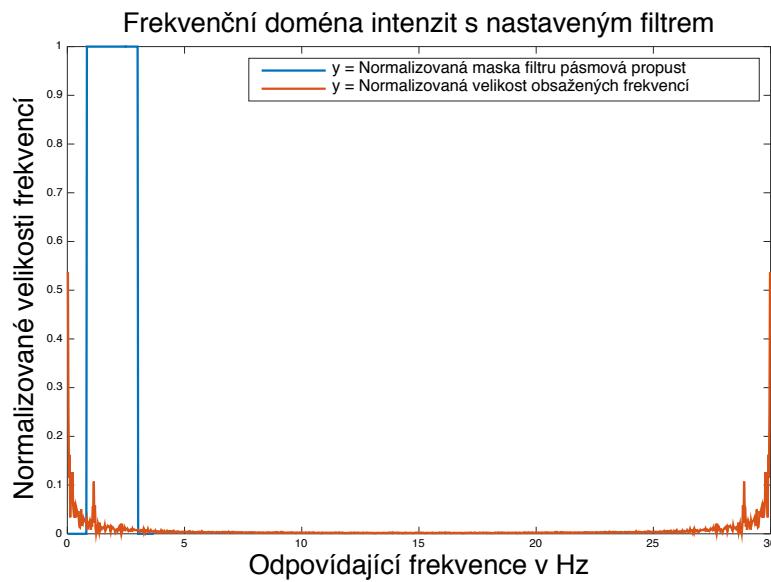
pásmovej propusť ve frekvenční doméně. Jako rozsah tohoto filtru jsou nastaveny hodnoty 55 resp. 180 tepů/min, které jsou převedeny do jednotky Hz, tedy 55/60 Hz a 180/60 Hz. V grafu na Obrázku 5.3 je zobrazená reálná složka frekvenční domény dohromady s filtrem pásmovej propusť. Na ose x jsou označeny 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 záběru. Jak pomalé, tak příliš rychlé změny byly filtrem odstraněny. Nyní po filtrace zbyly už jen frekvence, které potenciálně mohou odpovídat tepu. Nejsilnější z nich bude potom odpovídat výslednému tepu.

Na grafu na Obrázku 5.5 frekvenční domény po filtrace pásmovou propustí, je zřetelně vidět nejsilnější obsažená frekvence 0.8696 Hz (po přepočtení 69 tepů/min). Skutečná hodnota tepu pro tento záznam byla 67 tepů/min, při udávané chybě měření 5% jsou tedy hodnoty

## 5. POSTUP

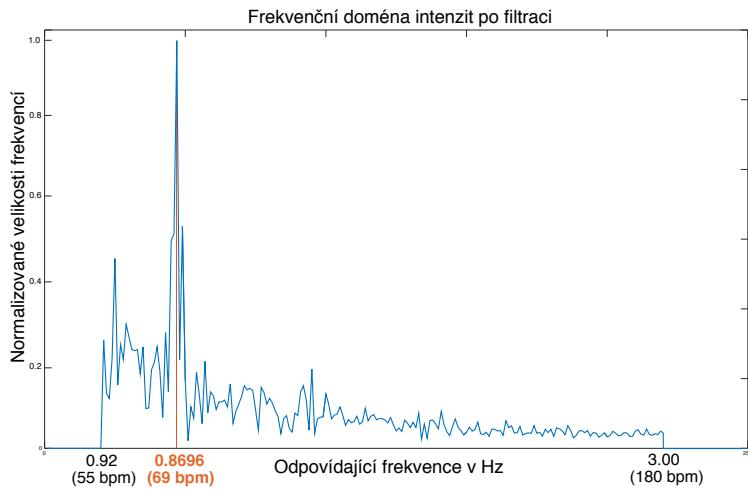
---



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 filtrace.



Obrázek 5.5: Frekvenční doména naměřených intenzit po filtraci.



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.

