



# VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

## FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

## ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

## OBRAZOVÁ ANALÝZA MITOTICKÝCH CHROMOSOMŮ

DIGITAL IMAGE ANALYSIS OF MITOTIC CHROMOSOMES

### DIPLOMOVÁ PRÁCE

MASTER'S THESIS

### AUTOR PRÁCE

AUTHOR

Bc. Luboš Jaroš

### VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Helena Škutková, Ph.D.

BRNO 2017

# Diplomová práce

magisterský navazující studijní obor **Biomedicínské inženýrství a bioinformatika**

Ústav biomedicínského inženýrství

**Student:** Bc. Luboš Jaroš

**ID:** 155580

**Ročník:** 2

**Akademický rok:** 2016/17

**NÁZEV TÉMATU:**

## Obrazová analýza mitotických chromosomů

**POKYNY PRO VYPRACOVÁNÍ:**

1) Proveďte literární rešerši cytogenetických metod pro identifikaci chromosomů z mikroskopických snímků a zhodnoťte jejich výhody a nevýhody pro analýzu počtu a struktury chromosomů. 2) Seznamte se s technikami zpracování obrazových snímků cytogenetické analýzy chromosomů pro určení karyotypu organismů a vyhodnocení změn v počtu a struktuře chromosomů způsobujících genetická onemocnění. 3) Navrhněte a realizujte algoritmus pro automatické vyhodnocení karyotypu člověka ze snímků mitotických chromosomů. 4) Algoritmus rozšiřte o vyhodnocení strukturálních změn chromosomů oproti typickému lidskému karyotypu a vytvoření chromosomových ideogramů. 5) Vytvořte programové rozhraní v Matlabu pro určení lidského karyotypu, chromosomových ideogramů a genetických změn. 6) Program otestujte na volně dostupných cytogenetických snímcích lidských chromosomů a proveďte diskuzi výsledků.

**DOPORUČENÁ LITERATURA:**

[1] GAJENDRAN, V., RODRI, X a J. J. GUEZ. Chromosome counting via digital image analysis. In Image Processing, 2004. ICIP '04. 2004 International Conference on. 2004, vol. 5, p. 2929-2932 Vol. 2925.

[2] PIPER, JIM a ERIK GRANUM. On fully automatic feature measurement for banded chromosome classification. Cytometry, 1989, 10(3), 242-255

**Termín zadání:** 6.2.2017

**Termín odevzdání:** 19.5.2017

**Vedoucí práce:** Ing. Helena Škutková, Ph.D.

**Konzultant:**

**prof. Ing. Ivo Provazník, Ph.D.**  
předseda oborové rady

**UPOZORNĚNÍ:**

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

## **ABSTRAKT**

Rozvoj moderní medicíny umožňuje studovat lidský genom a odhalovat predispozice pro odlišné nemoci. Jednou z takových technik je analýza lidského karyotypu, tedy souboru všech chromosomů. Nejdůležitějším krokem při analýze karyotypu je detekce chromosomů a jejich kategorizace. V práci jsem navrhl nový algoritmus pro detekci a kategorizaci chromosomů do sedmi skupin z mikroskopického snímku vzorku DNA. Implementace algoritmu byla provedena v prostředí Matlab. Pro ověření přesnosti segmentace a klasifikace chromosomů byl algoritmus testován na snímcích ze dvou databází (117 a 38 snímků). Sensitivita segmentace činí 88% a pozitivní prediktivní hodnota segmentace je 92%. Úspěšnost párování chromosomů je 77%.

## **KLÍČOVÁ SLOVA**

Chromosom, lidský karyotyp, cytogenetika, barvicí metody chromosomů, předzpracování obrazu, analýza obrazu

## **ABSTRACT**

The development in modern medicine has allowed to study human genome and detect predispositions to several diseases. One of very promising techniques is the analysis of human karyotype, i.e., the number and appearance of chromosomes in the cell nucleus. The most important step in the karyotype analysis is the chromosome detection and categorization. In this work, a new algorithm for detection of chromosomes from an image of microscopic DNA sample and their categorization into seven groups was developed. The algorithm was implemented in Matlab. The accuracy of segmentation and classification was tested on a set of images from two databases with 117 and 38 images, respectively. The sensitivity of the developed segmentation reached 88% while the value of positive predictivity of segmentation reached 92%. The success rate of chromosome pairing achieves 77%.

## **KEYWORDS**

Chromosome, human karyotype, cytogenetic, chromosome staining methods, image preprocessing, image analysis

JAROŠ, L. *Obrazová analýza mitotických chromosomů*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2017. 57 s. Vedoucí diplomové práce Ing. Helena Škutková, Ph.D..

## PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma *Obrazová analýza mitotických chromosomů* jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

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

V Brně dne .....

.....

(podpis autora)

## PODĚKOVÁNÍ

Děkuji vedoucímu diplomové práce Ing. Heleně Škutkové, Ph.D. za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé diplomové práce.

V Brně dne .....

.....

(podpis autora)

# OBSAH

<b>Úvod</b>	<b>8</b>
<b>1 Cytogenetika</b>	<b>9</b>
1.1 Chromosom.....	9
1.2 Lidský karyotyp .....	11
1.3 Chromosomové aberace.....	12
<b>2 Identifikace a vizualizace chromosomů</b>	<b>15</b>
2.1 Pruhovací barvicí metody .....	15
2.2 Molekulární metody.....	19
<b>3 Zpracování obrazových snímků</b>	<b>21</b>
3.1 Předzpracování obrazu.....	21
3.2 Detekce chromosomů .....	22
3.3 Separace chromosomů .....	24
3.4 Klasifikační znaky .....	27
<b>4 Realizace obrazové analýzy</b>	<b>32</b>
4.1 Předzpracování obrazu.....	32
4.2 Segmentace chromosomů od pozadí.....	34
4.3 Separace překryvů u chromosomů.....	36
4.4 Separace dotyků u chromosomů.....	39
4.5 Klasifikace chromosomů do karyotypu .....	41
4.6 Grafické uživatelské rozhraní .....	47
<b>5 Diskuze a výsledky</b>	<b>48</b>
5.1 Vyhodnocení a diskuze segmentace obrazu .....	49
5.2 Vyhodnocení a diskuze klasifikace obrazu.....	50
<b>6 Závěr</b>	<b>52</b>

<b>Literatura</b>	<b>54</b>
<b>Zdroje obrázků</b>	<b>57</b>
<b>Přílohy</b>	<b>I</b>

# ÚVOD

Rozvoj moderní medicíny umožňuje studovat lidský genom a odhalovat predispozice pro široké spektrum nemocí. Chromosomální vyšetření dnes patří mezi nedílnou součást genetických diagnostických vyšetření. Cílem chromosomálního vyšetření je sestavit lidský karyotyp. Z lidského karyotypu je možné určit strukturní nebo numerické aberace, které predikují genetické onemocnění. Chromosomy je možné pozorovat v optickém nebo fluorescenčním mikroskopu a dále třídit dle klasifikačních znaků.

Pro pracovníky v genetických laboratořích, kteří lidský karyotyp sestavují většinou manuálně, je tato činnost časově velmi náročná. Sestavení karyotypu je tedy nejen časově náročné, ale i drahé. Proto vzniká poptávka pro automatizovaných systémech pro detekci a klasifikaci obrazů mitotických chromosomů. Návrh těchto systémů vyžaduje specifické znalosti z oblasti digitálního zpracování signálu, bioinformatiky a výpočetní techniky.

Cílem práce je navrhnout nový algoritmus a implementovat ho v prostředí Matlab. Před samotnou implementací je provedena rešerše zabývající se cytogenetickými metodami a základními metodami barvení chromosomů. Důraz je kladen na možnosti zpracování lidského karyotypu a následné diagnostiky genetických onemocnění. Práce rovněž shrnuje algoritmy zabývající se segmentací chromosomů.

Navržený algoritmus by měl být schopen automaticky segmentovat a klasifikovat chromosomy mitotických chromosomů. Algoritmus by neměl být časově náročný a měl by umět sestavit lidský karyotyp. Dále je nutné vyhodnotit strukturní a numerické aberace. Navržený algoritmus je primárně zaměřen na často používané techniky barvení Q a G pruhování. Kvalita algoritmu je vyhodnocena na snímcích z veřejně dostupných databází.



# 1 CYTOGENETIKA

Cytogenetika je vědní obor, který se zabývá studiem buněčných struktur nesoucích genetickou informaci. Cytogenetickými metodami se zjišťuje počet, tvar a struktura chromosomu, a hledají se příčiny a následky změn na chromosomech. Základní vyšetřovací metodou cytogenetiky je chromosomová analýza. Výsledkem této analýzy je karyotyp, který označuje uspořádanou sadu chromosomů dané buňky. Při studiu jsou chromosomy pozorovány ve světelném mikroskopu v době, kdy dochází k dělení buňky (metafáze). Vyšetření se nejčastěji využívá pro stanovení karyotypu s vrozenými vadami v prenatalní diagnostice nebo při asistované reprodukci. [1], [2]

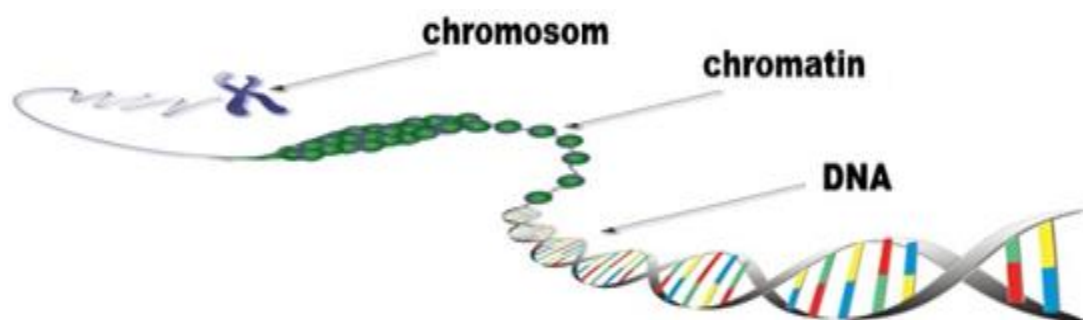
První historické zmínky o cytogenetice jsou již z roku 1888, kdy W. Waldeyer zavedl pojem chromosom. V první polovině 20. století byly studovány karyotypy nejrůznějších živočišných druhů a daly tak vznik prvním obsáhlým přehledům. Tyto cytogenetické výsledky však často nebyly správné díky technické nedokonalosti metodických postupů. Až v roce 1956 vědci Tjio a Levan vyvinuli efektivní metodiku analýzy chromosomů a stanovili počet lidských chromosomů na počet 46. Na konci 70. let došlo k dalšímu metodickému objevu, který spočíval v zavedení diferenciálního barvení chromosomů. Tyto metody umožňovaly podrobnější analýzu chromosomů a určení karyotypu. Zatím k poslední velké metodické inovaci došlo ke konci 80. let, kdy se cytogenetika převážně dostala na molekulární úroveň. Jedná se o techniku hybridizace a využití imunochemických metod značení sond pomocí fluorochromů. [2], [3]

## 1.1 Chromosom

Chromosomy jsou specifické barvitelné struktury, které lze pozorovat pod světelným mikroskopem při dělení v jádru buňky. Chromosomy jsou složeny z DNA, RNA a histonů. Zaručují rovnoměrné rozdělení genetické informace do dceřiných buněk.

Z chemického hlediska jsou chromosomy složeny z histonových bílkovin, které tvoří kostru chromosomu. Histonové bílkoviny se účastní replikací DNA, ochrany DNA, regulace replikace aj. Na tuto kostru se dále namotávají molekuly DNA (1,6-8,2 cm). Tento komplex DNA a histonů se nazývá chromatin. Struktura chromatinu se skládá z nukleozómu a solenoidu. Základní strukturou je nukleozóm, který je tvořen histonovými molekulami obtočených kolem DNA. Vyšší struktura Solenoid je tvořena spirálovitým uspořádáním nukleozómů. Solenoid je uspořádán do pravidelných smyček

kolem matrice a tvoří základní segment chromosomu. Chemickou strukturu chromosomu znázorňuje obrázek 1. [4],[5]



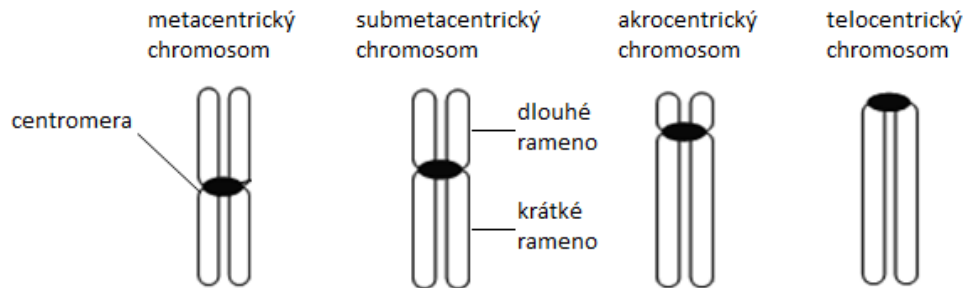
Obrázek 1: Struktura chromosomu [1]

Chromosom je pentlicovitý útvar zpravidla tvořený dvěma rameny, mezi nimiž je ztenčená oblast nazývaná centromera. Koncová část ramen se potom označuje jako telomera. Tato ramena nebývají stejně dlouhá, proto u chromosomu rozlišujeme krátké (p rameno) a dlouhé rameno (q rameno). Ramena se dále dělí na oblasti a pruhy. Pro určení konkrétního místa v karyotypu je použit čtyřmístný kód, kde první znak určuje chromosom, druhý rameno, třetí oblast a čtvrtý pruh. Například kód 2q12 označuje druhý chromosom, dlouhé rameno, oblast jedna a pruh dva. Při pozorování chromosomu v metafázi je chromosom tvořen dvěma chromatidami, které jsou spojeny v oblasti centromery, protože již došlo k replikaci DNA. Chromosomy se dělí podle umístění centromery a délky ramen do čtyř skupin, které jsou zobrazeny na obrázku 2:

- Metacentrický chromosom – centromera je umístěna přibližně uprostřed chromosomu a dělí chromatidy na přibližně dvě stejně dlouhá ramena. Chromosom připomíná tvarově písmeno X.
- Submetacentrický chromosom – centromera je umístěna směrem k telomerám a chromatidy jsou rozděleny na dlouhé a krátké rameno.
- Akrocentrický chromosom – centromera dělí chromatidy na jedno extrémně krátké rameno a jedno extrémně dlouhé rameno.
- Telocentrický chromosom – centromera je umístěna v oblasti telomer a připomíná tvarově písmeno V. Na rozdíl od předešlých chromosomů se tento typ chromosomu v karyotypu člověka nevyskytuje.

Běžné buňky v lidském organismu jsou diploidní, což znamená, že všechny chromosomy v organismu jsou v párech. Každý chromosom se tedy v buňce vyskytuje dvakrát. Výjimku v lidském organismu typicky tvoří pohlavní buňky, které jsou haploidní (obsahují pouze jednu sadu). U autozomů (nepohlavní chromosomy) se u obou párů

nachází genetická informace, která dává vznik stejným znakům. U heterozómů (pohlavní chromosomy) se u párů nachází jiné genetické informace, které dávají vznik odlišným znakům. [4],[5]



Obrázek 2: Typy chromosomu podle polohy centromery [2]

## 1.2 Lidský karyotyp

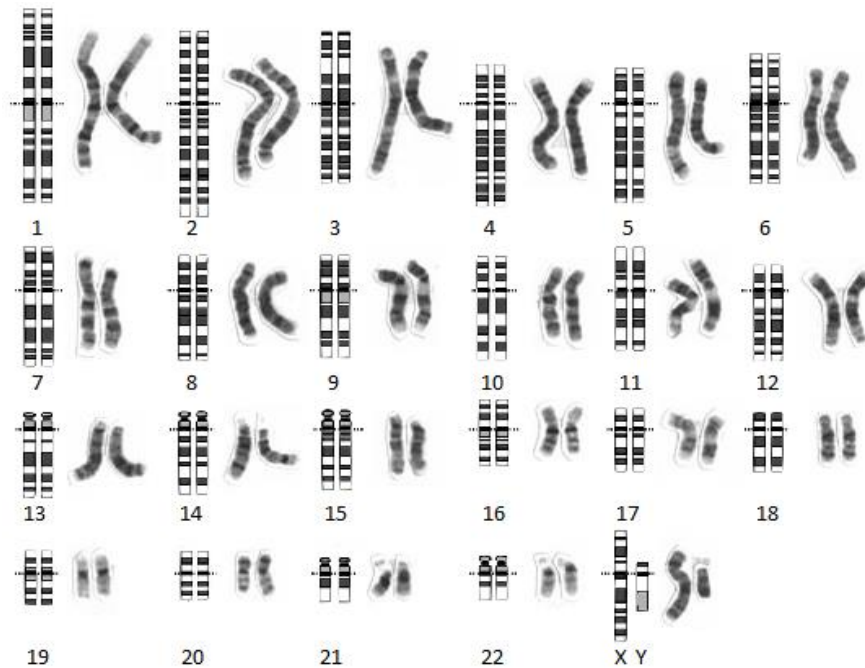
Normální lidský karyotyp se skládá z dvaceti tří párů chromosomů, celkem tedy ze 46 chromosomů. Z těchto 23 párů chromosomů je 22 párů tvořeno autozomy a tvoří homologní páry. Poslední pár je heterogenní a tvoří pohlavní chromosomy. V lidském karyotypu se u mužského pohlaví vyskytují chromosomy XY. Jedná se o heterogametické pohlaví. U ženského pohlaví se vyskytují chromosomy XX (homogametické pohlaví). Důležité je si uvědomit, že muž má pouze jeden chromosom X oproti ženě, která má dva. Z tohoto důvodu je náchylnější na některá genetická onemocnění oproti ženě, protože nemá potenciálně zdravou kopii.

Vyšetření karyotypu u dospělého člověka nebo dítěte je relativně snadné, stačí odebrat krev. Komplikovanější vyšetření ovšem nastává při vyšetření plodu matky, kdy musí být použita některá z invazivních metod odběru buněčného materiálu.

Obraz z vyšetření zdravého ženského karyotypu, který obsahuje 46 chromosomů, je možno vidět na obrázku 3. V pravé části u jednotlivých obrázků se nachází chromosomy seřazené do karyotypu. V levé části se nachází ideogramy daných chromosomů (jedná se o standardizaci chromosomu s přesně danými oblastmi a pruhy). Chromosomy v karyotypu třídíme podle morfologických znaků jako je velikost, typ chromosomu a pozice centromery do sedmi skupin [5], [6], [7], [8]:

- Skupina A – velké metacentrické chromosomy – 1, 2, 3
- Skupina B – velké submetacentrické chromosomy – 4, 5
- Skupina C – střední submetacentrické chromosomy – 6, 7, 8, 9, 10, 11, 12, X

- Skupina D – střední akrocentrické chromosomy – 13, 14, 15
- Skupina E – malé submetacentrické chromosomy – 16, 17, 18
- Skupina F – malé metacentrické chromosomy – 19, 20
- Skupina G – malé akrocentrické chromosomy – 21, 22, Y



Obrázek 3: Zdravý mužský karyotyp s ideogramy [3]

### 1.3 Chromosomové aberace

Pod pojmem chromosomové aberace si můžeme představit mutace, které mění jejich tvar a strukturu. Některé chromosomové mutace se u nositele neprojeví, jiné jsou neslučitelné se životem.