## 5. POSTUP

---

### Iniciální ověření na záznamech

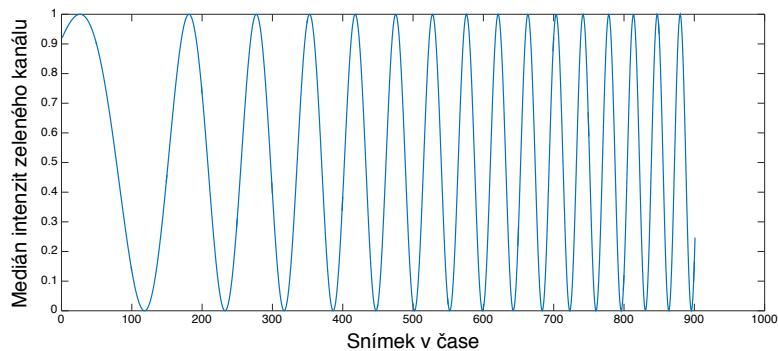
Mimo předchozího záznamu bylo potřeba pro iniciální ověření analyzovat různorodá videa. Jako netradiční případ se nabízí tep dítěte, který je podstatně vyšší než tep dospělého. Další atypický záznam může být obličej snědého člověka. Jako poslední záznam pro iniciální testování je nasnímaná ruka na podlaze. 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

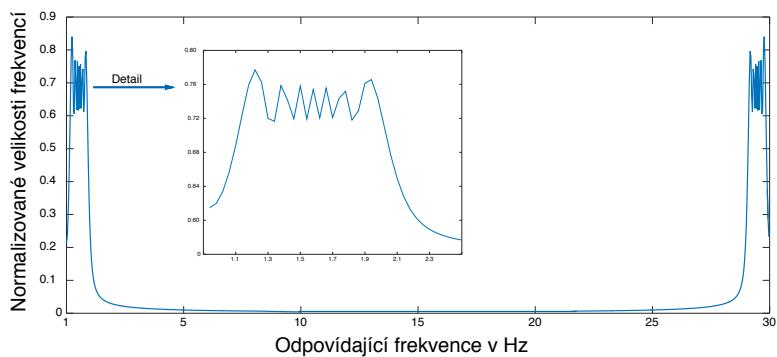
Filtrování ve frekvenční doméně má nesmírnou výhodu v podobě zaměření se na potřebné frekvence. Analýza frekvencí je imunní vůči globálním změnám světla. Globální změna světla je změna intenzit všech pixelů najednou v závislosti na čase — důležité je, že tato změna se musí týkat všech pixelů v obrazu. 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 mohel 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.

Vzhledem k tomu, že se práce zaměřuje na klidovou tepovou frekvenci, tak tento problém nebude vadit. Z testovacích videí vyplynulo,



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



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

## 5. POSTUP

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 implementace využívá datový typ *Mat*. Tento typ podporuje více typů, takže je využitelný pro šedotónové i barevné obrazy. Také umožňuje reprezentaci hodnot v datovém typu *float*.

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

## 6.2 Popis aplikace

Popis aplikace odpovídá použití jednoho z *reálných* režimů. 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. Samotná časová náročnost tohoto cyklu přímo ovlivňuje plynulost promítaného videa. V případě dodaného záznamu se jedná pouze o *trhání*, které se projevuje jen v uživatelském rozhraní, neovlivňuje tedy následné zpracování a výpočty. Při použití webkamery se ale jedná přímo o změnu výsledné vzorkovací frekvence, protože další snímek ze vsutpnáho proudu webkamery se vždy načte až po uplynutí jednoho cyklu.

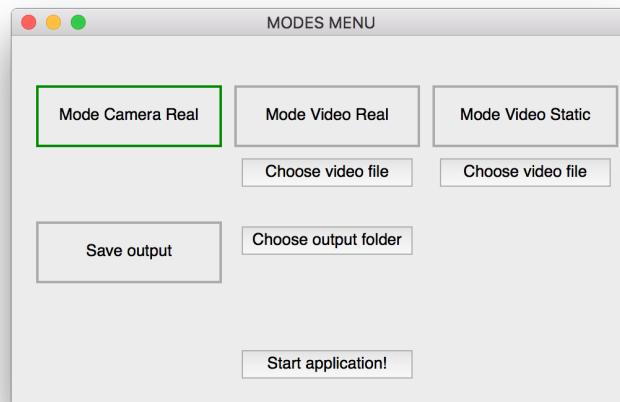
V každé iteraci cyklu je originální vstupní obraz vložen do bufferu<sup>1</sup>, který bude zpracován až najednou po naplnění.

---

1. Za buffer je považován záznam dle standardní definice 2.3.

## 6. IMPLEMENTACE

---



Obrázek 6.1: Uživatelské rozhraní pro volbu režimu aplikace.

### Vzorkovací frekvence

Už samotné zobrazování snímků z proudu webkamery je náročné, protože se jedná o bezzávratový formát bez komprese. Pravidelná maximální vzorkovací frekvence (podporovaná webkamerou) v prostředí OpenCV není zaručena ani při pouhém zobrazování v cyklu bez jakýchkoli úprav. Bohužel není k dispozici ani žádná funkce, která by vrátila hodnotu skutečné frekvence. Z tohoto důvodu aplikace pro určení využívá alternativní přístup, kdy počítá dobu každé iterace cyklu a aktualizuje aktuální frekvenci. Pro výpočet je využita knihovna *chrono*<sup>2</sup>, která nabízí funkce pro měření uplynulého času a je přenositelná napříč různými operačními systémy. Narození od některých ostatních funkcí měří opravdu uplynulý čas a ne jenom čas spotřebovaný procesorem. Přesnost 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 do bufferu se ukládají v plné velikosti). Konkrétně potom  $2^n$  krát - bez použití interpolace,

---

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

která je příliš náročná, aby mohla být prováděna v hlavním cyklu. Konstanta  $n$  je potom zvolena tak, aby šířka, respektive výška v případě vstupu orientovaného na výšku, co nejvíce odpovídala zadefinované hodnotě  $480px^3$ .

### Inicializace

Po startu nejdříve probíhá proces inicializace, kdy s prvními 20 snímků 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 levnějších (web)kamer, často také začíná záznam stisknutím spouště a s tím spojené záhvěvy scény.

### Detekce

Pro fungování algoritmu pro určení tepu je třeba izolovat dostatečně velkou oblast, ve kterou bude vyplňovat pokožka jak bylo popsáno v kapitole 3.1. V obličeji se jako největší plocha celistvé kůže nabízí čelo. Proto je třeba provádět detekci obličeje. Pro lepší určení pozice čela je vhodné použít taky detekci očí, potom se pozice čela určuje podstatně přesněji než jen ze znalosti pozice obličeje.

Pro detekci obličeje a očí využívá práce algoritmus *Haar Feature-based* kaskádové klasifikátory. Tuto metodu představili Paul Viola a Michael Jones v roce 2001. Detekce je založena na porovnávání jasových rozdílů pomocí různých konvolučních jader. Nejdříve je potřeba klasifikátor *trénovat* pomocí strojového učení z velkého množství vzorků (pozitivních i negativních). OpenCV pro použití nabízí již předtrénované klasifikátory. Mimo stabilitu 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 zabírá největší plochu, což 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.

---

3. Na použité hardwarové konfiguraci při použití načteného záznamu zaručuje plynulý obraz.

## 6. IMPLEMENTACE

---

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

Potom, co je detektován obličeji 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.

### 6.2.2 Výpočetní třída

Buffer 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. Velikost buffer je přímo určena potřebným počtem hodnot pro určení tepu,