Chromosomové aberace mohou být buď vrozené nebo získané. U vrozených chromosomálních aberacích jsou mutace přítomné již od počátku vývoje jedince. Tyto aberace jsou přítomny ve všech buňkách v těle a přenášejí se dále na potomky. Typicky se jedná o inverze nebo balancované translokace. Získané chromosomové aberace vznikají až během života jedince působením odlišných vnějších vlivů na jedince nebo jako důsledek poruchy reparačních mechanismů. V somatických buňkách je charakteristickým poškozením zlom v chromosomu a následné chybné spojení. Může dojít i ke ztrátě části chromosomu. Tyto mutace mohou vést k aktivaci onkogenů a mohou být příčinou vzniku řady nemocí, jako například diabetes,

ateroskleróza a nádorové onemocnění. Dalším důsledkem může být zrychlené stárnutí buněk a tkání, které snižují kvalitu života postiženého. V gametických buňkách se získané aberace chromosomů projeví sníženou plodností jedinců. Gametické aberace nepříznivě ovlivňují vývoj zárodku poruchami metabolismu a zvýšenou pravděpodobností k určitým onemocněním (nádorové onemocnění, hypertenzní nemoc, diabetes). [2], [9], [10], [12]

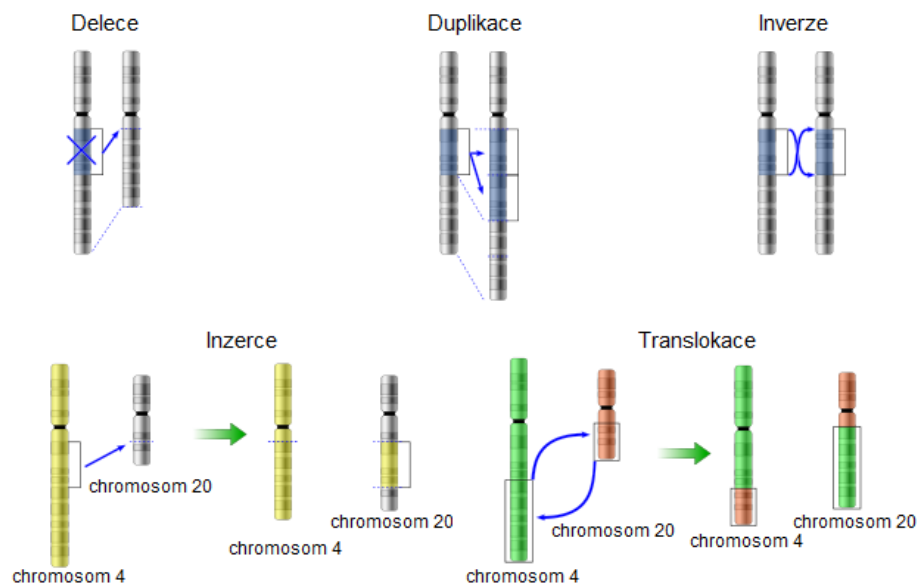
Chromosomové aberace dále rozdělujeme podle morfologie, na numerické aberace a strukturní aberace. Numerické aberace jsou způsobeny abnormálním počtem chromosomů v karyotypu, tedy jiný počet než čtyřicet šest. Aberace je způsobena chybou při rozdělení dceřiných buněk během buněčného dělení nebo poruchou fertilizace (oplozením vajíčka dvěma spermii). Chromosomální struktura je neporušena, ale obsahuje nestandardní množství genetického materiálu. Numerické aberace dále dělíme:

- Polyploidie – jedná se o zmnožení celé chromosomální sady. Lidský karyotyp normálně obsahuje 46 chromosomů ( $2n$ ), u triploide je v karyotypu obsaženo 69 chromosomů ( $4n$ ) a u tetraploidie dokonce 92 chromosomů ( $4n$ ). Polyploidie není slučitelná se životem.
- Aneuploidie – znamená, že v karyotypu chybí nebo přebývá nějaký chromosom. Konkrétní chromosom může být buď znásoben nebo ztracen. Při násobení se jedná o trisomii (tři kopie chromosomu) a tetrasomii (čtyři kopie chromosomu). Při ztrátě mluvíme o monosomii (jedna kopie chromosomu) nebo nuliosomii (žádná kopie chromosomu). Tyto aberace již jsou slučitelné se životem, ale nositel má zpravidla nějakou tělesnou vývojovou vadu nebo sníženou inteligenci a plodnost. Existuje celá řada klasických syndromů způsobené touto aberací. Mezi nejznámější patří Downův syndrom (trisomie 21. lidského chromosomu), Edwardsův syndrom (trisomie 18. chromosomu), Patauův syndrom (trisomie 13. chromosomu) a Turnerův syndrom (žena má jeden chromosom X).

Strukturní aberace jsou méně časté než numerické. Aberace jsou způsobeny chromosomovými zlomy, které jsou následkem chromosomální nestability. V klinické medicíně nemají velký význam, ale jsou spojeny s nádorovým onemocněním. Chromosomy jsou charakterizovány poruchou struktury, přičemž množství genetického materiálu může, ale i nemusí být zachováno. Pokud není zachováno, dochází ke změně znaků jedince. Strukturní aberace dělíme na balancované a nebalancované. U balancovaných strukturních aberací je zachováno původní množství genetického materiálu (může přejít na potomka a stát se nebalancovaným). Nebalancované aberace nemají původní množství genetického materiálu, buď chybí nebo přebývá. Podle charakteru změny ve struktuře chromosomu rozlišujeme základní druhy poruch. Mezi

nejčastější patří delece, duplikace, inverze, inserce a translokace.

Delece je způsobena zlomem chromosomu, kdy dochází ke ztrátě úseku chromosomu. Pokud dochází k jednomu zlomu, tak dojde ke ztrátě terminálního konce. Pokud ke dvěma, dochází ke ztrátě segmentu uloženého mezi centromerou a terminálním koncem. Delece způsobuje nebalancovaný karyotyp. Duplikace se projevuje znásobením segmentu chromosomu. Bývají méně nebezpečné než delece. Tento druh poruchy způsobuje také nebalancovaný karyotyp. U inverze dojde na jenom chromosomu ke dvou zlomům. Vzniklý segment se otočí o  $180^\circ$  a zpět se začlení do chromosomu. Pokud jsou zlomy na stejném raménku a úsek nezahrnuje centromeru, jedná se o paracentrickou inverzi. Pokud je na každém raménku jeden zlom a invertovaný úsek zahrnuje centromeru mluvíme o pericentrické inverzi. Inverze patří mezi balancované aberace a obvykle se neprojeví ve fenotypu (soubor všech pozorovatelných vlastností a znaku živého organismu). U inserce je z jednoho chromosomu úsek odstraněn a vložen do jiného chromosomu buď v původním směru nebo invertovaně. Ke vzniku této aberace jsou potřeba minimálně 3 zlomy, dva na jednom chromosomu a jeden na druhém. Inserce je poměrně vzácná a jedná se o nebalancovanou aberaci. Translokace je záměna dvou úseků dvou chromosomů. Může dojít jak k balancované, tak k nebalancované translokaci. Pokud je záměna dvou úseků mezi dvěma homologními chromosomy, tak vzniká reciproká translokace. Zde zůstává počet chromosomů stejný. Pokud ale dojde k fúzi dvou ramen akrocentrického chromosomu, tak jedinec s touto translokací má o chromosom méně a translokaci nazýváme robertsonskou. Na obrázku 4 jsou ilustrovány již popsané strukturní aberace. [10], [11], [12], [13]



Obrázek 4: Příklady strukturních chromosomových aberací [4]

## 2 IDENTIFIKACE A VIZUALIZACE CHROMOSOMŮ

Prvním krokem k získání lidského karyotypu je příprava chromosomů k mikroskopickému pozorování. Chromosomy, jak již bylo zmíněno v předchozí kapitole, jsou získány z dělících se buněk z periferní krve. Nejčastěji jsou použity lymfocyty. Buňky jsou dále kultivovány v mediu a následně zastaveny při dělení v metafázi mitózy. K zastavení dělení je použit mitotický jed kolchicin. Do suspenze je dále přidán hypotonický roztok z důvodu sedimentace a zvětšení chromosomů. Posledním krokem je fixace suspenze na mikroskopickém sklíčku a získání obrazu chromosomů.

Chromosomy jsou pro lepší vizuální reprezentaci barveny. Používají se pruhové barvicí metody nebo molekulární metody, které jsou popsány dále v této kapitole. Příprava a barvení chromosomů je velmi důležitým krokem. Při špatném provedení může dojít ke ztrátě informací v obraze. Při dlouhém působením hypotonického roztoku může dojít k rozpadu chromosomů. Pokud bude hypotonický roztok působit na chromosomy krátce, může dojít ke špatnému rozprostření chromosomů na mikroskopickém sklíčku. Pokud k tomuto dojde, postrádá další analýza smysl a je nutné vyšetření podstoupit znovu. Důležitým faktorem při analýze je i snímač, kterým je pořízen obraz chromosomů. Při nízkém rozlišení není možné správně identifikovat jednotlivé chromosomy do karyotypu. [12], [14]

### 2.1 Pruhovací barvicí metody

Chromosomové preparáty barvíme základními pruhovacími technikami z důvodu jejich zvýraznění. Při barvení se na chromosomech vytváří pruhy dané šířky a intenzity. Podle postupu barvení chromosomů rozlišujeme odlišná rozložení pruhů. Pruhovací metody byly zavedeny v letech 1968 až 1971. Do té doby byly používány jenom konvenční metody klasického barvení.

Klasicky obarvené chromosomy jsou homogenně obarvené barvou bez jakéhokoliv pruhování. Takto obarvené chromosomy můžeme rozdělit do několika skupin podle délky a polohy centromery. U této metody barvení je možné určit počet a hrubou stavbu chromosomů. Nalezenou aberaci není možné určit podrobněji, lze určit pouze zda je v dané buňce přítomna. Příklad klasického barvení je zobrazen na obrázku 5.



Obrázek 5: Snímek obarvených chromosomů pomocí klasického barvení [5]

### **G pruhování**

G pruhování je nejčastěji používané barvení chromosomů. Důvodem je rychlost a cena při dostatečném zachování informace. G pruhování je založeno na působení trypsinu a následném obarvení chromosomu Geimsovým roztokem. Počet pruhů je dán stupněm kondenzace chromosomu. Tmavě zbarvené proužky jsou tvořeny úseky DNA s převládajícím adeninem a thyminem. Světlé proužky odpovídají DNA s převládajícím guaninem a cytosinem. Každý pruh má své číslo, které slouží k popisu strukturních změn a lokalizaci genů a jiných sekvencí DNA. Na rozdíl od klasického barvení lze odhalit strukturní i numerické abnormality. Obrázek 6 zobrazuje snímek chromosomů barvené pomocí G pruhování.

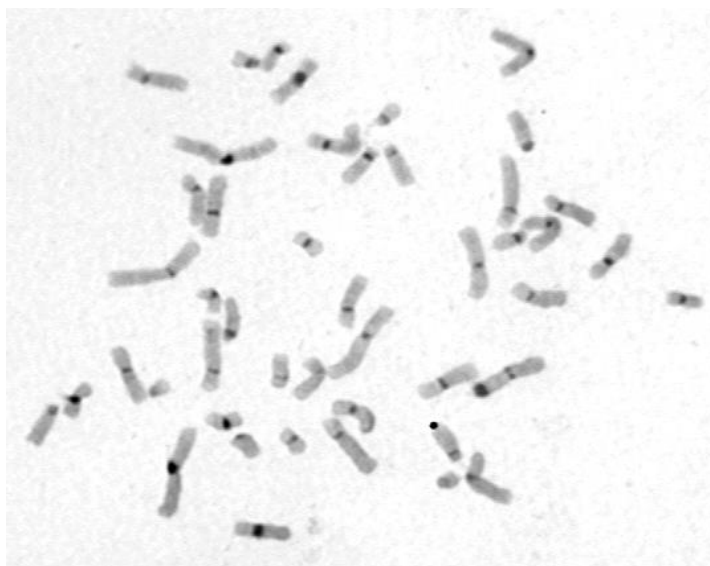


Obrázek 6: Snímek obarvených chromosomů pomocí G pruhování [6]



## C pruhování

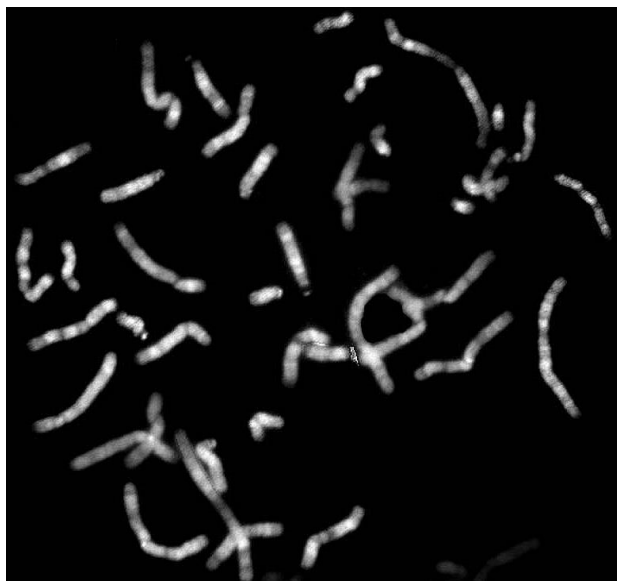
C pruhování docílíme použitím nasyceného hydroxidu barnatého při vyšší teplotě a následným barvením Giemsovým barvivem. Pruhy vznikají v centromerických oblastech všech autozomů, X chromosomu a v distální části dlouhých ramen Y chromosomu, což je možné vidět na obrázku 7. Tyto oblasti jsou bohaté na repetitivní sekvence DNA. Nejvýraznější bloky centromerického chromatinu jsou na chromosomech 1, 9 a 16. Toto barvení pomáhá sledovat chromosomové polymorfismy postihující okolí centromery. Snímek získaný tímto pruhováním slouží jako doplňkové vyšetření. Ze snímku lze sice určit karyotyp, ale není možné určit strukturní abnormality.



Obrázek 7: Snímek obarvených chromosomů pomocí C pruhování [6]

## Q pruhování

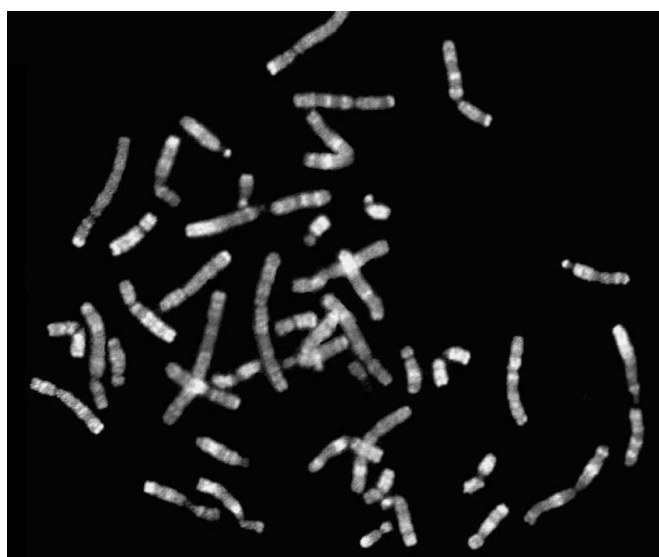
Při Q pruhování jsou chromosomy barveny ve vodném roztoku chinakrinu. Používají se i fluorescenční barviva akridinová oranž nebo 33258 Hoechst. Tyto barviva vytváří na chromosomech pruhy podobné G pruhování, ale snímek lze vidět pouze ve fluorescenčním mikroskopu, který má filtry odpovídající excitačním a emitovaným vlnovým délkám. Obrázek 8 zobrazuje snímek z fluorescenčního mikroskopu, kde je použito Q pruhování. Jedna z nevýhod je, že není trvalé a s časem se snižuje intenzita fluorescenčního záření. Tudíž je proces barvení časově náročnější. Další nevýhodou je, že pro snímání je potřebný fluorescenční mikroskop, tudíž je celý proces i podstatně dražší než podobné G pruhování. I přes tyto zápory se toto barvení v praxi používá. Důvodem je šetrné zachování morfologie chromosomů. Snímek je tedy vhodný nejen pro sestavení karyotypu ale i pro zjištění strukturních abnormalit.



Obrázek 8: Snímek obarvených chromosomů pomocí Q pruhování [7]

### **R pruhování**

Při R pruhování jsou na chromosomy při vyšších teplotách ještě před barvením aplikovány fosfátové pufrů. Působením pufrů dochází k denaturaci převážně adeninu a guaninu. Následně jsou preparáty chromosomů barveny Geimsovým barvivem. Tímto způsobem jsou získány reverzně obarvené pruhy (proto R pruhování), které jsou přesným opakem G pruhů a Q pruhů. Tedy tam, kde jsou G pruhy a Q pruhy světlé, tam jsou R pruhy tmavé a opačně. R se pruhy vytváří v oblastech bohatých na chromatin, a tak lze této barvicí metody využít při detekci malých strukturních abnormalit. Obrázek 9 znázorňuje chromosomový snímek barvený R pruhováním. [15], [16], [17], [18]



Obrázek 9: Snímek obarvených chromosomů pomocí R pruhování [7]

## 2.2 Molekulární metody

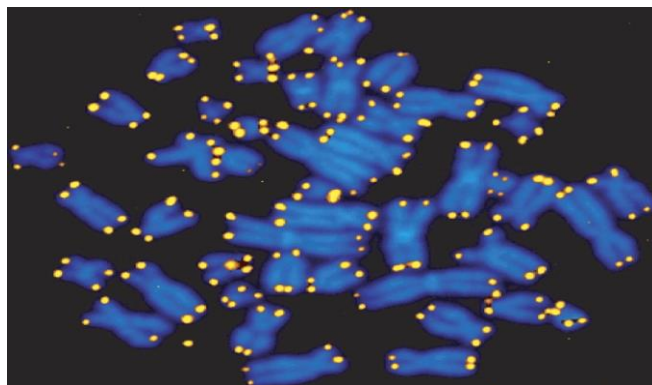
S rozvojem molekulární biologie přichází nové techniky vizualizace chromosomů. Metody molekulární cytogenetiky slouží k detekci submikroskopických chromosomálních aberací, které by nebyly zachytitelné v pruhovacích metodách.

V molekulární cytogenetice se jedná především o hybridizaci in situ. To je metoda, která využívá lokalizaci a identifikaci specifického úseku DNA pomocí denaturace a následné reasociace DNA. Denaturace DNA je dosaženo zahříváním molekul DNA na teplotu 94 °C až 96 °C, kdy zanikají vodíkové můstky a vlákna DNA se tak oddělí. K Hybridizaci dochází podle pravidel komplementarity, kde je spojeno vlákno DNA s druhým vláknem s označenou sondou. K hybridizaci dochází při teplotě přibližně 37°C. K Hybridizaci dochází v biologickém materiálu, tedy in situ. Sondy byly dříve značeny radioaktivně, ale v současné době se používají fluorescenčně značené sondy (Hybridizace in situ fluorescenční-FISH). Výsledky lze dále pomocí počítačových programů upravovat ve smyslu zesílení signálu a potlačení nespecifického pozadí.

Molekulární studium buněk v interfázi přináší nové informace o biologicky významných vrozených i získaných chromosomových abnormalitách. Základními metodami lze studovat poruchy funkce genu, určovat jejich roli v patogenezi u vrozených onemocnění, popřípadě určovat maligní buněčné transformace. Molekulární metody zvyšují citlivost cytogenetických metod o několik řádů. Mezi nejčastější používané molekulární metody patří FISH a komparativní genomová hybridizace (CGH).

### Fluorescenční in situ hybridizace (FISH)

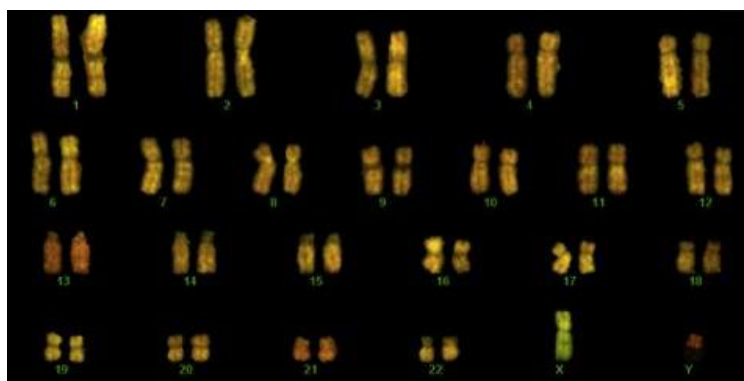
Tato metoda je založená na hybridizaci in situ, kde sonda využívá imunofluorescence. U přímo značených sond je detekce možná ve fluorescenčním mikroskopu. Platí, čím více jsou chromosomy kondenzované, tím vyšší je rozlišení, ale stává se méně přehledným. Pro barvení se nejčastěji používá propidium jodid (červeně zbarvené chromosomy ve fluorescenčním mikroskopu) nebo 4,6-diamidino 2-phenylindol (tzv. DAPI), který chromosomy barví modře. Sondy dělíme na centromerické, lokus specifické a telomerické. Pokud chceme například zjistit translokaci na chromosomu 16, tak použijeme sondu specifickou pro danou oblast chromosomu a centromerické sondy pro chromosom 16. Tím zjistíme, zda jsou všechny alely na svých místech. Metoda FISH se převážně využívá při studiu karyotypu buněk v metafázi a v interfázi, při určování početních a strukturních chromosomových odchylek a při detekci reziduálních chorob. Obrázek 10 zobrazuje snímek získaný kvalitativní FISH metodou s použitím peptidových sond. Telomery jsou zobrazeny žlutě, zatímco DNA v chromosomech je modrá.