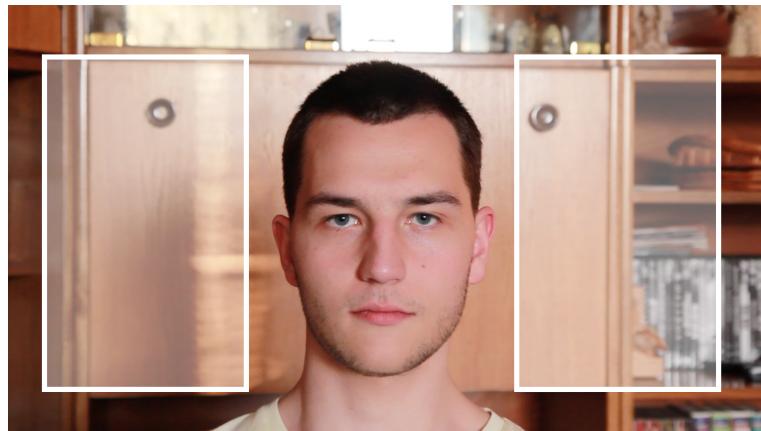
### 6.2.3 Inicializace

Při inicializaci třídy *BpmVideoProcessor* se provádí důležitá část určení plochy čela. Do třídy předáváme pozici a velikost již detektovaného obličeje z uživatelského rozhraní. Detekce očí v některých případech nefunguje 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 posunutím středu mezi očima přibližně o jednu desetinu výšky obličeje nahoru.

Pozice čela je korektní, je-li rozdíl mezi horizontálním středem obličeje a středem určeného čela menší než  $0.15 \times$  šířka obličeje, neboli:

$$abs\left(\frac{x_{face} + w_{face}}{2} - \frac{x_{forehead} + w_{forehead}}{2}\right) \leq 0.15 * w_{face}. \quad (6.1)$$



Obrázek 6.2: Vhodná místa pro analýzu globálních světelných změn.

### 6.2.4 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 tedy imunní vůči filtrování ve frekvenční doméně. Pro potlačení je tentokrát třeba detektovat prostor mimo obličej, který je po celou dobu záznamu pozičně fixovaný a projevují se v něm jen globální změny. Vhodná místa jsou zobrazena na Obrázku 6.2.

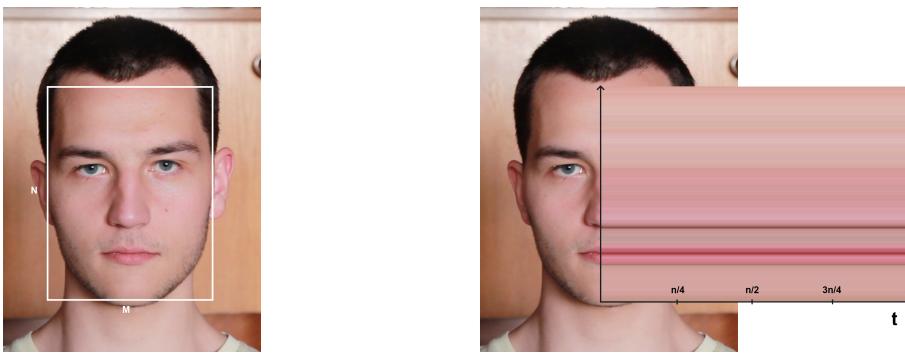
V prostoru je vytvořeno pole s charakteristikou intenzit snímků shodným způsobem jako v prostoru čela. Vhodným odečtením polí od sebe a následným normalizováním by se mely globální frekvence potlačit a kýžené frekvence odpovídající tepu naopak zvýraznit. Během měření a výsledků ovšem nebyly zaznamenány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

Další - nakonec nevyužitý pokus - spočíval ve stabilizaci záznamu pomocí sledování optického toku<sup>4</sup>. Před analýzou intenzit byl nejdřív záznam detekovaného čela stabilizován. Zpřesnění výsledků nebylo

---

4. Stabilizace byla inspirována algoritmem dostupným z <http://nghiaho.com/?p=2093>.

## 6. IMPLEMENTACE



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

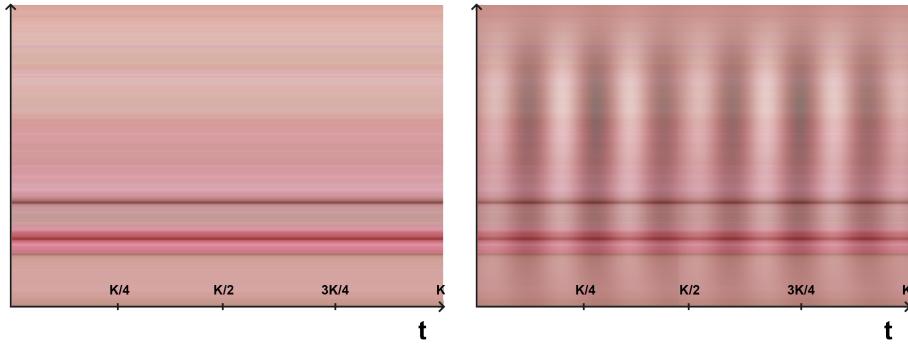
zaznamenáno, naopak v některých případech, došlo ke zhoršení. K nařušení výsledků došlo pravděpodobně kvůli affinním transformacím prováděných pro každý snímek během stabilizace videa a s nimi spojenými interpolacemi.

### 6.2.5 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 *Eulerian magnification*. 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 2.3). Vytvoření jednoho snímku v časové řadě je ilustrováno v Obrázku 6.3.

Pro každý snímek je sestrojena Gaussova pyramida<sup>5</sup>. Na jednotlivé úrovně v pyramidě je aplikován filtr pásmová propust. Jako střed intervalu propusti je nastaven již vypočítaný tep a rozsah intervalu je nastaven na 10 tepů/min (nastaveno podle subjektivního dojmu

5. Gaussova je pyramida, která tvoří obraz ve více stupních rozlišených. Každá úroveň je podvzorkována na polovinu rozlišení úrovně o stupeň vyšší a následně je rozostřena Gaussovým filtrem.



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 bufferu.

nejlepšího výsledku). Aby bylo možné algoritmus použít v kontextu této práce, bylo nutné jej přeimplementovat, tak aby bylo možné každou úroveň v pyramidě zpracovávat ve zvláštním vlákně. Po tomto rozdělení trvá filtrování Gaussovy pyramidy stejnou dobu, jak kdyby byla zpracovávána jen nejvyšší úroveň.

Filtrováním snímků časoprostorové řady je dosaženo potlačení frekvencí, které neodpovídají tepové frekvenci (podobně jako při určování tepu). Po filtrace vždy odpovídající snímek normalizujeme. Gaussova pyramida je nejdříve rekonstruována na časoprostorovou řadu, která je posléze převedena zpět do standardního záznamu. Ten nazveme jako *vizualizace tepu*.

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. Při dalším průchodu hlavního cyklu je použit další zpracovaný snímek.

V bufferu je po celou dobu maximální počet snímků, při vložení snímků nad kapacitu je nejstarší snímek odstraněn. Po dokončení prvního výpočtu aplikace ihned předává buffer s novými snímkami

## 6. IMPLEMENTACE

---



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.

k zpracování. Při takovémto opakovaném zpracování je možné detektovat i pomalu se měnící tep během snímání. Takto aplikace funguje dokud ji uživatel neukončí.

### 6.2.6 Rozdíl mezi režimy

Fungování *reálného* režimu při použití webkamery webkamery a při načteném externím záznamu je velmi podobné. 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 mezi *reálným* a *statickým* režimem je především ve sběru snímků do bufferu. Zatímco v *reálném* režimu je zobrazen obraz spolu s ukládáním snímků do bufferu, ve *statickém* režimu jsou snímky vloženy do bufferu aniž by bylo původní video viditelné. Pro rychlejší vkládání dat je program vyvíjen ve verzi OpenCV s podporou software *FFmpeg*<sup>6</sup>.

---

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

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 nastaví 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 vstup přetáčet. Tím by byla by porušena snaha 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 10 rozdílných záznamech v rozmezí tepu 58 – 82 tepů/min. Pro testovací účely není nutné testovat záznamy se stejnými reálnými tepy, protože výsledky vycházejí stejně. 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 6.3 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 6.3 je přístup pro tvorbu charakteristických informací o intenzitách v každém snímku (popsáno v kapitole 5.2). Měření opravdu ukázalo, že se barevné 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. Buffer by samozřejmě mohl obsahovat i více snímků, ale pro rychlosť zpracování je vhodné použít co nejnižší počet snímků.