Obrázek 10: Značení telomer v snímku FISH metodou [8]

### Komparativní genomová hybridizace (CGH)

Komparativní genomová hybridizace (CGH) umožňuje mapovat a detekovat relativní počet kopií odlišných genomů. Metoda je založena na porovnávání dvou rozdílných DNA úseků. Jednou barvou značíme standartní chromosomální výbavu a druhou barvou vyšetřovanou. Za normálních podmínek jsou oba vzorky zbarveny stejně. Oblasti zmnožení nebo ztráty DNA sekvencí, jako jsou delece, duplikace nebo amplifikace jsou zvýrazněny rozdílnou barvou homologů (párových chromosomů). CGH nelze využít u jednoduchých translokací, inverzí a inzercí, při kterých se nemění poměr počtu kopií sekvencí DNA. Obrázek 11 znázorňuje výsledek metody CGH, kde snímek obsahuje červeně a zeleně barvené chromosomy. Pokud je chromosom zbarvený do zelena značí větší počet chromosomů než v referenčním vzorku. Pokud je chromosom zbarvený do červena značí menší počet chromosomů než v referenčním vzorku. Chromosomy 13 a 21 jsou červené barvy, což značí monosomii. Chromosom X je zelený, tudíž obsahuje více chromosomů. Chromosom Y je červený z čehož plyne, že obsahuje méně chromosomů než v referenčním vzorku. Zkoumaný snímek tedy znázorňuje ženský karyotyp (referenční snímek obsahuje mužský karyotyp). [19], [20], [21], [22]



Obrázek 11: Vyhodnocený snímek ženského karyotypu s monosomií na 13. a 21. chromosomu metodou CGH [9]

### 3 ZPRACOVÁNÍ OBRAZOVÝCH SNÍMKŮ

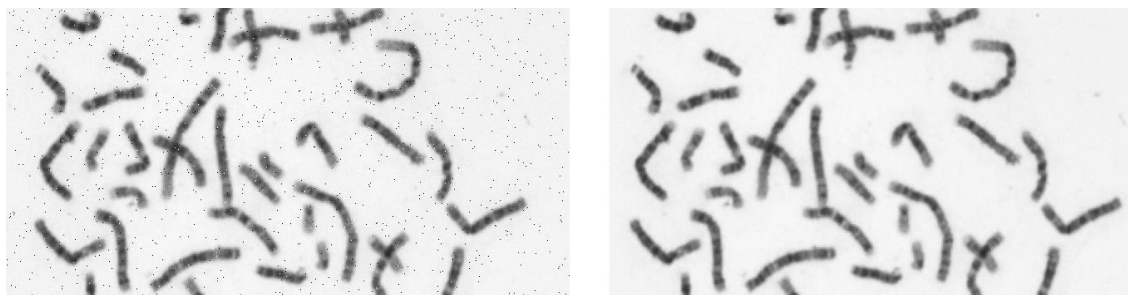
S vývojem cytogenetických metod se zvyšují i nároky na rychlost automatické analýzy cytogenetických obrazů. Analýzu mitotických chromosomů můžeme provádět na snímcích získaných základními vizualizačními technikami. Proto je nutné použít odlišný postup při detekci a klasifikaci snímků získaných pruhovacími metodami a snímků získaných molekulárními metodami.

Tato práce se zaměřuje na zpracování chromosomů získaných pruhovacími technikami barvení. Zpracování obrazových snímků se skládá z několika kroků. Nejdříve je nutné snímek mitotických chromosomů předzpracovat, aby detekce jednotlivých chromosomů byla co nejsnazší. Dalším krokem je detekce a separace chromosomů od pozadí. Protože se jedná o automatickou analýzu, je v této části vždy nutné chromosomy zkontrolovat, zda-li se vyseparovaly správně. Pokud by došlo k chybné separaci, byla by následná analýza bezcenná. Posledním krokem zpracování snímku mitotických chromosomů je klasifikace a určení, zda se jedná o zdravý karyotyp. Klasifikace je založena na získání klasifikačních znaků z jednotlivých chromosomů. [22], [23], [24]

#### 3.1 Předzpracování obrazu

Úkolem předzpracování obrazu je zvýšit výtěžnost informace z obrazu pro nejjednodušší identifikaci chromosomů a posléze znaků v jednotlivých chromosomech. Některé obrazy bývají zarušeny odlišnými druhy rušení, které je nutné před analýzou odstranit.

V prvním kroku se musí obraz převést z barevného na šedotónový, protože další analýza probíhá v šedotónovém obrazu. Po převedení nastává filtrace. Hojně využívaným filtrem je mediánový filtr. Filtr úspěšně eliminuje zkreslení známe jako „pepř a sůl“ (náhodné černé nebo bílé pixely v obraze), viz. obrázek 12. Nejčastěji používaná velikost masky mediánového filtru je 3 x 3.



Obrázek 12: Snímek s rušením „sůl a pepř“ (vlevo), snímek po mediánové filtraci (vpravo)

Další filtrace spočívají v lineárních nebo nelineárních transformacích jasových hodnot. Tyto transformace vedou k zvýraznění vyšších frekvencí v obraze, tedy k zvýraznění hran pro lepší detekci. Další možností úpravy jsou průměrovací filtry, které mají tendenci rozmazávat, a tedy znehodnocovat informaci v obraze. [25], [26]

## 3.2 Detekce chromosomů

V literatuře se vyskytuje řada segmentačních metod pro detekci chromosomů a separaci od pozadí. Jedná se o metody segmentace založené na detekci hranic, segmentaci rozvodím, segmentaci pomocí aktivních kontur nebo segmentace metodou Otsu. V práci budou blíže popsány metody Otsu a aktivních kontur. [27], [28], [29]

### Otsu metoda

Otsu metoda vychází z histogramu. Na obraz se díváme jako na dvě skupiny bodů s danými intenzitami a snažíme se najít takový práh, při kterém jsou v obou skupinách mezi pixely minimální rozptyly, zatímco mezi skupinami maximální rozptyly. Metoda se snaží minimalizovat chybu mezi pozadím a popředím. Optimálního prahu je docíleno pomocí snižování společných rozptylů obou skupin podle vzorců:

$$\sigma_b^2(T) = n_b(T)n_o(T)[\mu_b(T) - \mu_o(T)]^2 \quad (1)$$

$$n_b(T) = \sum_{i=0}^{T-1} p(i) \quad (2)$$

$$n_o(T) = \sum_{i=T}^{N-1} p(i) \quad (3)$$

$$\mu_b(T) = \sum_{i=0}^{T-1} \frac{ip(i)}{n_b(t)} \quad (4)$$

$$\mu_o(T) = \sum_{i=T}^{N-1} \frac{ip(i)}{n_o(t)}, \quad (5)$$

kde  $\sigma_b^2(T)$  je vzájemný rozptyl obou skupin,  $T$  je aktuální zkoumaná hodnota prahu a  $[0, N-1]$  je rozsah úrovní intenzity.  $p(i)$  je počet pixelů o dané intenzitě.

Pro každou uvažovanou hladinu je nutné rozdělit pixely podle hodnoty prahu do skupin. V každé skupině se počítá průměr. Rozdíl umocníme na druhou

a vynásobíme počtem pixelů v každé skupině. Určením maximální hodnoty  $\sigma_b^2(T)$  zjistíme i práh. Tento postup je nicméně výpočetně náročný a častěji se využívá výpočtu podle histogramu. Při zvyšování prahu dochází k přesunu pixelů z jedné skupiny do druhé skupiny. Stačí tedy upravovat jenom  $n$  a  $\mu$  podle vzorců:

$$n_b(T + 1) = n_b(T) + n_T \quad (6)$$

$$n_o(T + 1) = n_o(T) + n_T \quad (7)$$

$$\mu_b(T + 1) = \frac{\mu_b(T)n_b(T) + n_T T}{n_b(T + 1)} \quad (8)$$

$$\mu_o(T + 1) = \frac{\mu_o(T)n_o(T) + n_T T}{n_o(T + 1)}, \quad (9)$$

kde  $n_T$  je četnost pixelů o intenzitě rovné prahu.

### Aktivní kontury

Metoda aktivních kontur je založena na minimalizaci hodnoty energie, která přísluší dané křivce. Energie je složená s externí a interní energie. Externí energie je závislá na vlastnostech obrazu a lokální pozici bodu křivky. Interní energie je nezávislá na segmentovaném obraze a závisí jenom na tvaru a vlastnostech křivky. V praxi je možné aktivní kontury rozdělit podle výpočtu externí energie na metody s využitím gradientu a bez využití gradientu. Dále bude popsána metoda s využitím gradientu, tzv. metoda hada.

Metody využívající gradient jako parametr pro výpočet externí energie je výhodné použít v nezašuměném obraze s výraznými hranami. Metoda hada patří k nejstarším a nejjednodušším metodám aktivních kontur. Hada je z definice spilne, který se snaží o minimalizaci energie. Energie je dána externí silou, která má tendenci posouvat hada směrem k rostoucímu gradientu do místa hrany. Interní složka energie má snahu zachovat hladkost křivky. Celá energie křivky je zadána parametrem  $s$  a je dána rovnicí:

$$E_{snake} = \int_0^1 E_{int}(v(s)) + E_{ext}(v(s)) ds, \quad (10)$$

kde  $v(s)$  je parametricky zadaná křivka s normalizovaným parametrem  $s$ .  $v(s)$  je možné zapsat i v kartézských souřadnicích jako  $v(s) = \{x(s), y(s)\}$ .  $E_{int}$  je interní energie,  $E_{ext}$  je externí energie. Tyto energie je možné rozepsat jako:

$$E_{snake} = \int_0^1 E_{int}(v(s)) + E_{image}(v(s)) + E_{com}(v(s)) ds, \quad (11)$$

kde  $E_{image}$  reprezentuje externí energii v nezměněné podobě. Interní energie je rozdělena do  $E_{int}$ , která reprezentuje sílu pro zachování křivky a  $E_{com}$ , která vyjadřuje pružné síly křivky. Následně je potřebné rozepsat jednotlivé energie z důvodu explicitního výpočtu. První část interní energie lze rozepsat jako:

$$E_{int} = \frac{1}{2}(\alpha(s)|v_s(s)|^2 + \beta(s)|v_{ss}(s)|^2), \quad (12)$$

kde  $\alpha$  a  $\beta$  jsou zadané empirické váhy.  $v_s$  a  $v_{ss}$  jsou první a druhou derivací křivky podle parametru  $s$ . Energie obrazu je dána součtem energií z dat z obrazu, které je možné rozepsat rovnicí:

$$E_{image} = w_{line}E_{line} + w_{edge}E_{edge} + w_{term}E_{term}, \quad (13)$$

kde  $w$  odpovídá empiricky zadaným vahám, které specifikují daný typ problému. Dílčí energie mohou být vypočítány daným způsobem.  $E_{line}$  můžeme nahradit v nejjednodušší formě intenzitou obrazu,  $E_{edge}$  se počítá pomocí první derivace v obraze a  $E_{term}$  je možné vypočítat pomocí směru gradientu rozmazaného objektu. Poslední potřebnou energii lze vypočítat na základě analogie s pružinou, kterou teoreticky představuje. Energie pružiny je dána rovnicí:

$$E_{com} = -k(s_1 - s_2)^2, \quad (14)$$

kde  $s_1$  a  $s_2$  jsou dva body křivky. Spojením všech výše zmíněných energií získáme iterativní řešení pro jednotlivé body křivky, které se budou posunovat směrem k hranám křivky. [28], [30], [31]

### 3.3 Separace chromosomů

Některé detekované chromosomy obsahují dva či více překrývajících se chromosomů. Dalším krokem při zpracování obrazu je nutné tyto shluky dvou či více chromosomů rozdělit. V této podkapitole budou popsány dvě vybrané metody, které se danou problematikou zabývají. První metoda využívá konkávních bodů v obryse chromosomu. Druhá metoda je založena na výpočtu jednochromosomového skóre. [32], [33], [34]

#### Rozdělení chromosomů pomocí konkávních a konvexních bodů

Rozdělení dotýkajících se chromosomů pomocí konkávních a konvexních bodů vychází z binárního obrazu získané pomocí prahování Otsu. Z binárního obrazu je následně



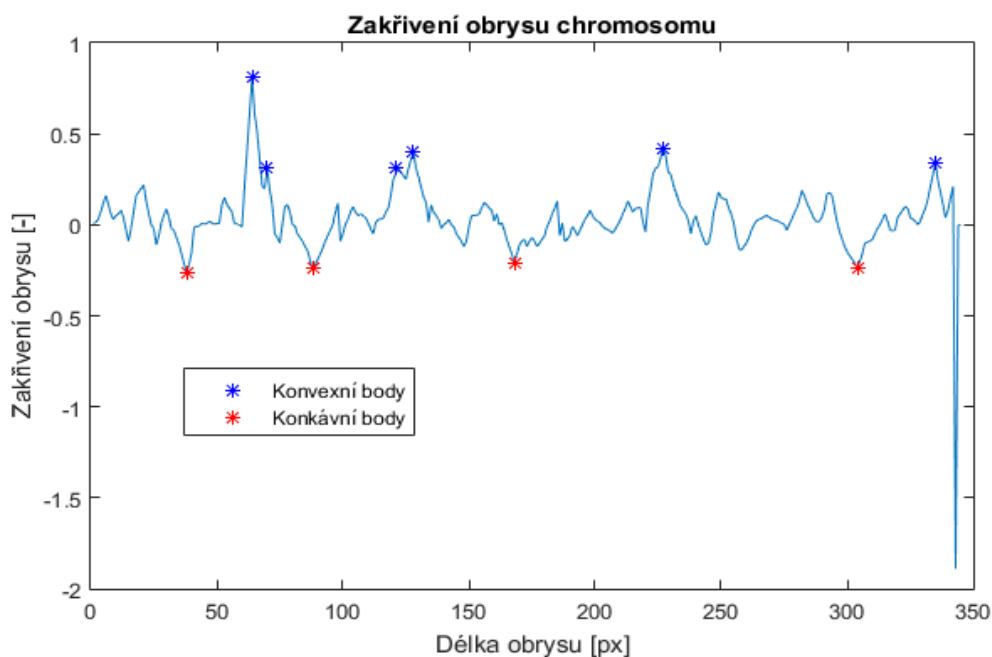
získán obrys chromosomu. Pro vyhlazení obrysu chromosomu je využito aproximace polygonem. Obrys chromosomu má tvrdé přechody mezi pixely, a proto je pro vyhlazení použito polygonu a mediánového filtru. Body zakřivení jsou poté počítány z funkce zakřivení v celé délce obrysu chromosomu. Body  $x$  a  $y$  jsou vyjádřeny parametrickou rovnicí s parametrem  $t$ :

$$x = x(t) \text{ a } y = y(t), \quad (14)$$

kde  $x$  a  $y$  jsou polohy bodu v obryse a  $t$  je pozice v obryse. Tyto body jsou následně postupně derivovány. Druhá derivace nám říká, zda první derivace klesá nebo stoupá. Bod, kde je druhá derivace větší než nula, je konvexní. Konkávnost nastává v bodě, kde druhá derivace menší než nula. Druhá derivace zakřivení obrysu je definována rovnicí:

$$C(l) = \frac{x'y'' - y'x''}{(x'^2 + y'^2)^{\frac{3}{2}}}, \quad (15)$$

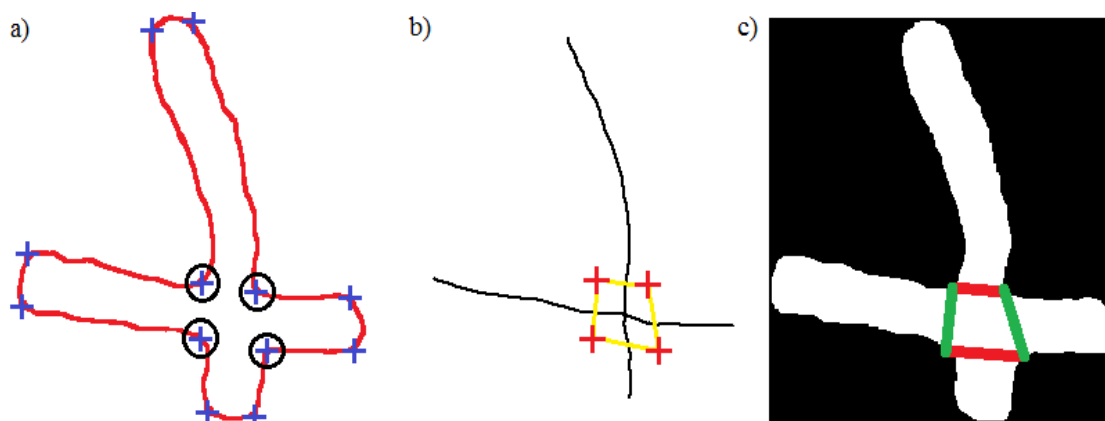
kde  $C$  je bod funkce zakřivení obrysu a  $l$  je pozice zakřivení souřadnic. Výsledná funkce zakřivení je vykreslena v grafu s označenými konkávními a konvexními body (obrázek 13). Konvexní body, označené v obrázku modře, určují konce chromosomu. Konkávní body označené v obraze červeně určují čtyři body, kde dochází k dotyku chromosomu (obrázek 14a). Tyto body pomáhají následnému rozdělení chromosomu.



Obrázek 13: Zakřivení obrysu chromosomu s konkávními a konvexními body [10]

Po získání čtyř konkávních bodů je vypočtena středová linie chromosomu. Pomocí středové linie jsou následně vždy spojeny dva a dva body, které jsou nejbližší ke středové linii. Tímto způsobem jsou získány čtyři fiktivní přímky, které slouží

k rozdělení chromosomu. Obrázek 14b znázorňuje středovou linii spolu s červeně vyznačenými konkávními body a žlutě vyznačenými fiktivními přímkami. Obrázek 14c znázorňuje binární obraz společně s fiktivními přímkami, které rozdělují překrývající chromosomy na dva samostatné chromosomy. [33], [39]



Obrázek 14: Rozdělení chromosomu pomocí konkávních bodů [10]

a) obrys chromosomu se zvýrazněnými konkávními a konvexními body, b) středová linie chromosomu s body s vyznačenými přímkami. c) binární obraz s vyznačenými přímkami rozdělení

### Rozdělení chromosomů na základě výpočtu SCS (Jednochromosomové skóre)

Výpočet jednochromosomového skóre slouží jako rozhodovací kritérium, zda se jedná o více chromosomů nebo o osamocený chromosom. Obraz je stejně jako v minulé metodě segmentován pomocí Otsu metody. Z binárních obrazů jsou následně vypočítány středové linie chromosomů. Na výpočet středových linií byl použit modifikovaný algoritmus původně určený pro snímky sítnice. Výchozí bod je umístěn do středu shluku a následně je odhadnut lokální průměr a směr chromosomu. Bod je poté posunut ve směru chromosomu a výpočet se iterativně opakuje, dokud není vypočtená celá středová linie ve všech možných směrech.

U každého shluku je vypočítáno skóre, které odhaduje tvar chromosomu. Skóre se počítá pomocí střední linie a průměru chromosomu. Postupným rozšiřováním oblastí okolo středové linie až do průměru chromosomu získáme odhad plochy shluku chromosomů. Jednochromosomové skóre je vypočteno pomocí rovnice:

$$SCM(C) = \frac{A(C)}{A(\hat{C})} \geq t_{SCM}, \quad (16)$$

kde  $A(C)$  je plocha segmentovaného shluku,  $A(\hat{C})$  je plocha shluku, která je odhadnutá pomocí středové linie a průměru,  $t_{SCM}$  je heuristicky zvolený práh na hodnotě 0,2. Pokud SCS překročí tuto hodnotu jedná se o shluk dvou a více chromosomů.

Jedná-li se o shluk, následuje hledání bodů překryvu nebo dotyku chromosomu. Hledání spočívá v nalezení lokálních minim v obrysu chromosomu vždy mezi dvěma body. Každý segment dvou chromosomů  $s_{ij} = [k_j - k_i]$  je definován počátečním bodem  $k_i$ , koncovým bodem  $k_j$ , a orientací  $v_{ij}$ . Vzdálenost mezi dvěma segmenty je definována rovnicí:

$$(s_{12}, s_{34}) = \frac{s_{12} \times s_{23}}{|s_{12}|} \quad (17)$$

Všechny body, jejichž vzdálenost je menší než průměr chromosomu a zároveň se jejich orientace liší méně než  $\pi/6$ , jsou body považovány za body dotyku. Poté jsou hledovými algoritmem hledány dvojice nejbližších dvou přímek, které spojují čtyři body a tvoří mnohoúhelník. Pokud jsou tyto dvě přímky rovnoběžné, tak jsou tyto čtyři body vyhodnoceny jako překryv a následnými fiktivními přímkami rozděleny. Pokud jsou nalezeny pouze dva body jedná se o dotyk dvou chromosomů. U dotyku jsou chromosomy rozděleny pomocí trackovacího algoritmu. Trackovací algoritmus hledá cestu mezi dvěma body s možnou nejnižší intenzitou a poté dojde k rozdělení. Rozdělení shluku čtyř chromosomů je možné vidět na obrázku 15. Na obrázku jsou barevně rozděleny dva překryvy a jeden překryv. [34], [35]



Obrázek 15: Barevně rozdělené chromosomy [11]

### 3.4 Klasifikační znaky

Klasifikační znaky jsou důležitými prvky sloužícími k automatickému vyhodnocení karyotypu. Mitotické chromosomy vykazují značnou tvarovou rozmanitost, proto je některé klasifikační znaky zjistit obtížnější než jiné. U některých klasifikačních znaků

může dojít ke nesprávnému určení, které vede ke špatnému vyhodnocení karyotypu. Klasifikační znaky můžeme dělit na geometrické znaky a znaky pruhovací. Mezi geometrické znaky patří obsah chromosomu, jeho délka nebo pozice centromery chromosomu. Pruhovací klasifikační znaky jsou typické pro pruhovací techniky a dávají nám informace o struktuře daného chromosomu. Mezi tento klasifikační znak patří jasový profil pruhů, vypočtený z kolmic středové linie.

Nejčastější kombinace klasifikačních znaků nám umožňují chromosomy zařadit do sedmi skupin podle velikosti, správně chromosom otočit, ale také napárovat homologické chromosomy, čímž se dá určit pohlaví, druh onemocnění aj.

### **Obsah chromosomu**

Nejsnáze získatelným klasifikačním znakem je obsah chromosomu. Obsah chromosomu se získá z binárního obrazu mitotického chromosomu a jako suma všech pixelů v dané kontuře v obrazu. Binární obraz, ze kterého se získá klasifikační znak obsah, je možné vidět na obrázku 16. Správné určení tohoto klasifikačního znaku závisí pouze na předchozí segmentaci chromosomů od pozadí. Pokud segmentace proběhne úspěšně, tak i výpočet tohoto klasifikačního znaku bude správný. Tento klasifikační znak pomáhá k rozdělení chromosomů podle velikosti.



Obrázek 16: Binární obraz pro výpočet obsahu chromosomu

### **Délka chromosomů**

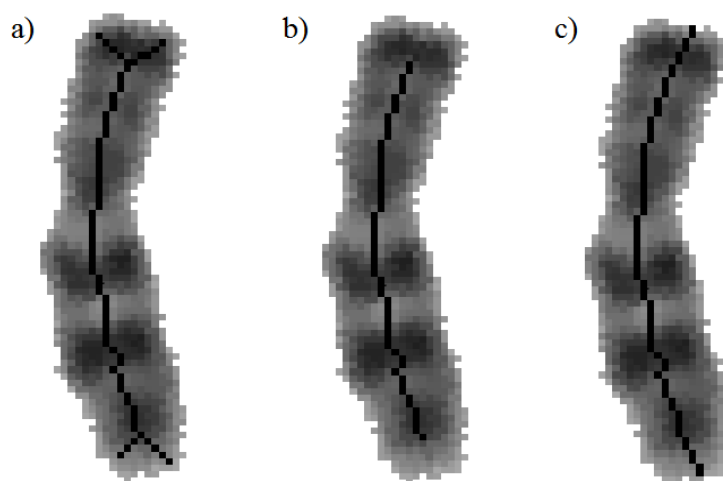
Druhým klasifikačním znakem, který lze z chromosomů získat je délka chromosomu. Výpočet délky chromosomu již není tak jednoduché, jako je zjištění obsahu chromosomu. Pro získání délky chromosomu je nutné vypočítat středovou linii chromosomu, protože chromosomy jsou ohnuté pod odlišnými úhly a detekce v jedné linii tedy není možná.

Středová linie chromosomů se počítá z binárního obrazu, který je znázorněn na obrázku 16. Existují další postupy pro získání středové linie. Jeden ze způsobů je iterativní ztenčování binárního obrazu. Operace ztenčování vede k odebrání pixelů z obrysu objektu. Operace je zastavena, jakmile se obraz nemění. Další možností, jak získat

středovou linií je aproximace výpočtu tzv. linií dohoření. Idea metody vychází z analogie s hašením požárů pomocí zakládání protipožárů (Jakmile nekontrolovaný požár dorazí k místu spáleném protipožárem, nemůže se dále šířit). Binární obraz je považován za tzv. travnaté pole, které je po okrajích (obrysu obrazu) zapáleno. Požár se šíří konstantní rychlostí ke středu objektu. V místě, kde se dva požáry v opačném směru setkají vzniká linie dohoření. Další metoda k získání středové linie spočívá ve výpočtu distanční mapy. Výpočet distanční matice je vytvořen na základě vzdáleností pixelů. Výpočet vzdálenosti je pomocí euklidovské vzdálenosti. Jedná se o transformaci, kdy je každému bodu objektu v obraze přiřazena hodnota, která se rovná jeho vzdálenosti k nejbližšímu bodu pozadí. Distanční mapa je tedy šedotónový obraz s nezápornými hodnotami, který popisuje vlastnosti binárního obrazu. Přičemž hodnoty uprostřed objektu mají nejvyšší hodnotu. Dalo by se říci, že lokální maxima v distanční mapě korespondují se středovými liniemi.

Popsané metody detekují i falešné linie, a proto je nutné tyto falešné linie odstranit. Detekované falešné linie je možné vidět na obrázku 17a. K odstranění falešných linií se používají trackovací algoritmy. Po odstranění falešných linií je dalším nedostatkem absence středové linie až k obrysu chromosomu (obrázek 17b). Tudíž dalším krokem je dopočítání středové linie až k obrysu. Dopočítaná středová linie je možná vidět na obrázku 17c.

Tento klasifikační znak slouží také k zjištění velikosti, jak tomu bylo v případě obsahu chromosomu. Zároveň tento klasifikační znak slouží k určení zbylých dvou klasifikačních znaků. Tudíž při nesprávném výpočtu středové linie dochází ke špatné detekci i polohy centromery a jasového profilu chromosomu.

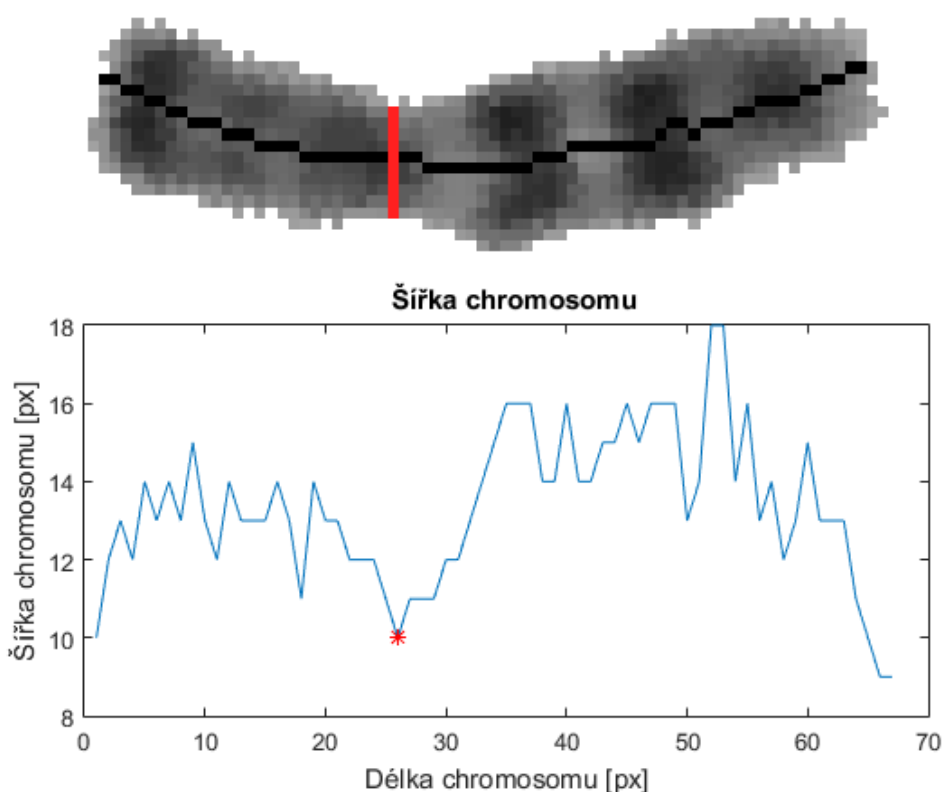


Obrázek 17: Chromosomy se středovou linií

- a) chromosom se středovou linií a falešnými konci, b) chromosom středovou linií a bez falešných konců, c) chromosom s prodlouženými konci až k obrysu chromosomu

## Poloha Centromery

Z morfologického hlediska je centromera nejužší místo chromosomu v podélném směru. Poloha centromery vychází ze středové linie. Na středovou linii jsou pomyslně kladeny kolmice a je zjišťována šířka v podélném směru chromosomu. Tímto způsobem je získán 1 D signál, kde globální minimum je právě pozice centromery. Chromosom se středovou linií a červenou vyznačenou centromerou je zobrazen v horní části obrázku 18. Ve spodní části obrázku 18 je zobrazena šířka chromosomu společně s vyznačeným místem globálního minima.



Obrázek 18: Horní část chromosom se středovou linií a pozicí centromery a spodní část šířka signálu chromosomu s označeným místem centromery

Pomocí polohy centromery a délky chromosomu lze zařadit každý chromosom do jedné ze sedmi skupin. Další využití pozice centromery spočívá v zjištění, jak je chromosom otočený. Vychází se ze vzorce:

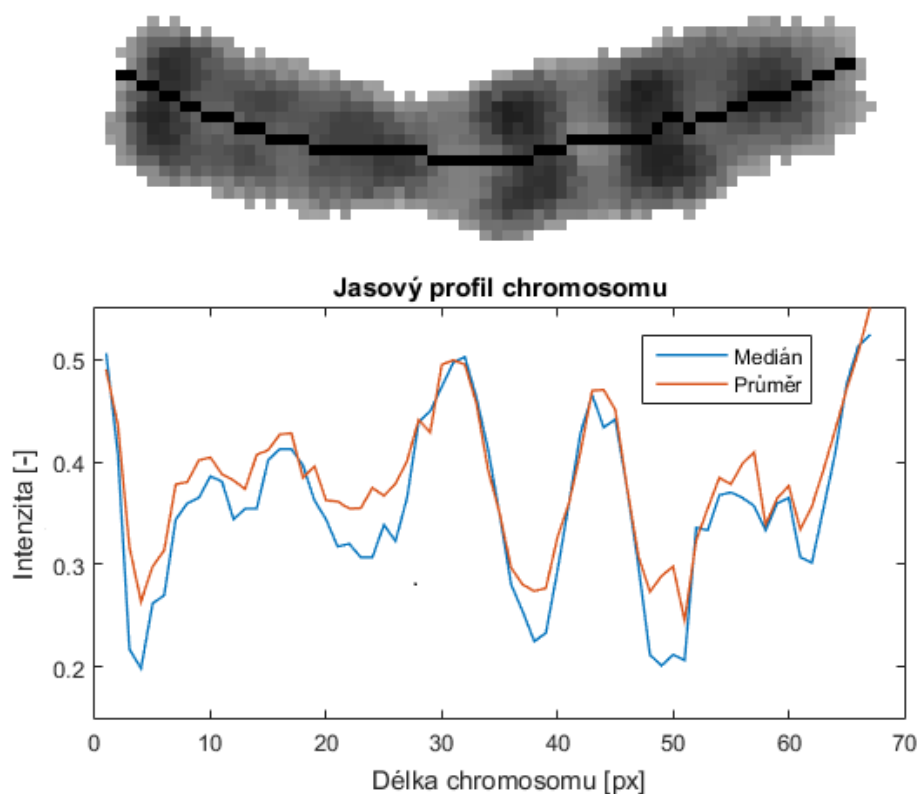
$$l = \frac{p}{d}, \quad (18)$$

kde  $p$  je pozice centromery v signále,  $d$  je celková délka chromosomu a  $l$  poměr mezi pozicí centromery a délkou chromosomu. Pokud  $l$  překročí práh 0,5, tak dojde k otočení chromosomu o  $180^\circ$ . Při určování pozice centromery dochází k chybnému určení, jestliže

signál neobsahuje pouze jedno globální minimum, ale více globálních minim. V tomto případě je šířka chromosomu ve všech bodech středové linie podobná. Tato chyba se vyskytuje u chromosomů, které jsou příliš ohnuté nebo špatně segmentovány od pozadí.

### Jasový profil chromosomu

Posledním znakem, který se taky zjišťuje ze středové linie, je jasový profil chromosomu. Podobně jako u pozice centromery, tak i u jasových profilů chromosomu jsou pomyslně kladeny kolmice na středovou linii (horní část obrázku 19). Z kolmic je následně vypočítána průměrná intenzita a mediánová intenzita. Jasové profily chromosomů mohou sloužit k porovnání homologních párů chromosomu, k určování trisomie a monosomie nebo k zjištění pohlaví jedince. Průměrnou a mediánovou intenzitu je možné vidět v spodní části obrázku 19. [22], [36], [37], [38], [39]

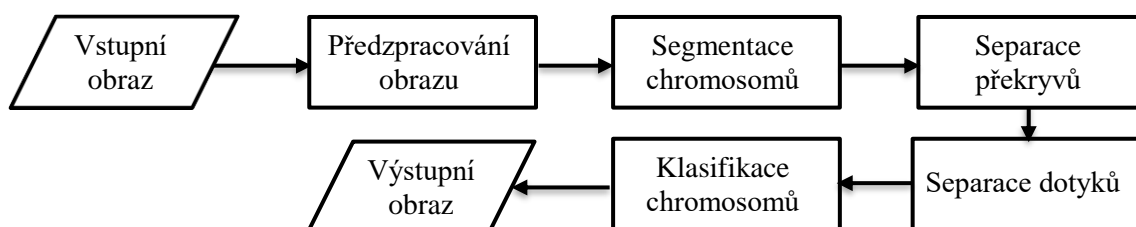


Obrázek 19: Horní část chromosom se středovou linií a spodní část průměrný a mediánový jasový profil

## 4 REALIZACE OBRAZOVÉ ANALÝZY

Cílem práce je návrh a realizace algoritmu pro automatické vyhodnocení karyotypu ze snímku mitotických chromosomů. Navržené řešení je realizováno v prostředí Matlab tvořené standartními knihovnamí. Algoritmus je testován na obrazech ze dvou volně dostupných zdrojů. [12], [13]

Protože je G a Q pruhování nejčastější metodou, zaměřil jsem se primárně na obrazy mitotických chromosomů z těchto barvicích technik. Molekulární metody vizualizují přímo určené strukturní aberace, a proto se jako vstupní obraz do algoritmu nehodí. Velmi záleží i na způsobu pořízení snímku a následné kompresi. Snímky s nízkou kvalitou již neobsahují potřebné informace a je velmi obtížné snímky dále zpracovávat.

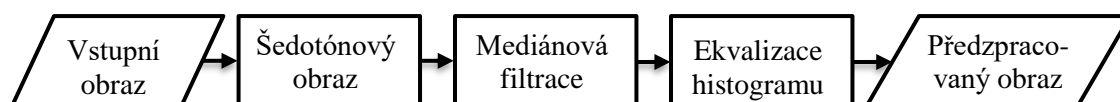


Obrázek 20: Blokový diagram obrazové analýzy mitotických chromosomů

Obrázek 20 znázorňuje blokový diagram algoritmu. Celá analýza se skládá z pěti modulů, které jsou podrobněji popsány v následujících podkapitolách. Jako vstup do algoritmu jsou použity snímky ve formátu jpg a bmp. Po načtení obrazu nastává předzpracování a segmentace obrazů chromosomů od pozadí obrazu. Dalším krokem je separace překrývajících se a dotýkajících se chromosomů. Následuje analýza obrazu mitotických chromosomů. Analýza spočívá v zařazení chromosomů do sedmi skupin podle velikosti. V každé skupině jsou následně párovány homologní páry chromosomů a je vykreslen výsledný obraz lidského karyotypu. [10], [18]

### 4.1 Předzpracování obrazu

První modul je předzpracování vstupního obrazu. Předzpracování obrazu je rozděleno do několika kroků. Obrázek 21 znázorňuje blokový diagram ve kterém jsou zobrazeny kroky předzpracování.



Obrázek 21: Blokový diagram předzpracování obrazu

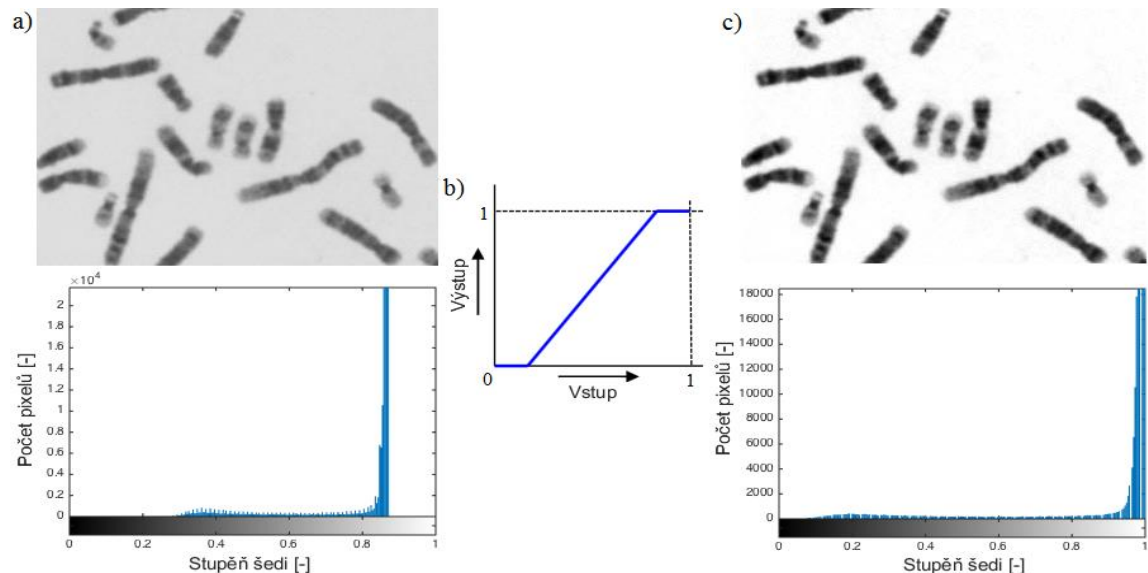


Vstupem algoritmu je šedotónový obraz. Barevné obrazy jsou automaticky převedeny na šedotónové obrazy. Odlišné barvicí techniky způsobují to, že obraz má jinou barvu pozadí. Pokud obraz obsahuje tmavé pozadí, dalším krokem je invertování jasových hodnot obrazu. Algoritmus dále zpracovává obraz se světlým pozadím. Mediánová filtrace se v obraze využívá z důvodu eliminace zkreslení typu „Pepř a sůl“, vznikající při pořízení snímku. V algoritmu je použita mediánová filtrace s velikostí masky 3 x 3. [33]

Posledním krokem je úprava jasu obrazu pomocí ekvalizace histogramu. Jedná se o transformaci s využitím všech hladin šedi rovnoměrně. Cílem této transformace je „roztáhnout“ histogram na celou škálu šedi, aby došlo ke zvýšení dynamického rozsahu. V šedotónovém obraze je černá barva reprezentovaná nulou a bílá barva reprezentovaná jedničkou. Například pokud se nejvyšší hodnota jasu vyskytuje jako 0,85 a nejnižší hodnota jako 0,15, pak je obraz transformován podle transformační funkce, viz. obrázek 22b. Obrázek 22a zobrazuje originální vstupní obraz společně s histogramem. Obrázek 22c znázorňuje obraz po ekvalizaci společně s histogramem. Tato lineární transformace lze matematicky vyjádřit pomocí rovnice:

$$g_{i,k} = \frac{f_{i,k} - \min(f)}{\max(f) - \min(f)}, \quad (19)$$

kde  $f_{i,k}$  je hodnota aktuálního pixelu,  $\min(f)$  je minimum vstupního obrazu a  $\max(f)$  je maximum vstupního obrazu. [39]



Obrázek 22: Transformace s využitím ekvalizace histogramu

- a) originální snímek mitotických chromosomů (nahore) a histogram z originálního obrazu (dole), b) transformační funkce pro změnu intenzity, c) výsledný snímek se zvýšeným kontrastem (nahore) a histogramem z výsledného obrazu (dole).

## 4.2 Segmentace chromosomů od pozadí

Vstupem do druhého modulu je předzpracovaný obraz. Tento modul slouží k segmentaci chromosomů od pozadí snímku. Segmentace chromosomů se skládá z několika kroků:

1. Získání masky pruhovací metodou Otsu a odstranění nečistot
2. Aplikace aktivních kontur s maskou z předchozího kroku
3. Ohraničené kontury v obraze a výběr chromosomů

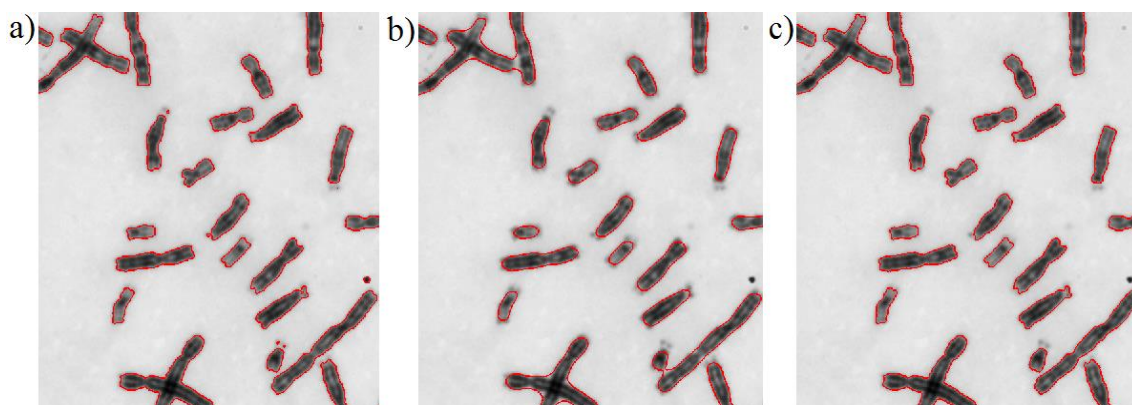
V prvním kroku je nutné nejprve spočítat práh Otsu metodou, která byla popsána v minulé kapitole. Tímto krokem získáme práh pro tvorbu binární masky, která je nutná pro další zpracování. Ukázka binární masky je zobrazena v levé části v obrázku 23. Z obrázku je patrné, že binární obraz obsahuje nečistoty nebo chybně vybrané pixely uprostřed chromosomu. Proto je na masku použit mediánový filtr. Velikost masky mediánového filtru byla nastavena na velikost 2% původní velikosti vstupního obrazu. Tímto krokem získáme binární masku (pravá část obrázku 23) s hladšími přechody v obraze, který neobsahuje zmíněné nedostatky. Binární obraz však spojil některé chromosomy dohromady. Tento problém je řešen v dalším kroku.