## 6. IMPLEMENTACE

---

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 frekvenci jako dvojnásobek tepové frekvence snímané osoby, neboli:

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

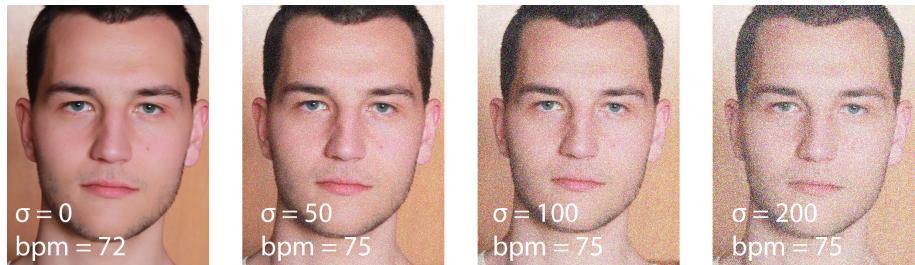
kde  $bpm_{real}$  je reálná tepová frekvence. Např. při tepu 60 tepů/min (1 Hz) je potřeba nasnímat záznam minimální vzorkovací frekvencí 2 snímky/s (2 Hz). Toto omezení platí pro samotné určení tepu, nicméně pro uživatele je video už příliš trhané.

Určení a vizualizace tepu funguje dobře při denním osvětlení. Během měření zobrazeném na Obrázku 6.6 se ukázalo, že postup pro určení tepu není příliš náchylný na šum. Při použití určitého umělého osvětlení na bázi zářivek algoritmus nepracuje příliš dobře i při dostatečném nasvícení. Dá se říci, že je pro aplikaci vhodnější zašumělé video za denního světla než záznam s kvalitním zářivkovým nasvícením.

Největším problémem pro postup je, když se během záznamu obličej příliš hýbá. Takováto videa zpravidla vedou k nekorektním výsledkům. Bohužel nepomáhá ani stabilizace obrazu před výpočty jak bylo diskutováno v Kapitole 6.2.4.

Každé testovací video je k dispozici v příloze v repozitáři ve složce **testing**. Pro každé testovací video je vytvořena jedna složka, ve které se nachází: originální video, výstupní video s vizualizací tepu a zobrazenou hodnotou detekovaného tepu, vygenerovaný snímek ze vstupního video, detekovaný obličej, detekované čelo:

- originální video,
- výsledné video s vypočítaným tepem a jeho vizualizací,
- snímek detekovaného obličeje a čela,
- textový soubor s průměrem intenzit zeleného kanálu pro každý snímek použity pro detekci tepu,
- ukázkou jak ovlivňuje použitá maska čela výsledek,



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  $\sigma$  a výsledný tep  $bpm$ . I z velmi silně zašumělého videa, ilustrovaného pravým snímkem, se podařilo vypočítat tep spadající do korektního intervalu.

- porovnání výsledků při použití různých přístupů z Kapitoly 5.2 pro určení tepu v závislosti na počtu použitých snímků v tabulce ve formátu .csv.

## 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 6.1: 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 hlavici. V tomto měření ještě nebyl kladen důraz na výkon aplikace, takže nebyl kladen důraz na co nejmenší počet snímků pro určení.

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 6.2: 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.

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.



## 7 Závěr

Podařilo se vytvořit funkční aplikaci pro detekci a následnou vizualizaci tepu, která funguje ve více režimech. Samotný režim s využitím webkamery je velice náročný na hardwarové vybavení stroje, na kterém je používán, proto pravděpodobně nebude fungovat na starších sestavách. Postup řešení byl obecně velmi náročný na výkon, proto bylo nutné hojně využívat vícevláknové programování pro ulehčení výpočtů. 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.



## **A An appendix**

Here you can insert the appendices of your thesis.