Obrázek 23: Binární obraz po naprahování Otsu metodou (vlevo) a binární obraz po naprahování a následné mediánové filtraci (vpravo)

Druhý krok je založen na metodě aktivních kontur, která byla rovněž blíže představena v předchozí kapitole. Metoda využívá uzavřenou křivku, která aproximuje hranici chromosomu podle okolních pixelů. Vstupem pro funkci je binární maska a šedotónový obraz. Aktivní kontury se počítají v iteracích. Algoritmus využívá pouze 50 iterací, kdy dosahuje dostatečné přesnosti. Obrázek 24 znázorňuje srovnání ohraničení chromosomů s odlišnými metodami a postupy. Obrázek 24a zobrazuje ohraničení po Otsu prahování. Je patrné, že obraz obsahuje detekované nečistoty a zbytky kratších ramen

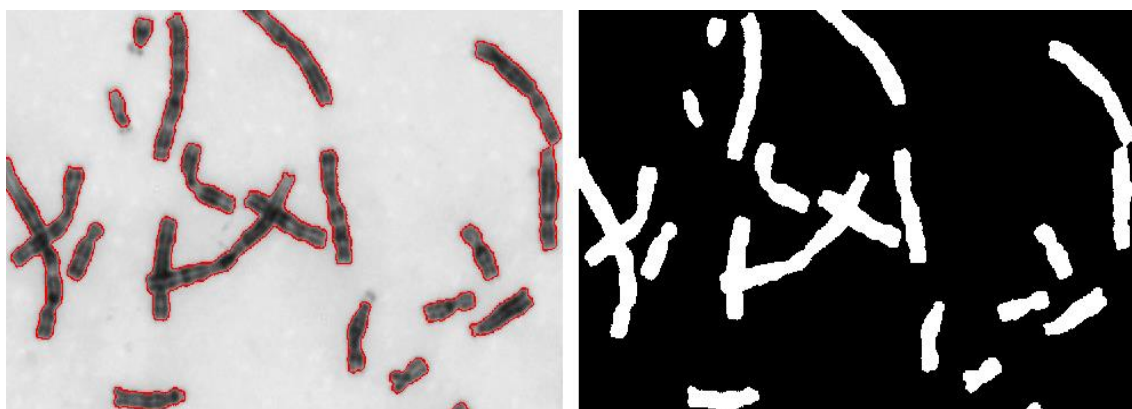
chromosomu. Obrázek 24b zobrazuje Otsu prahování s mediánovou filtrací. Z obrazu je patrné, že se některé chromosomy spojily, ale došlo k nutné eliminaci nečistot. Obrázek 24c obsahuje obě metody a přidává aktivní kontury, které spolehlivě rozdělí spojené chromosomy. Bohužel nedošlo k ohraničení kratších ramen chromosomů, které nejsou dostatečně jasově výrazné. Obrázek 25 zobrazuje obraz kontury chromosomů v předzpracovaném obraze a obraz jeho binární masky.



Obrázek 24: Srovnání segmentace chromosomů

a) ohraničení chromosomů pomocí Otsu metody, b) ohraničení chromosomů pomocí Otsu metody s mediánovou filtrací, c) ohraničení chromosomů pomocí aktivních kontur.

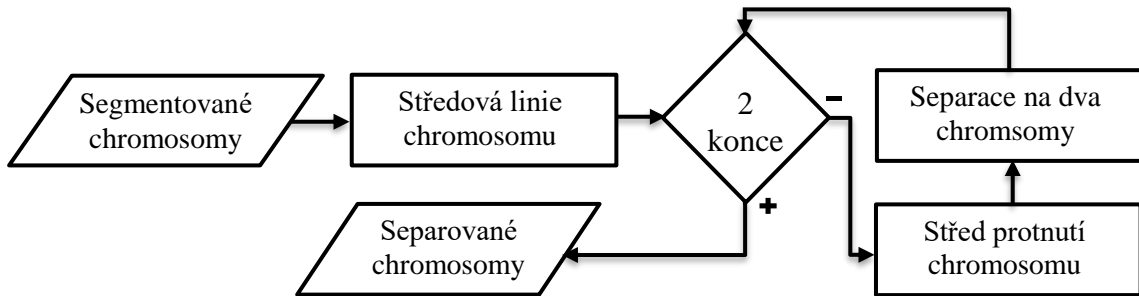
Posledním krokem je výběr vždy jedné uzavřené kontury do nového obrazu. Vstupní obraz je rozdělen do několika skupin podle uzavřené kontury (pokud obsahuje 46 chromosomů, je rozdělen do 46 skupin). Jednotlivé skupiny jsou vyříznuty do nového obrazu a jsou uloženy jako binární masky, ale také jako šedotónové obrazy s informacemi z předzpracovaného obrazu. Ideálně by mělo být vyseparováno 46 chromosomů, tedy 46 nových obrazů. Mezi chromosomy se objevují také objekty jako jádérko, které je mnohonásobně větší než chromosom. Tyto objekty jsou z další analýzy vyřazeny.



Obrázek 25: Ohraničené chromosomy pomocí aktivních kontur (vlevo) a binární maska s výběrem chromosomů (vpravo)

### 4.3 Separace překryvů u chromosomů

Tento modul realizuje detekci překryvů a posléze separaci chromosomů. Separace chromosomů je založena na získání středové linie obrazu. Obrázek 26 popisuje jednotlivé části algoritmu pro rozdělení chromosomů.

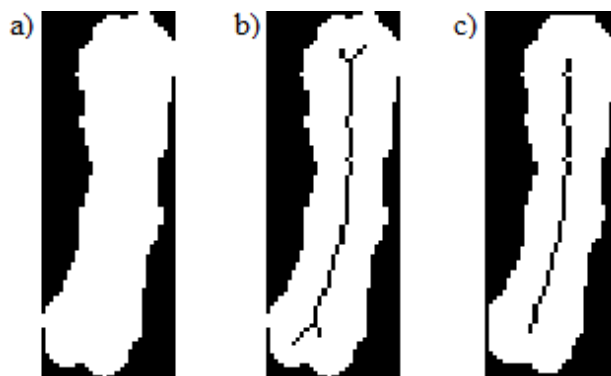


Obrázek 26: Blokový diagram pro separaci překrývajících se chromosomů.

Do této části algoritmu vstupují jednotlivé binární obrazy segmentovaných chromosomů z předešlé části. Z binárních obrazů (obrázek 27a) jsou následně počítány středové linie, pomocí iterativního ztenčování obrysu binárního obrazu. Jakmile se obraz nemění je cyklus zastaven. Tento druh výpočtu středové linie občas vede k vytvoření falešných konců středové linie (obrázek 27b). K odstranění falešných konců středové linie je použit morfologický operátor uzavření, který je aplikován na vstupní binární obraz a znovu spočítá středovou linii (obrázek 27c). Tento krok odstraní většinu falešných konců. Pokud ne, následuje trackovací algoritmus. Odstranění falešných konců pomocí trackovacího algoritmu je využito pouze pokud je ve středové linii víc jak dva konce. Trackovací algoritmus hledá od středu protnutí dvou středových linií konce, a počítá jejich vzdálenost. Jakmile vzdálenost konců středové linie od středu středové linie překročí zadaný práh, je tato část středové linie odstraněna. Práh je heuristicky určen na vzdálenost 10 pixelů. Vzdálenost mezi středy a konci středové linie je počítána pomocí euklidovské vzdálenosti, která lze vyjádřit rovnicí [41]:

$$e(k, s) = \sqrt{(k_x - s_x)^2 + (k_y - s_y)^2}, \quad (20)$$

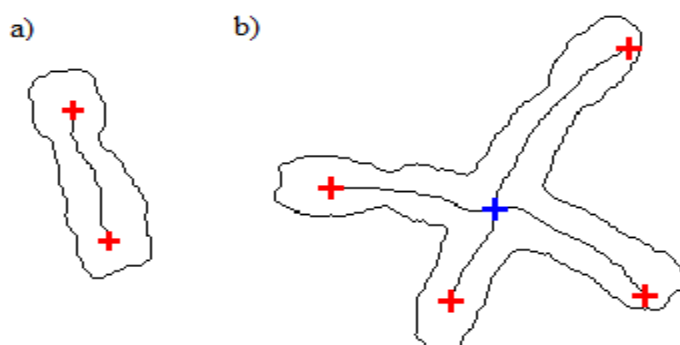
kde  $k_x$  a  $k_y$  vyjadřují pozici konce středové linie a  $s_x$ ,  $s_y$  vyjadřují pozici středu středové linie a  $e$  vyjadřuje vypočtenou euklidovskou vzdálenost.



Obrázek 27: Středová linie v binárních obrazech

- a) binární obraz chromosomu, b) chromosom se středovou linií bez použití morfologické operace uzavření, c) chromosom se středovou linií s použitím morfologické operace uzavření

Jakmile je určena středová linie, je nutné zjistit počet konců ve středové linii. Pokud jsou nalezeny dva konce, jedná se o jeden chromosom a výpočet končí. Pokud je nalezeno ve středové linii více než dva konce, tak obraz postupuje dále ke zpracování. Na obrázku 28a je demonstrován osamocený chromosom s nalezenými dvěma konci středové linie. Obrázek 28b zobrazuje obraz chromosomu s čtyřmi konci a s nalezeným středem překryvu, který postupuje dále do této části algoritmu.

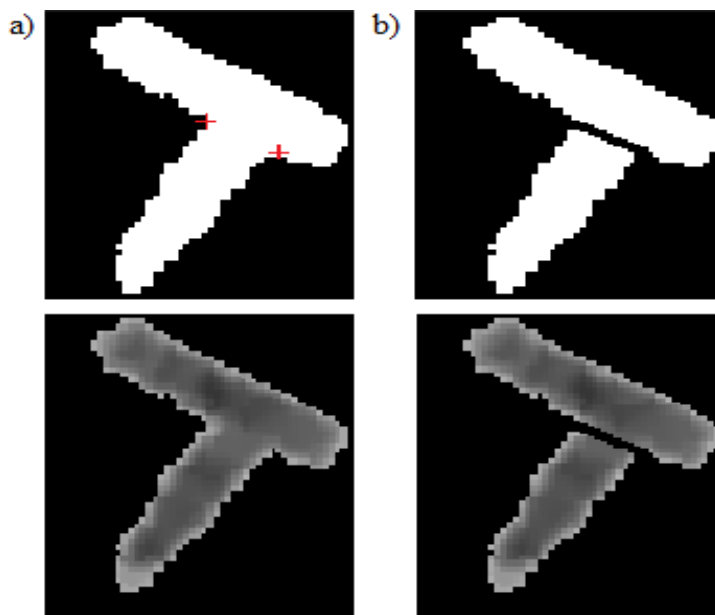


Obrázek 28: a) Obrys chromosomu se středovou linií a detekovanými dvěma konci, b) obrys chromosomu se středovou linií a detekovanými čtyřmi konci a jedním středem.

Střed překryvu je vyhodnocen jako místo, kde se setkají alespoň tři části středové linie. Po zjištění středu následuje detekce okolních bodů u středu chromosomu, tak aby mohlo dojít k separaci. V okolí středu dochází k nárůstu okolních bodů vždy o jeden pixel z každé strany v cyklu. Dochází tedy k jakémusi zvětšování čtverce. Cyklus je zastaven dvěma způsoby. První způsob ukončení je po dosažení okraje obrazu. Druhý způsob ukončení spočívá v nalezení čtyř od sebe oddělitelných skupin bodů.

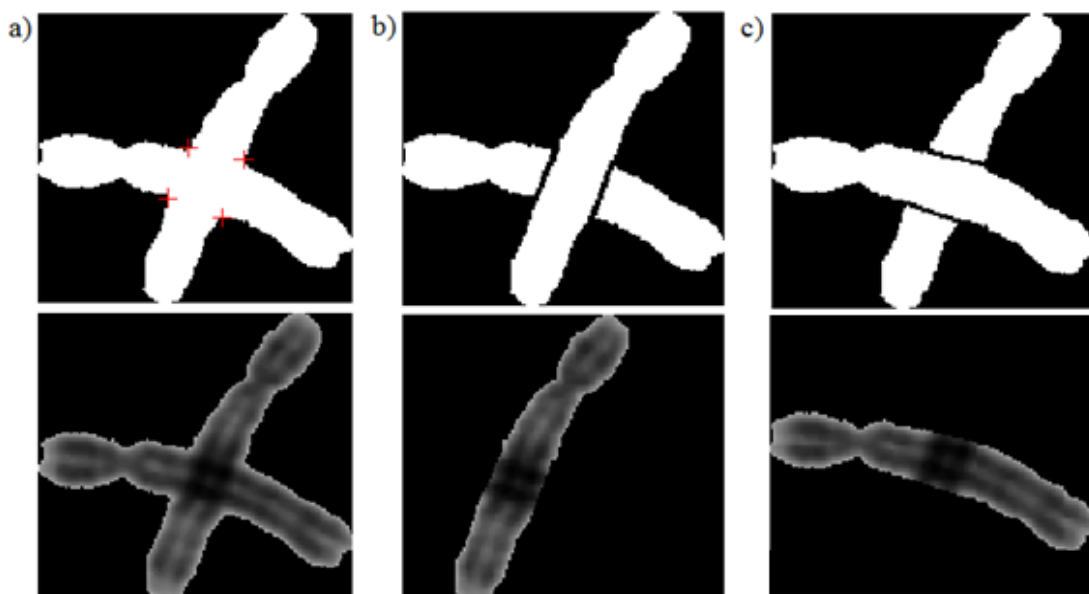
V prvním případě je v čtverci určen počet oddělitelných skupin bodů s opačnou intenzitou pixelů, než je intenzita u středu. Pokud jsou nalezeny tři skupiny bodů, jedná

se o překryv, který si můžeme představit jako písmeno T. Z překryvu jsou dále vybrány dvě skupiny bodů, které mají nejmenší obsah. Následně je v těchto dvou skupinách vždy nalezen bod, který je nejbližší ke středu. Tyto dva body tvoří fiktivní přímku, kterou je shluk chromosomů rozdělen. Na obrázku 29a je vidět binární obraz s nalezenými dvěma body a šedotónový shluk dvou chromosomů. Obrázek 29b již znázorňuje binární a šedotónový obraz, který byl rozdělen fiktivní přímkou. Při překryvu do písmene T dochází ke ztrátě části chromosomu.



Obrázek 29: Rozdělení překrývajících se chromosomů do písmene T s červenými značkami

Druhý způsob ukončení cyklu spočívá v nalezení čtyř od sebe oddělitelných skupin bodů s opačnou intenzitou pixelů, než je intenzita ve středu. Jedná se tedy o překryv do písmene X. Z každé skupiny je vybrán bod, který je nejbližší ke středu. Takto nalezené body je možno vidět v horní části obrázku 30a. Nalezené body následně separují chromosomy ve dvou krocích. V prvním kroku je náhodně vybrán jeden bod. K tomuto bodu jsou spočteny vzdálenosti zbylých bodů. Nejbližší bod k náhodně vybranému bodu určuje první dva body, které se mají spojit. Zbylé dva body spojí druhou přímkou, jak je možno vidět na obrázku 30b v horní části. V druhém kroku je spojen náhodný bod s druhým nejbližším bodem. Tyto body tvoří první přímkou a zbylé dva body druhou přímkou. Tímto dochází k separaci druhého chromosomu, viz horní část obrázku 30c. Dolní části obrázku 30 znázorňují tři obrazy. Obrázek 30a původně překřížený chromosom. Obrázek 30b separovaný první chromosom. Obrázek 30c zobrazuje separovaný druhý chromosom. Z obrazu je zřejmé, že překřížené části obrazu byly znehodnoceny a v následné analýze mohou způsobovat chyby.



Obrázek 30: Rozdělení překrývajících se chromosomů do písmene X s červenými značkami

Po separaci chromosomy podstupují znovu výpočet středové linie a výpočet počtu konců středové linie. Pokud je počet konců větší než dva, tak výsledný obraz postupuje znovu do separace překryvů chromosomu.

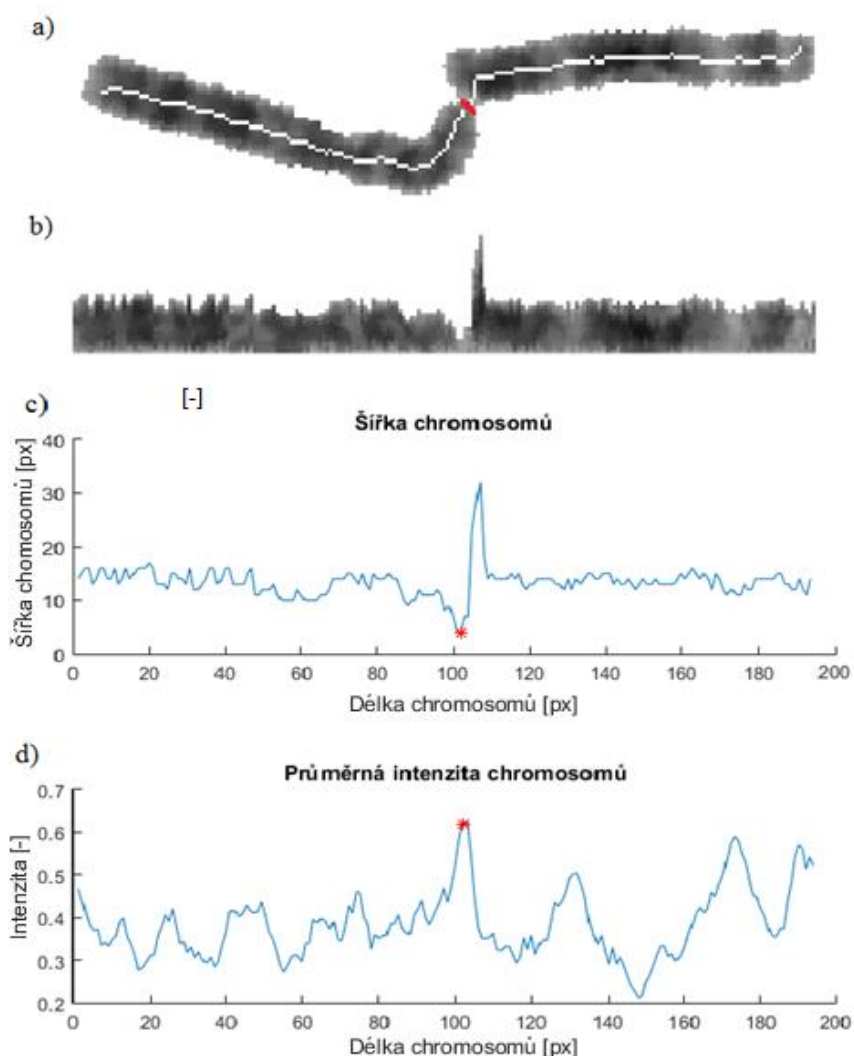
#### 4.4 Separace dotyků u chromosomů

Dalším krokem algoritmu je separace dotýkajících se chromosomů (většinou se jedná o spojení dvou chromosomů). Právě proto tento modul řeší rozdělení dvou dotýkajících chromosomů. Chromosomy jsou v místě dotyku charakterizovány zúžením a zvýšením jasové hodnoty oproti zbytku chromosomu, čehož je v algoritmu využito. Centromera je také charakterizována zúžením, ale obsahuje nižší jasové hodnoty.

Do čtvrtého modulu vstupují všechny obrazy z předešlé části. U těchto obrazů se dále pracuje se středovou linií (obrázek 31a). Středová linie je procházena po třech pixelech od prvního bodu do posledního bodu. Na prostřední pixelu ze tří je spočtena kolmice v obou směrech a je vybrána úsečka jasových hodnot z šedotónového obrazu. Jedná se tedy o narovnání chromosomu. Narovnaný chromosom je možné vidět na obrázku 31b. Počet pixelů v dané úsečce značí šířku chromosomu. Z jasových hodnot v úsečce je dále vypočtena průměrná hodnota jasu. Tímto krokem získáme dva 1 D signály, a to šířku chromosomů (obrázek 31c) a průměrnou intenzitu chromosomů (obrázek 31d). V šířce chromosomu je následně hledáno minimum (červené značky v obrázku 31a a obrázku 31d). Pokud je minimum menší než práh a průměrná jasová hodnota větší než práh, tak je daná pozice vyhodnoceno jako místo dotyku. Prah je



nastaven pro šířku jako dvě třetiny z průměrné šířky chromosomu a pro intenzitu jako jeden a půl násobek průměrné hodnoty z průměru intenzit. Místo dotyku je zobrazeno červeně na obrázku 31a. Místem dotyku je vedena opět kolmice a chromosom je rozdělen na dva samostatné chromosomy, pro které je spočítán obsah. Obsah chromosomu je kontrolním prvkem velikosti chromosomu. Pokud je obsah chromosomu menší než 2% velikosti původního obrazu, tak k rozdělení chromosomů nedojde.

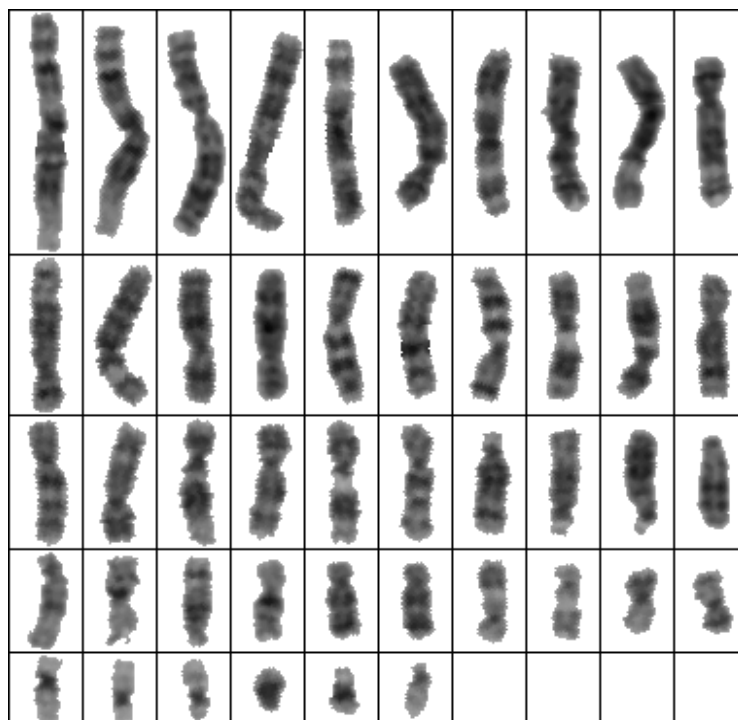


Obrázek 31: Separace dotýkajících se chromosomů

a) chromosom se středovou linií s červeně zvýrazněným místem dotyku, b) narovnaný chromosom, c) šířka chromosomu v 1 D signálu s vyznačeným místem dotyku, d) průměrná intenzita chromosomu s vyznačeným místem dotyku

Separace dotýkajících se chromosomů je posledním krokem algoritmu, který se řadí k segmentaci a separaci chromosomů. Separované chromosomy je možné vidět na obrázku 32, kde jsou vykresleny podle obsahu binárního obrazu do nového obrazu. Po segmentaci dojde k vypsání počtu chromosomů a zda se jedná o zdravý karyotyp.



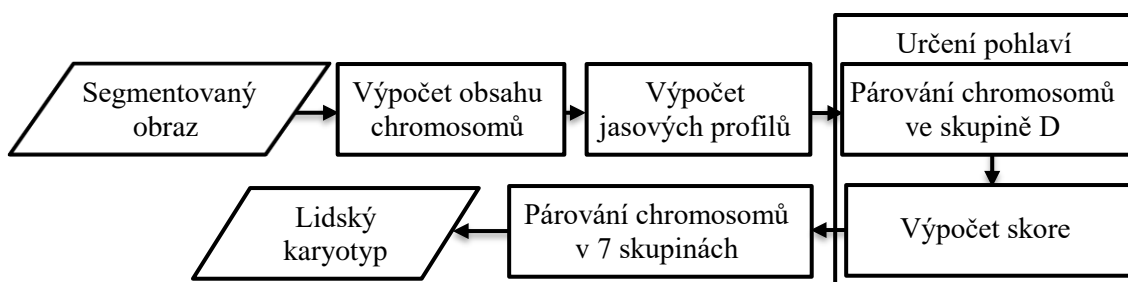


Obrázek 32: Výsledný obraz po segmentaci chromosomů

## 4.5 Klasifikace chromosomů do karyotypu

Klasifikace chromosomů probíhá pouze na zdravé sestavě chromosomů. Chromosomy jsou zařazeny do lidského karyotypu podle klasifikačních znaků, které byly podrobněji popsány v kapitole tři. V této práci je pro vyhodnocení karyotypu použito dvou klasifikačních znaků: obsah chromosomů a jasový profil.

Obrázek 33 zobrazuje blokový diagram popisující vyhodnocení lidského karyotypu. Prvním krokem je výpočet obsahu chromosomů na základě binárních obrazů a jasových profilů na základě středové linie. Následným krokem je určení pohlaví. Pohlaví je určeno pomocí skóre. Skóre je vypočteno pomocí spárovaných chromosomů ve skupině D. Dle pohlaví jsou chromosomy rozděleny do sedmi skupin podle obsahu. V každé skupině následně dochází k párování chromosomů. Jakmile jsou spárovány všechny chromosomy, dojde k vykreslení obrazu s lidským karyotypem.



Obrázek 33: Blokový diagram vyhodnocení lidského karyotypu

### Obsah chromosomu

Obsah chromosomu je počítán z binárních obrazů segmentovaných chromosomů. V binárním obraze jsou chromosomy znázorněné světlými pixely (hodnota 1) a okolí tmavými pixely (hodnota 0). Výpočet obsahu chromosomů spočívá v součtu všech hodnot pixelů v daném obraze. Obsah chromosomů slouží k zařazení jednotlivých chromosomů do jedné ze sedmi skupin (A až G). Ve skupině A jsou chromosomy s největším obsahem a ve skupině G s nejmenším obsahem.

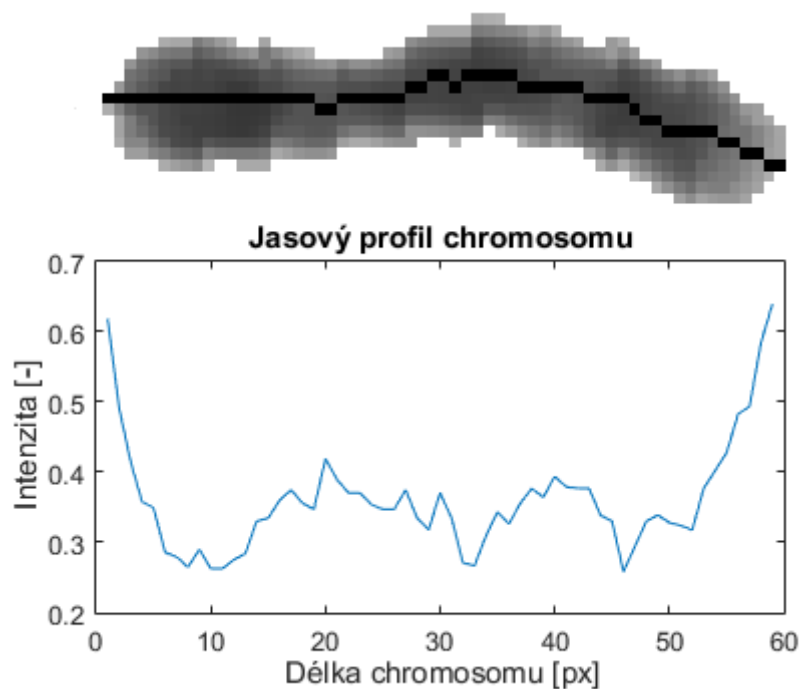
### Jasový profil chromosomu

Pro výpočet jasových profilů se znovu využije středová linie chromosomu. Na rozdíl od segmentace, je nutné v klasifikaci středovou linii prodloužit až do konce obrysu obrazu, viz obrázek 34a. Prodloužení středové linie až ke konci spočívá v zjištění směrnice konců středové linie. Směrnice je počítána posledními třemi body z těchto konců. Ve směru směrnice jsou postupně dopočítány pixely až k obrysu chromosomu. Dopočítaná středová linie až k obrysu chromosomu je zobrazena na obrázku 34b.



Obrázek 34: Středová linie (vlevo) a středová linie s prodlouženými konci (vpravo)

Následně jsou určeny jasové profily chromosomů. Jasové profily jsou počítány analogicky jako v minulé podkapitole s jedním rozdílem. Místo výpočtu průměru je z dané kolmice s jasovými hodnotami vybrána mediánová hodnota jasu. Mediánový jasový profil chromosomu je zobrazen v spodní části obrázku 35. V horní části je vykreslen šedotónový obraz chromosomu společně se středovou linií, kterou jsou vedené kolmice.



Obrázek 35: Chromosom se středovou linií společně s mediánovým jasovým profilem

### Určení pohlaví

Ve třetím kroku je nutné nastavit pohlaví manuálně, nebo využít funkci k automatickému určení chromosomu. Při automatické detekci se využívá faktu, že ženský a mužský karyotyp se liší v počtu chromosomů, které jsou zařazeny ve skupině C. Ženský karyotyp obsahuje 16 chromosomů ve skupině C a 4 chromosomy ve skupině G. Zatímco mužský karyotyp obsahuje 15 chromosomů ve skupině C a 5 chromosomů ve skupině G. [5]

Rozdílný počet chromosomů u pohlaví ve skupině C nám vytváří odlišné sestavy chromosomů ve skupině D. Proto je využito párování ve skupině D a následně vypočteno skóre. Nejprve jsou párovány homologní chromosomy pro ženu (27. až 32. chromosom) a poté pro muže (26. až 31. chromosom). Pohlaví, které dosáhne menšího skóre je vyhodnoceno jako výsledné. Skóre je počítáno z jasových profilů a pomocí kvadratické odchylky signálu (RMSD) s váhováním.

Párování homologních chromosomů spočívá v postupném porovnávání vždy dvou jasových profilů ve skupině D. Chromosomy mají odlišnou délku jasových profilů a zároveň mohou být otočeny o  $180^\circ$ . Proto je v prvním kroku nutné kratší chromosom otočit o  $180^\circ$  a počítat vše dvakrát. Jednou pro otočený signál a podruhé pro původní signál. Vybrán je signál s menší hodnotou kritériální funkce. V druhém kroku je signál s menší délkou jasového profilu posouván po signálu s větší délkou jasového profilu

v okně. Velikost okna je právě délka kratšího chromosomu. Vždy dva jasové profily jsou postupně porovnávány podle kvadratické odchylky, která je váhována podílem obsahu chromosomu. Kvadratická odchylka lze matematicky vyjádřit [41]:

$$RMSD = \sqrt{\frac{\sum_{t=1}^N (s_1(t) - s_2(t))^2}{N}}, \quad (21)$$

kde  $RMSD$  značí kvadratickou odchylku mezi dvěma jasovými profily,  $s_1(t)$  je jasový profil prvního chromosomu v pozici  $t$  a  $s_2(t)$  je jasový profil druhého chromosomu v pozici  $t$ ,  $N$  je délka kratšího chromosomu. Poměr obsah chromosomů lze vyjádřit pomocí rovnice:

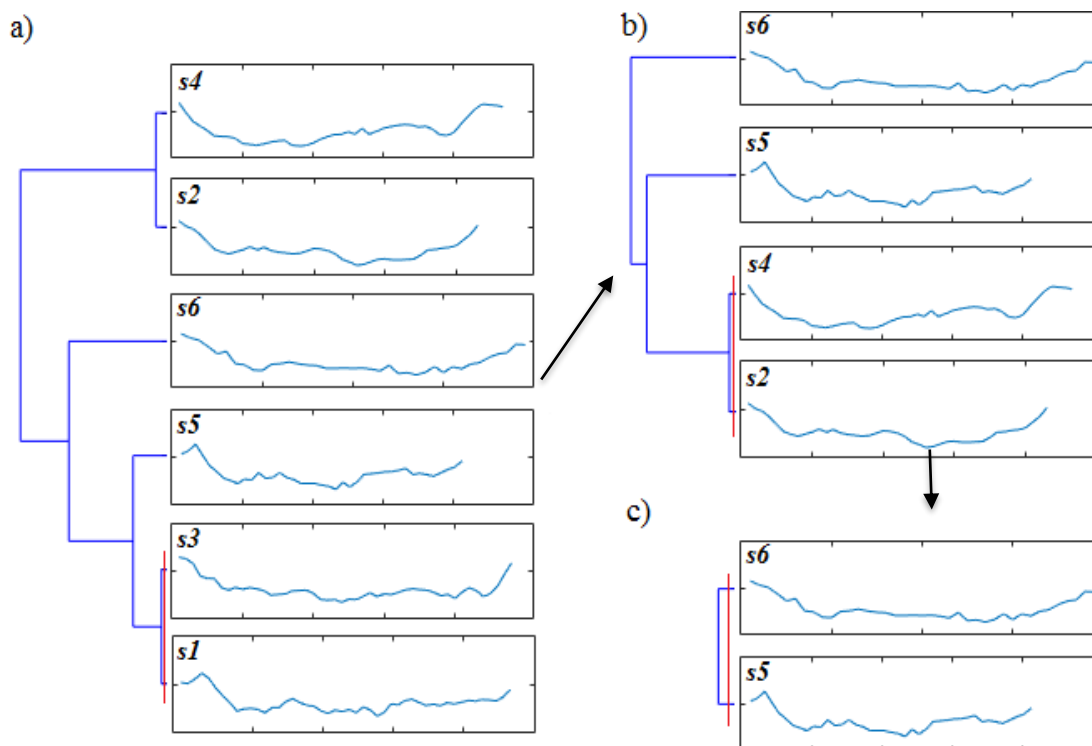
$$w = \frac{O_{s1}}{O_{s2}}, \quad \text{kde } O_{s1} < O_{s2} \quad (22)$$

kde  $O_{s1}$  je obsah prvního chromosomu a  $O_{s2}$  je obsah druhého chromosomu. Tyto dvě rovnice společně vyjadřují kritériální funkci:

$$C = \frac{RMSD}{w}, \quad (23)$$

kde  $RMSD$  značí kvadratickou odchylku mezi dvěma jasovými profily a  $w$  je poměr obsahu dvou chromosomů.

Tímto způsobem dochází k porovnání každého chromosomu s každým v rámci skupiny D. Hodnoty jsou zapisovány do matice 6x6. Z matice je následně vybráno minimum a chromosomy s minimem jsou spárovány a hodnota kritériální funkce je uložena. Výpočet se opakuje, dokud nejsou spárované všechny chromosomy. Princip je vysvětlen na obrázku 36. Na obrázku 36a je zobrazeno šest jasových profilů s dendrogramem (diagram znázorňující podobnost chromosomů), které jsou vzájemně porovnávány. Červenou čarou je zobrazen práh. Na obrázku 36b je zobrazen další krok, kde jsou dva chromosomy už spárovány a porovnávají se jenom čtyři chromosomy. Obrázek 36b je poslední krok, kde jsou zařazeny zbylé dva chromosomy. Výsledné skóre se rovná sumě všech tří hodnot minim kritériální funkce. Na základě menší hodnoty skóre je následně rozhodnuto o pohlaví.

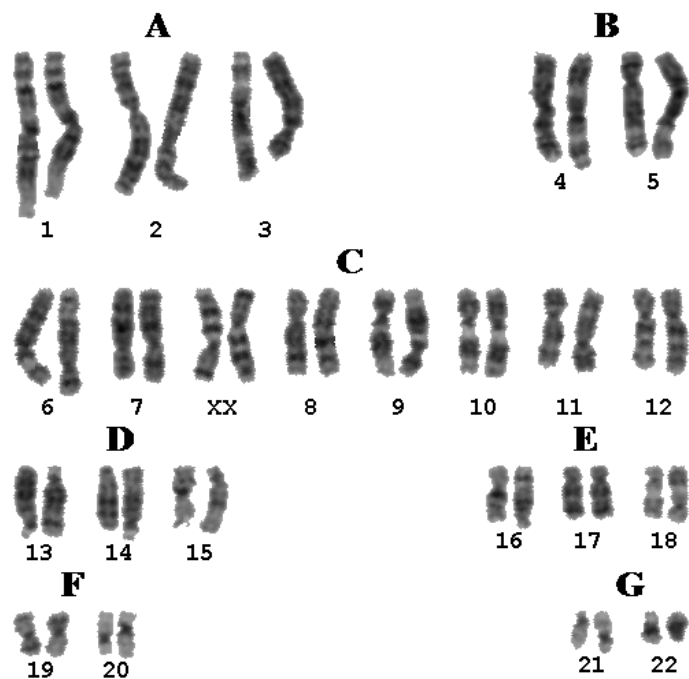


Obrázek 36: Princip párování homologních chromosomů

### Párování chromosomů v 7 skupinách

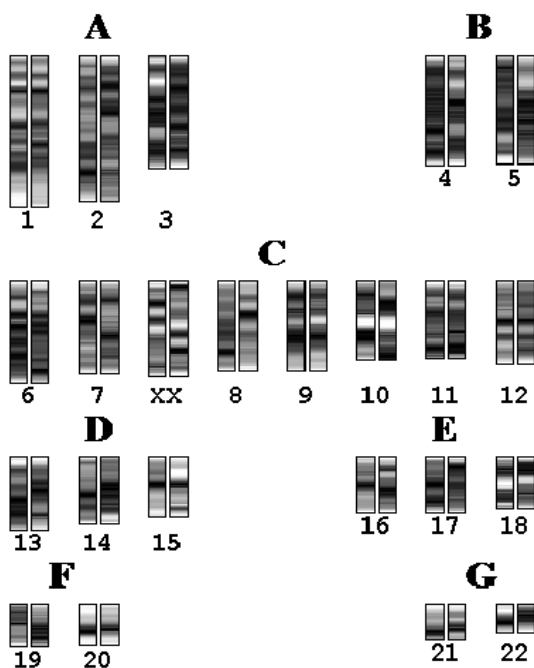
Po určení pohlaví následuje rozdělení chromosomů do sedmi skupin podle obsahu. V jednotlivých skupinách jsou chromosomy párovány stejným způsobem jako při určování karyotypu. Pomocí kriteriální funkce z jasových profilů a obsahu je hledáno minimum v matici. Matice je postupně zmenšována o dva chromosomy, dokud neobsahuje pouze poslední dva chromosomy. Tento proces je uplatněn u všech skupin kromě skupiny A. Ve skupině A jsou párovány pouze čtyři největší chromosomy. Je to z důvodu, že tyto chromosomy mají podobný obsah. Pátý a šestý chromosom je podstatně menší a v algoritmu docházelo k chybnému vyhodnocení.

Jakmile jsou spárovány všechny chromosomy ve všech skupinách je vykreslený obraz karyotypu. U ženského karyotypu se ve skupině C na třetím místě nachází párový chromosom X. U mužského karyotypu, je určen jako chromosom X ten chromosom, který není spárován s ostatními (lichý počet chromosomů - jeden zůstane). Chromosom Y je určen obdobně, ale ve skupině G. Výsledný zdravý ženský karyotyp je zobrazen na obrázku 37.



Obrázek 37: Výsledný zdravý ženský karyotyp

Posledním krokem této části algoritmu, který je volitelný, je vykreslení ideogramů z jasových hodnot jednotlivých chromosomů. Ideogramy byli upraveny na stejnou délku v homogních párech. Z důvodu zlepšení kontrastu bylo použito ekvalizace histogramu na jednotlivé chromosomy. Ideogramy z předchozího klasifikovaného obrazu je možné vidět na obrázku 38.

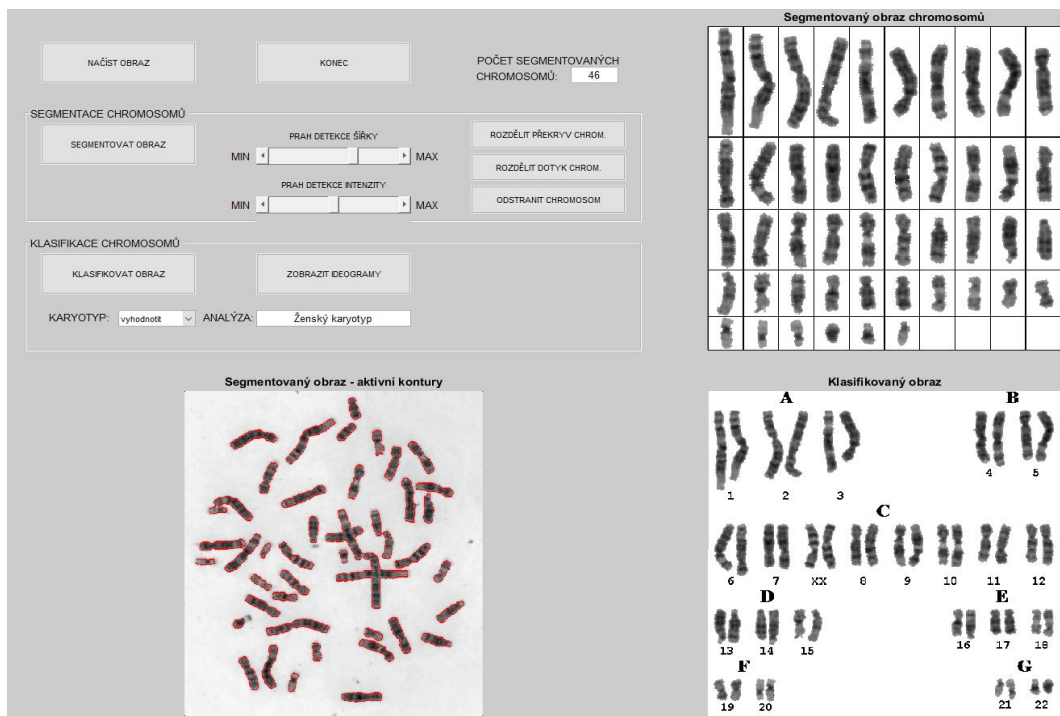


Obrázek 38: Ideogramy zdravého ženského karyotypu

## 4.6 Grafické uživatelské rozhraní

Algoritmus byl pro jednodušší obsluhu vybaven grafickým uživatelským rozhraním (GUI), které je zobrazeno na obrázku 39. Podrobnější návod je popsán v příloze B. GUI je rozděleno do 4 kvadrantů. V prvním kvadrantu se zobrazuje načtený obraz. Třetí kvadrant zobrazuje segmentovaný obraz a čtvrtý kvadrant vykresluje klasifikovaný lidský karyotyp. Druhý kvadrant obsahuje ovládací prvky se základními funkcemi:

- **Načíst obraz** – spouští dialogové okno, pro výběr vstupního obrazu
- **Konec** – tlačítko ukončení program Matlab.
- **Segmentace obrazu** – automatické segmentace chromosomu od pozadí.
- **Prah detekce šířky** – manuální práh šířky pro detekci dotýkajících se chromosomů (výchozí hodnota 0,66).
- **Prah detekce intenzity** – manuální práh intenzity pro detekci dotýkajících se chromosomů (výchozí hodnota 1,5).
- **Rozdělit překryv chromosomů** – po výběru dojde k rozdělení překryvů
- **Rozdělit dotyk chromosomů** – po výběru dojde k rozdělení dotyků
- **Smazat chromosom** – po výběru odstraní vybraný chromosom.
- **Klasifikace chromosomů** – automatické klasifikaci lidského karyotypu.
- **Karyotyp** – výběrový seznam, pro výběr pohlaví nebo vyhodnocení.
- **Zobrazit ideogramy** – v novém okně se zobrazí ideogramy chromosomů.



Obrázek 39: Grafické uživatelské rozhraní algoritmu

## 5 DISKUZE A VÝSLEDKY

Algoritmus byl testován na dvou databázích obrazů mitotických chromosomů. První databáze [12] obsahuje 117 snímků mitotických chromosomů s Q pruhu. Druhá databáze [13] se skládá z 38 snímků mitotických chromosomů s G pruhu. Obě databáze obsahují snímky zdravých lidských karyotypů. Každý obraz tedy obsahuje přesně 46 chromosomů. Úspěšnost detekce je rozdělena na úspěšnost segmentace chromosomů (segmentace od pozadí, rozdělení překryvů a dotyků chromosomů) a úspěšnost klasifikace chromosomů (párování chromosomů a určení pohlaví karyotypu).

Pro vyhodnocení segmentace je nutné si vysvětlit následující zkratky, pomocí kterých se hodnotí úspěšnost detekce:

- *SP* (skutečně pozitivní detekce) – jedná se o správně detekované chromosomy.
- *FN* (falešně negativní detekce) – jedná se o falešně detekované objekty, které byly detekovány (shluk více chromosomů, nečistoty).
- *FP* (falešně pozitivní detekce) – jedná se o chromosomy, které nebyly detekovány.

Senzitivita  $S^+$  udává pravděpodobnost pozitivní detekce, je-li v analyzovaném obrazu detekovaný chromosom. Výpočet senzitivity je matematicky vyjádřen pomocí rovnici:

$$S^+ = \frac{SP}{SP + FN}, \quad (24)$$

kde *SP* je skutečně pozitivní detekce a *FN* falešně negativní detekce.

Pozitivní prediktivní hodnota  $P^+$  pak udává pravděpodobnost přítomnosti chromosomů v případě pozitivní detekce. Výpočet pozitivní prediktivní hodnoty probíhá pomocí vztahu:

$$P^+ = \frac{SP}{SP + FP}, \quad (25)$$

kde *SP* je skutečně pozitivní detekce a *FP* je falešně pozitivní detekce. [42]

Pro vyhodnocení úspěšnosti párování chromosomů je úspěšnost počítána jako podíl správně klasifikovaných mitotických chromosomů ku všem klasifikovaným párům mitotických chromosomů. Úspěšnost určení pohlaví je pak vyhodnocena jako podíl správně určených pohlaví ku všem testovaným obrazům mitotických chromosomů.



## 5.1 Vyhodnocení a diskuze segmentace obrazu

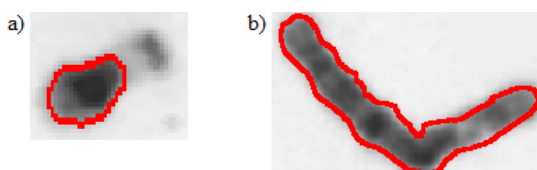
Segmentační část algoritmu byla testována, na zmíněných dvou databázích se snímky mitotických chromosomů. První databáze obsahovala snímky s Q pruhováním. Tyto snímky byly v dobré kvalitě uloženy v bezeztrátovém formátu. Druhá databáze obsahovala snímky s G pruhováním. Snímky G pruhování byly v horší kvalitě, protože byly uloženy v ztrátové formátu JPG. Vyhodnocení bylo provedeno na každé databázi zvlášť. Výsledky testování jsou zobrazeny v tabulce 1.

Tabulka 1: Úspěšnost segmentace chromosomů

	S <sup>+</sup> [%]	P <sup>+</sup> [%]
Q pruhování	89,2	91,3
G pruhování	86,4	92,4

Senzitivita snímků s Q pruhy je vyšší než senzitivita snímků z G pruhy. Nižší segmentační schopnost byla způsobena kvalitou snímku s G pruhy. Tyto výsledky jsou srovnatelné s jinými autory [33] a [34], kteří se zabývají podobnou problematikou. Publikované metody nebyly na takto velkém počtu obrazů. Dále se vždy zaměřili pouze na jeden druh pruhovací techniky.

Segmentace chromosomů se skládá z dílčích částí, kde mohlo docházet k odlišným druhům selhání detekce. Segmentace od pozadí je založena na vícefázové detekci založené na metodě Otsu a aktivních kontur. Po metodě Otsu prahování algoritmus odstraňuje drobné nečistoty. V této fázi docházelo k zmenšování akrocentrických chromosomů (obrázek 40a) a tudíž k zhoršování klasifikace chromosomů. Dalším častým problémem bylo nalezení středové linie. U některých obrazů docházelo k odstranění nejen falešných konců, ale i pravých, které definovaly překryv. Z tohoto důvodu nebyl chromosom správně rozdělen. Další chybné segmentace byly způsobeny chromosomy, které se dotýkaly. Proto byly vyhodnoceny jako ohnutý chromosom. Takové chromosomy nebyly rozděleny, protože jsou překryty konci chromosomů, což je možné vidět na obrázku 40b. V místě dotyku je šířka chromosomu i intenzita srovnatelná se zbylými částmi chromosomu. Pokud došlo k selhání algoritmu, je možný manuální zásah a chromosomy tak rozdělit ručně.



Obrázek 40: Nesprávná segmentace a separace chromosomů

## 5.2 Vyhodnocení a diskuze klasifikace obrazu

Klasifikace chromosomů je silně závislá na předchozí segmentaci obrazu. Pokud dojde ke ztrátě dat u některých chromosomů, dojde k chybnému párování. U snímků z obou databází byla počítána úspěšnost párování dvou homologních chromosomů a úspěšnost správného určení pohlaví. Do výpočtu úspěšnosti nebyly zahrnuty snímky, které neobsahovali 46 chromosomů, nebo část chromosomu chyběla. Úspěšnost klasifikace je zobrazena v tabulce 2.

Tabulka 2: Úspěšnost klasifikace chromosomů

	Úspěšnost párování [%]	Úspěšnost určení pohlaví [%]
Q pruhování	78,6	83,2
G pruhování	76,7	80,9

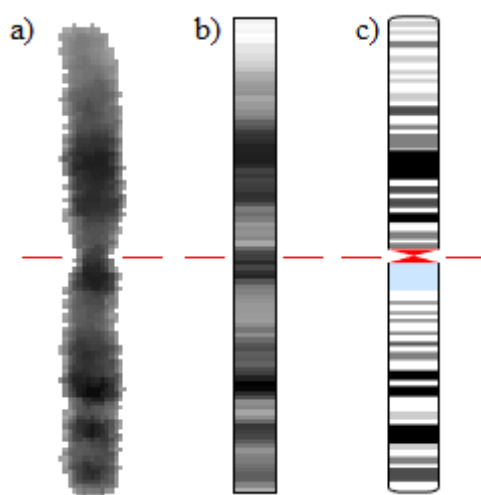
Úspěšnost párování dvou homologních chromosomů je silně ovlivněna několika faktory. Prvním je nízké rozlišení pruhů ve snímcích. U snímků není uvedeno ani rozlišení pruhů, tudíž není možné automatické srovnání s ideogramem. Kdyby všechny snímky měly pevně dané rozlišení pruhů, byla by možná analýza s ideogramem a úspěšnost by byla vyšší. Dalším faktorem je využití pouze dvou klasifikačních znaků. V algoritmu je využit obsah chromosomů a jasový profil chromosomu. Pozice centromery ani délka chromosomu není využita. Pozice centromery není v algoritmu využita z důvodu absence zúžených míst na chromosomech (úspěšnost nalezení byla pod 50%). Délka není využita, z důvodu lepších výsledků při zjištění obsahu chromosomu.

Karyotyp se správně dělí do 7 skupin právě podle délky chromosomu a pozice centromery. Jelikož detekce pozice centromery dosahovala nízké úspěšnosti, bylo rozdělení do 7 skupin modifikováno na rozdělení pouze podle obsahu chromosomů. Tento klasifikační znak se počítá ze segmentovaného chromosomu, který může být nesprávně segmentován a může docházet ke špatnému zařazení do skupiny. Tudíž v jedné skupině může být špatně zařazen chromosom z jiné skupiny a poté dochází ke špatnému párování ve zbylých skupinách.

V jednotlivých skupinách dochází k párování homologních chromosomů pomocí kvadratické odchylky signálu, která je váhována podílem obsahů. Tento vztah se ukázal jako nejúspěšnější při párování. V práci bylo použito i párování chromosomů podle korelací dvou jasových profilů, ale výsledky byly horší než využitá metoda (úspěšnost párování homologních chromosomů byla pod 65%).

Vyhodnocení karyotypu se zaměřuje pouze na zdravý karyotyp. Pokud je před klasifikací segmentováno 45 chromosomů, je vypsáno, že se jedná o monosomii, ale už není určeno, který chromosom není obsažen. To stejně platí i při segmentaci 47 chromosomů. Dojde k vypsání, že se jedná o trisomii, ale není už určeno, o který chromosom je obsažen třikrát. Monosomie a trisomie není v práci zjišťována, protože databáze neobsahují snímky s tímto druhem onemocnění a algoritmus by nemohl být řádně testován.

V práci jsem se dále zaměřil na výpočet ideogramu chromosomů. Ideogramy jsou počítány z mediánových jasových hodnot v chromosomu. Intenzity v ideogramu jsou následně upraveny pomocí ekvalizace histogramu z důvodu zvýraznění pruhů. Porovnáním těchto ideogramů s normalizovanými ideogramy nám určuje strukturní aberace. Snímky z databázi ale obsahují snímky chromosomů s nízkým rozlišením, a proto není tento druh analýzy možný. Srovnání vytvořených ideogramů s normalizovanými je možný vidět na obrázku 41. Na obrázku 41a můžeme vidět chromosom s červeně vyznačenou centromerou. Na obrázku 41b jeho ideogram, který je počítán ekvalizace histogramu z mediánových jasových hodnot. Obrázek 41c zobrazuje normalizovaný ideogram. Z obrázku je patrné, že přechody mezi jednotlivými pruhy nejsou tak zřetelné jak pruhy u normalizovaného ideogramu a detekce strukturních aberací tak není možná.



Obrázek 41: Chromosom s ideogramy

a) obraz chromosomu segmentovaný z obrazu, b) ideogram tvořený z jasových profilů, c) normalizovaný ideogram

## 6 ZÁVĚR

Cílem této práce bylo navrhnout a zrealizovat algoritmus pro obrazovou analýzu mitotických chromosomů. V úvodních kapitolách je popsána problematika, která se zaměřuje na cytogenetické metody, metody identifikace chromosomů a odlišná genetická onemocnění. Dále jsou s prací popsány postupy a techniky zpracování obrazu mitotických chromosomů.

Metody pro zobrazení chromosomů dělíme na pruhovací techniky a molekulární techniky. Práce se zaměřuje na analýzu chromosomu pomocí pruhovacích technik. Konkrétně se jedná o G pruhovací a Q pruhovací techniku. Tyto metody barvení nám poskytnou informaci o numerických a strukturních aberacích.

Navrhnutý algoritmus je složen z pěti modulů a je implementován v prostředí Matlab. První modul slouží k předzpracování obrazu. Druhý až čtvrtý modul slouží k segmentaci chromosomů od pozadí, segmentaci překrývajících se a dotýkajících se chromosomů. Pátý modul se zaměřuje na klasifikaci chromosomů do lidského karyotypu.

Segmentace chromosomů od pozadí (druhý modul) je založen na vícefázové detekci, která vychází z metody Otsu a technice aktivních kontur. Algoritmus dále řeší segmentaci překrývajících se chromosomů (třetí modul) pomocí výpočtu středové linie. Při překryvu středové linie dojde k označení, že se jedná o dva či více překrývajících se chromosomů. Čtvrtý modul algoritmu je separace dotýkajících se chromosomů. Chromosomy jsou rozděleny na základě nižší šířky a vyšší intenzity v místě dotyku. Tímto krokem získáme osamocené chromosomy, které můžeme analyzovat.

Pátý modul využívá k analýze dvou klasifikačních znaků. Prvním klasifikačním znakem je obsah chromosomů a druhým jasový profil chromosomu. Pomocí obsahu chromosomů jsou chromosomy rozděleny do sedmi skupin. V každé skupině dochází k párování homologních chromosomů pomocí hledání minima kriteriální funkce. Jakmile jsou všechny chromosomy spárovány, dochází k vykreslení lidského karyotypu. Součástí analýzy obrazu je i určení pohlaví.

Algoritmus je doplněn o uživatelské grafické rozhraní, které zpřehledňuje algoritmus. V případě špatné segmentace může dojít k manuální opravě a správné separaci chromosomu. Robustnost algoritmu byla testována na snímcích ze dvou databází, které obsahují snímky mitotických chromosomů získané G pruhováním a Q pruhováním. Tyto databáze obsahují celkem přes 150 snímků s obrazy mitotických chromosomů. Výsledky segmentace jsou určeny pomocí senzitivity a pozitivní

prediktivní hodnoty. Senzitivita segmentace pro Q pruhování je 89,2%, a pro G pruhování je 86,4%. Pozitivní prediktivní hodnota je pro Q pruhování 91,3% a pro G pruhování 92,4%. Úspěšnost párování chromosomů je pro Q pruhování 78,6% a pro G pruhování 76,7%. Úspěšnost určení pohlaví činí pro Q pruhování XX 83,2 % a G pruhování 80,9%.

Navrhnutý algoritmus zvládá vyhodnotit, zda se jedná o numerickou aberaci, nebo o zdravého jedince. Strukturní aberaci není možné vyhodnotit, protože snímky mají velkou variabilitu kvality a také rozlišení pruhů. Právě podle počtu pruhů se chromosomy porovnávají s ideogramy a tím vyhodnocují strukturní aberace.

Mezi hlavní přínosy navrženého algoritmu patří možnost zpracovávat dva typy obrazu, dalším přínosem je vyhodnocení lidského karyotypu s dostatečnou přesností. Vyšší úspěšnosti by mohlo být docíleno normalizací obrazu mitotických chromosomů.

# LITERATURA

- [1] ZEMANOVÁ, Z. MICHALOVÁ, K. In: *Klasická a molekulární cytogenetika v klinické praxi. Klinická biochemie a metabolismus* [online]., 2005. [cit. cit. 2016-11-09] Dostupné z: <http://nts.prolekare.cz/cls/odkazy/KBM0502-63.pdf>
- [2] Cytogenetika. In: *Cytogenetika* [online]. 2012 [cit. 2016-11-09]. Dostupné z: <https://web.natur.cuni.cz/~muncling/CYTO.DOC>
- [3] Chromozom. In: *Chromozom* [online]. 2014 [cit. 2016-11-09]. Dostupné z: <http://referaty-seminarky.cz/chromozom/>
- [4] ŠÍPEK Antonín. Genetika - biologie. In: *Chromozomy* [online]. 2010 [cit. 2016-11-19]. Dostupné z: <http://referaty-seminarky.cz/chromozom/>
- [5] ŠÍPEK, Antonín. Genetika - biologie. In: *Karyotyp člověka* [online]. 2010 [cit. 2016-11-19]. Dostupné z: <http://www.genetika-biologie.cz/karyotyp-cloveka>
- [6] Karyotyp člověka. In: *Karyotyp člověka* [online]. 2012 [cit. 2016-11-19]. Dostupné z: [http://is.muni.cz/do/1499/el/estud/prif/ps09/genetika/web/pdf/karyotyp\\_cloveka\\_tisk.pdf](http://is.muni.cz/do/1499/el/estud/prif/ps09/genetika/web/pdf/karyotyp_cloveka_tisk.pdf)
- [7] RIEGEL, M., *Human molecular cytogenetics: From cells to nucleotides*. Genet MolBiol, 2014, 194-209
- [8] DOUGLAS, L.N., A.B. MCGUIRE, A.M. MANZARDO a M.G. BUTLER. *High-resolution chromosome ideogram representation of recognized genes for bipolar disorder*. GENE. 2016, 136-147. DOI: 10.1016/j.gene.2016.04.011. ISSN 0378-1119.
- [9] PRITCHARD, DORIAN J. a BRUCE R. KORF. *Základy lékařské genetiky*. 1.. vydání. Praha: Galén, 2007. 182 s. s. ISBN 978-80-7262-449-2.
- [10] HANÁKOVÁ. In: *Vrozené chromosomové aberace* [online]. 2009 [cit. 2016-11-24]. Dostupné z: <http://telemedicina.med.muni.cz/pdm/genetika/res/f/2-vrozene-chromosomove-aberace.pdf>
- [11] KUGLÍK P.: *Vybrané kapitoly z cytogenetiky*, Masarykova univerzita v Brně, vydání, 2000, ISBN 80-210-2334-1
- [12] HANÁKOVÁ. In: *Získané chromosomové aberace* [online]. 2009 [cit. 2016-11-30]. Dostupné z: <http://telemedicina.med.muni.cz/pdm/genetika/res/f/3-ziskane-chromosomove-aberace.pdf>
- [13] KASHEVAROVA, A., a I. N. LEBEDEV. *Genomic architecture of human chromosomal diseases*. RUSSIAN JOURNAL OF GENETICS. 2016, 447-462. DOI: 10.1134/S1022795416040062. ISSN 1022-7954.
- [14] HANÁKOVÁ. In: *Příprava preparátů metodami klasické cytogenetiky* [online]. 2009 [cit. 2016-12-04]. Dostupné z: <http://telemedicina.med.muni.cz/pdm/genetika/res/f/4-priprava-preparatu-metodami-klasicke-cytogenetiky-hodnoceni.pdf>
- [15] O'CONNOR, C. *Karyotyping for Chromosomal Abnormalities*. Nature Education. 2008, [cit.2016-12-17]. Dostupné z: <http://www.nature.com/scitable/topicpage/karyotyping-for-chromosomalabnormalities-298>
- [16] JANEČKOVÁ. In: *Lidský genom* [online]. 2011 [cit. 2016-12-17]. Dostupné z: <http://slideplayer.cz/slide/2862713/>
- [17] HANÁKOVÁ. In: *Metody klasické cytogenetiky* [online]. 2009 [cit. 2016-12-17].

- Dostupné z:  
<http://telemedicina.med.muni.cz/pdm/genetika/res/f/6-metody-klasicke-cytogenetiky.pdf>
- [18] MICHALOVÁ, K. Encyklopedie laboratorních medicíny pro klinickou praxi. In: *Metody klasické cytogenetiky*. 2006 [cit. 2016-12-19]. Dostupné z: <http://www.enclabmed.cz/encyklopedie/A/KMABF.htm>
- [19] MICHALOVÁ, K. Encyklopedie laboratorních medicíny pro klinickou praxi. In: *Molekulární cytogenetika*. 2006 [cit. 2016-12-19]. Dostupné z: <http://www.enclabmed.cz/encyklopedie/A/KMAAF.htm>
- [20] KOČÁREK, E. *Molekulární biologie v medicíně*. 1. vydání. Brno: Národní centrum ošetrovatelství a nelékařských zdravotnických oborů v Brně, 2007. 218 s. ISBN 978-80-7013-450-4.
- [21] IVF1. In: *Comparative Genomic Hybridization* [online]. 2013 [cit. 2016-12-20]. Dostupné z: <http://www.ivf1.com/cgh-comparative-genomic-hybridization/>
- [22] LERNER, B., *Toward a completely automatic neural-network-based human chromosome analysis*. IEEE Trans Syst Man Cybern B Cybern, 1998. 28(4): p. 544-52.
- [23] CAO, Hongbao, Hong-Wen DENG a Yu-Ping WANG. *Segmentation of M-FISH Images for Improved Classification of Chromosomes With an Adaptive Fuzzy C-means Clustering Algorithm*. IEEE TRANSACTIONS ON FUZZY SYSTEMS. 2012, 1-8. DOI: 10.1109/TFUZZ.2011.2160025. ISSN 1063-6706.
- [24] SOMASUNDARAM, D a V.R.V. KUMAR. *Straightening of highly curved human chromosome for cytogenetic analysis*. MEASUREMENT. 2014, 880-892. ISSN 0263-2241.
- [25] VŠETIČKA, V. *Digitální filtry pro obrazová data*, Brno: Vysoké učení technické, Fakulta elektrotechniky a komunikačních technologií, 2010. 30 s. Vedoucí bakalářské práce Ing. Petr Číka, Ph.D.
- [26] HUANG, G. Yang, and G. Tang, *A fast two-dimensional median filtering algorithm*, IEEE Trans. Acoust., Speech, Signal Processing, vol. 27, no. 1, pp. 13–18, 1979.
- [27] TSU N., "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man, and Cybernetics, Vol. 9, No. 1, 1979, pp. 62-66.
- [28] KASS, M., WITKIN, A. a TERZOPOULOS, D. *Snakes: Active contour model*. 1988, Palo Alto, USA. ISSN: 1573-1405.
- [29] NJMAN, L. a M. SCHMITT. *Watershed of a continuous function*. In *Signal Processing* (Special issue on Mathematical Morphology.), Vol. 38 (1994), pages 99–112
- [30] TAO W.B., J.W. TIAN a J. LIU. *Image segmentation by three-level thresholding based on maximum fuzzy entropy and genetic algorithm*. PATTERN RECOGNITION LETTERS. 2003, 3069-3078. ISSN 0167-8655.
- [31] GAJENDRAN V., RODRI, X a J. J. GUEZ. *Chromosome counting via digital image analysis*. In *Image Processing*, 2004. ICIP '04. 2004 International Conference on. 2004, vol. 5, p. 2929-2932 Vol. 2925.
- [32] GRISAN, E., A. RUGGERI a E. POLETTI. *Automatic Segmentation and Disentangling of Chromosomes in Q-Band Prometaphase Images*. IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE. 2009. DOI: 10.1109/TITB.2009.2014464.
- [33] MADIAN, N., K.B. JAYANTHI a S. SURESH. *Contour Based Segmentation of Chromosomes in G-Band Metaphase Images*. IEEE GLOBAL CONFERENCE ON

- SIGNAL AND INFORMATION PROCESSING*. 2015, 943-947. ISBN 978-1-4799-7591-4.
- [34] GRISAN, E., E. POLETTI, C. TOMELLERI a A. RUGGERI. *Automatic segmentation of chromosomes in Q-band images*. *ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY*. 2007, 5513-5516. DOI: 10.1109/IEMBS.2007.4353594. ISSN 1094-68X.
- [35] GRISAN, E., A. PESCE, A. GIANI, M. FORACCHIA, and A. RUGGERI. "A new tracking system for the robust extraction of retinal vessel structure," in *Proc. 26 th Annual International Conference of IEEE-EMBS*, (New York), pp. 1620–1623, IEEE, 2004.
- [36] SOILLE, P. *Morphological image analysis*. Berlin: Springer Verlag, 2003, ISBN 3-540-42988-3, 391 s.
- [37] SOMASUNDARAM, D. a V KUMAR. *Separation of overlapped chromosomes and pairing of similar chromosomes for karyotyping analysis*. *MEASUREMENT*. 274-281, 2014. DOI: 10.1016/j.measurement.2013.11.024. ISSN 0263-2241
- [38] UHLMANN, V, R. DELGADO-GONZALO, M. UNSER, P.O. MICHEL a L. BALDI. *USER-FRIENDLY IMAGE-BASED SEGMENTATION AND ANALYSIS OF CHROMOSOMES*. *IEEE SYMPOSIUM ON BIOMEDICAL IMAGING (ISBI)*. 2016, 395-398. ISSN 1945-7928.
- [39] JAN, J. *Medical image processing, reconstruction and restoration: concepts and methods*. Boca Raton: Taylor, 2006, 730 s. ISBN 08-247-5849-8.
- [40] DENZA, E., DENZA Mi. M., *Encyclopedia of Distances*, Springer p. 94,2009
- [41] COUTSIAS E.A., SEOK C, DILL K.A. "Using quaternions to calculate RMSD". *J Comput Chem*. 25 (15).2004: 1849–1857. doi:10.1002/jcc.20110. PMID 15376254.
- [42] BENCKO, Vladimír, et al. *Epidemiologie, výukové texty pro studenty 1. LFUK, Praha*. 2. vydání. Praha Univerzita Karlova v Praze, 2002. 168 s. s. 77-78. ISBN 80-246-0383-7.



## ZDROJE OBRÁZKŮ

- [1] <https://santhoshproject.wordpress.com/2014/11/16/human-cells/>
- [2] <http://mmp.vfu.cz/opvk2014/?title=teorie-cytogenetika>
- [3] <http://www.genetika-biologie.cz/karyotyp-cloveka>
- [4] [http://www.wikiskripta.eu/index.php/Soubor:Chromosomes\\_mutations-en.svg](http://www.wikiskripta.eu/index.php/Soubor:Chromosomes_mutations-en.svg)
- [5] <http://www.baso.com.cn/en/view.asp?id=164>
- [6] [http://www.wikiskripta.eu/index.php/Identifikace\\_chromozom%C5%AF](http://www.wikiskripta.eu/index.php/Identifikace_chromozom%C5%AF)
- [7] <http://www.pathology.washington.edu/galleries/Cytogallery/main.php?file=human%20karyotypes>
- [8] <http://physrev.physiology.org/content/88/2/557>
- [9] <http://www.ivf1.com/cgh-comparative-genomic-hybridization/>
- [10] <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7418336&tag=1>
- [11] <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=477319>
- [12] <http://bioimlab.dei.unipd.it/Automatic%20Karyotyping%20Segmentation.htm>
- [13] snímky poskytnuté vedoucí DP z laboratoře

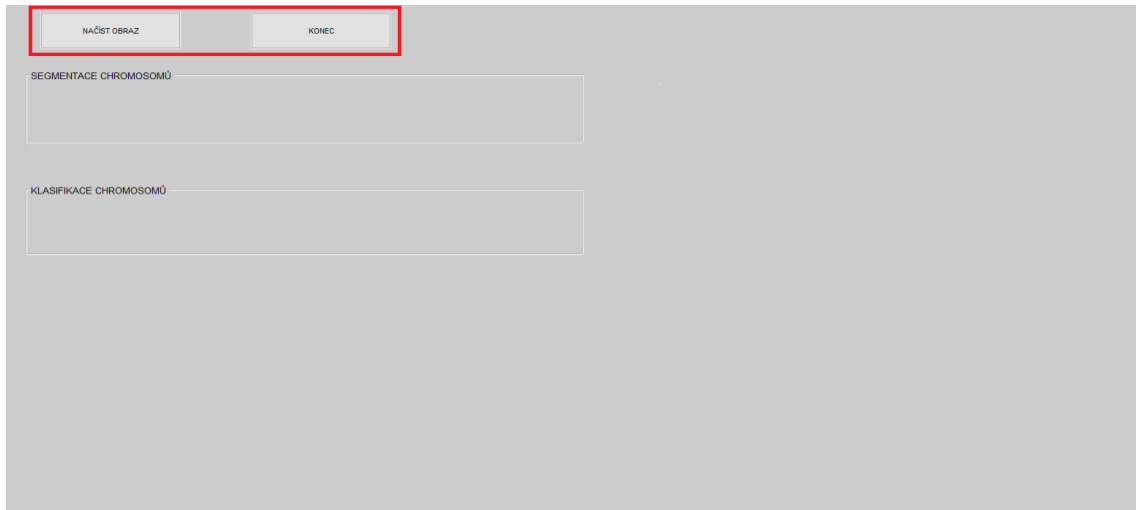
# PŘÍLOHY

## A. Složky a soubory uložené na CD

- Lubos\_Jaros\_DP.pdf – text diplomové práce.
- Program – složka obsahuje skripty a funkce vytvořeného algoritmu spolu s testovacím obrazem.

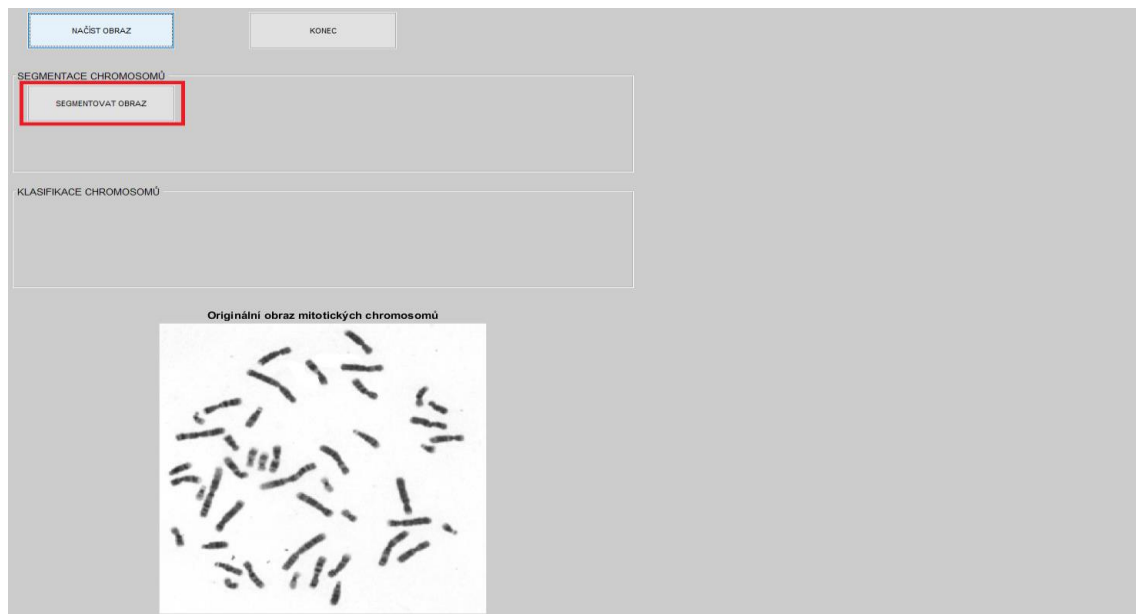
## B. Návod k GUI pro vyhodnocení lidského karyotypu

Všechny programové části byly vytvořeny a otestovány v prostředí Matlab R2015a na displeji s rozlišením 1920x1200. Program se spouští v prostředí Matlab pomocí skriptu `ObrazovaAnalyza_GUI.m`. Po spuštění se objeví okno se dvěma tlačítky, které je vidět na obrázku 42. Tlačítko **NAČÍST OBRAZ** spouští dialogové okno, kde je nutné si vybrat obraz k analýze. Druhé tlačítko **KONEC** ukončuje celý program i prostředí Matlab.



Obrázek 42: GUI po spuštění

Po načtení obrazu se zobrazí tlačítko **SEGMENTOVAT OBRAZ** v ovládací části segmentace obrazu, které je možné vidět na obrázku 43. Tlačítko spouští automatickou segmentaci obrazu a separaci překryvů a dotyků chromosomů.

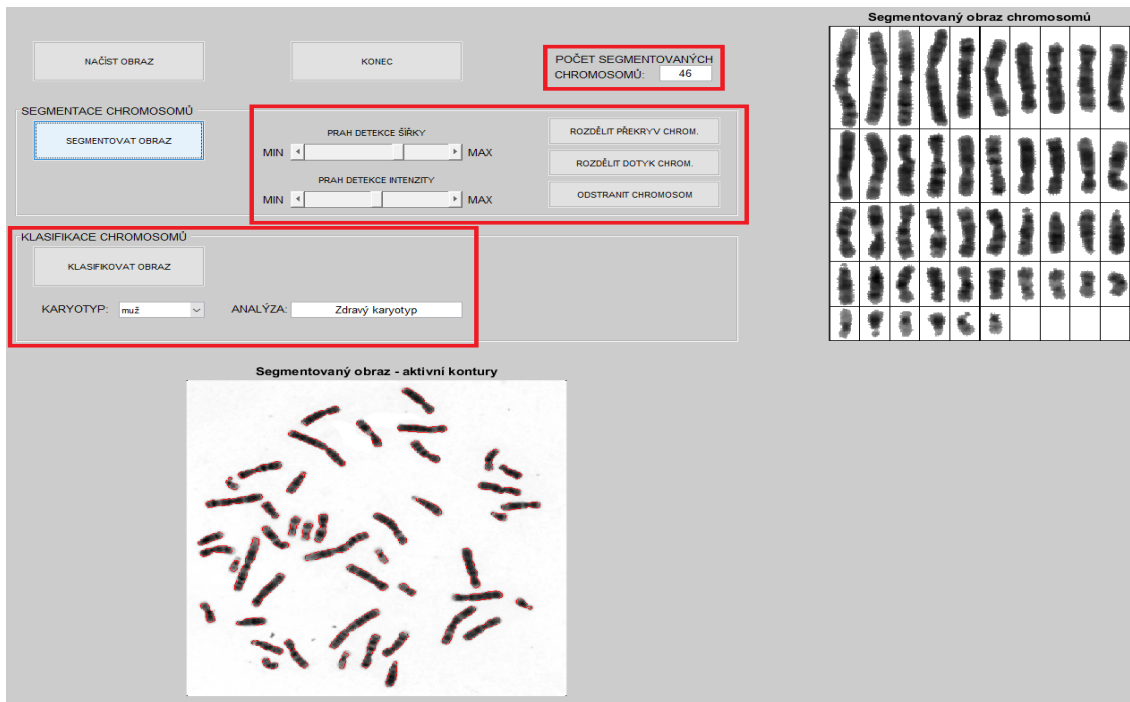


Obrázek 43: GUI po načtení obrazu

Po segmentaci obrazu se vykreslí segmentovaný obraz a zobrazí další ovládací prvky, které je vidět na obrázku 44. Pole **POČET SEGMENTOVANÝCH CHROMOSOMŮ** vypisuje hodnotu, kolik bylo celkem segmentováno chromosomu.

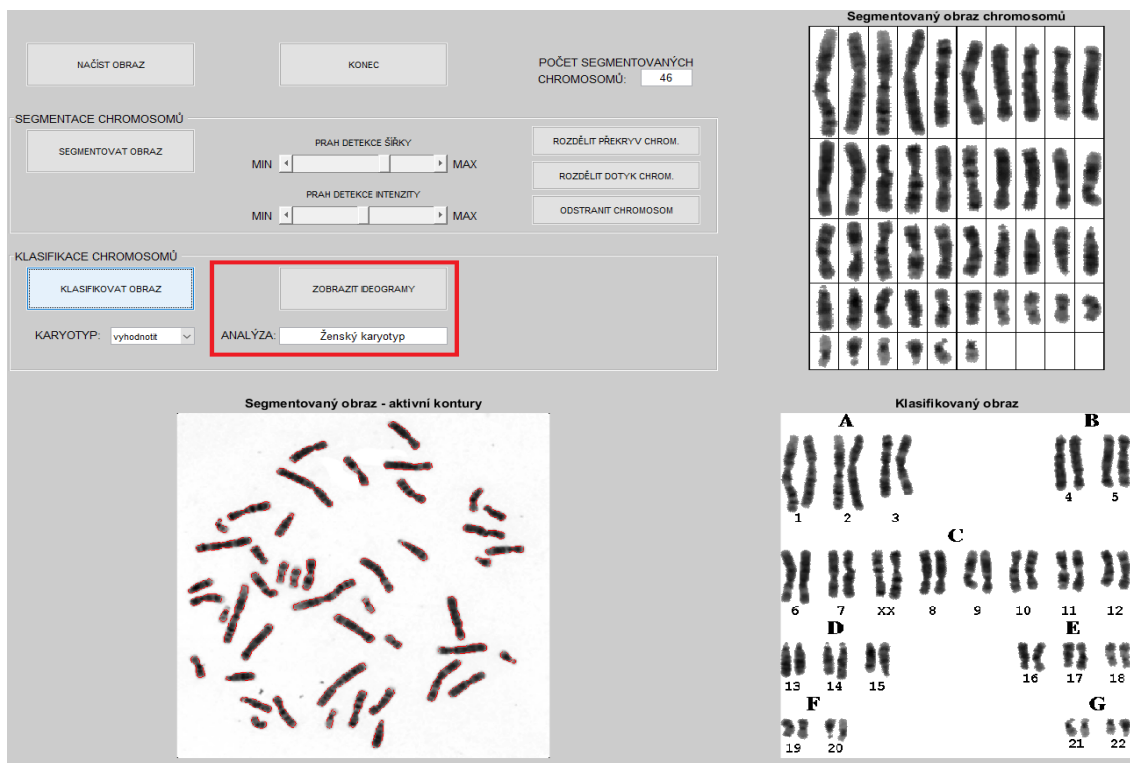
V segmentační části se zobrazí dva posuvníky a tři tlačítka. Posuvníky **PRAH DETEKCE ŠÍŘKY** a **PRAH DETEKCE INTENZITY** slouží k posouvání prahů při chybné separaci dotyků chromosomů. Maximální hodnota prahu detekce šířky je 1, minimální hodnota šířky je 0, výchozí hodnota je 0,66. Maximální hodnota prahu intenzity chromosomu je 3, minimální hodnota intenzity je 0, výchozí hodnota je 1,5. Při změně posuvníků je nutné znovu kliknout na tlačítko **SEGMENTOVAT OBRAZ**. Tlačítko **ROZDĚLIT PŘEKRYV V CHROMOSOMU** slouží k separaci překryvů. Po kliknutí na tlačítko se objeví kurzor, kterým musí obsluha kliknout na střed překrývajících se chromosomů v obraze segmentovaných chromosomů. Chromosomy se automaticky rozdělí a obraz segmentovaných chromosomů překreslí. Tlačítko **ROZDĚLIT DOTYK CHROMOSOMŮ** slouží k separaci dvou dotýkajících se chromosomů. Po kliknutí se objeví rovněž kurzor, kde musí obsluha vybrat dva body v obraze. Body tvoří pomyslnou přímkou, kterou bude veden řez. Tento řez rozdělí chromosom ve dva a opět vykreslí do segmentovaného obrazu. Posledním tlačítkem se segmentační části je **ODSTRANIT CHROMOSOM**. Opět se objeví kurzor, který odstraní vybraný chromosom. Tlačítko slouží k odstranění nečistot v obraze.

V klasifikační části se zobrazí tlačítko, pole a výběrový seznam. V poli **ANALÝZA** se po segmentaci vypíše, zda se jedná o zdravý karyotyp, monosomii nebo trisomii. Obsluha musí zkontrolovat, zda segmentace proběhla úspěšně. Ve výběrovém poli **KARYOTYP** je na výběr ze tří možností (muž, žena, vyhodnocení). Vyhodnocení slouží k určení pohlaví. Tlačítko **KLASIFIKACE OBRAZU** slouží k zařazení lidského karyotypu. Klasifikace je možná jenom pokud se jedná o zdravý karyotyp. Pokud je ve výběrovém poli **KARYOTYP** vybráno vyhodnocení, tak se do pole **ANALÝZA** vypíše pohlaví jedince. Dále se vykreslí obraz chromosomů zařazených do sedmi skupin.

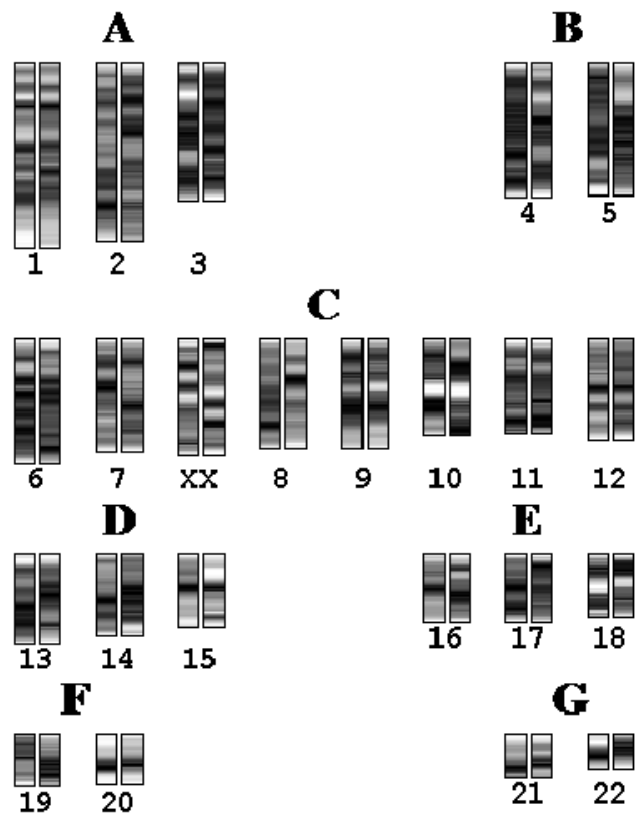


Obrázek 44: GUI po segmentaci obrazu

Po klasifikaci (obrázek 45) dojde k zobrazení posledního tlačítka **ZOBRAZIT IDEOGRAMY**. Tlačítko slouží k vykreslení ideogramů v novém okně (obrázek 46).



Obrázek 45: GUI po klasifikaci obrazu



Obrázek 46: Ideogramy chromosomů