



**ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE**

---

**Fakulta elektrotechnická  
Katedra radioelektroniky**

**Optická kontrola měděných originálů mechanického záznamu zvuku**

**Video inspection of master copper discs for mechanical sound recording**

Diplomová práce

Studijní program: Komunikace, multimédia a elektronika (magisterský)

Studijní obor: Multimediální technika (magisterský)

Vedoucí práce: Ing. Karel Fliegel, Ph.D.

**Bc. Martin Mohl**

---

**Praha 2017**



České vysoké učení technické v Praze  
Fakulta elektrotechnická  
katedra radioelektroniky

## ZADÁNÍ DIPLOMOVÉ PRÁCE

Student: **Martin Mohl**

Studijní program: Komunikace, multimédia a elektronika  
Obor: Multimediální technika

Název tématu: **Optická kontrola měděných originálů mechanického záznamu zvuku**

Pokyny pro vypracování:

Podejte přehled metod pro optické snímání signálu zaznamenaného na mechanických zvukových nosičích. Navrhněte systém pro bezkontaktní optické snímání drážky z měděných originálů pomocí kamery a algoritmy pro následný převod kontur drážky na zvukový signál. Analyzujte parametry snímacího systému a ověřte funkčnost navržených postupů aspoň na úrovni počítačové simulace.

Seznam odborné literatury:

- [1] Dušek, K.: Záznam a reprodukce zvuku, ČVUT, 1982.
- [2] Šonka, M., Hlaváč, V., Boyle, R.: Image processing, analysis, and machine vision, Thomson, 2008.

Vedoucí: Ing. Karel Fliegel, Ph.D.

Platnost zadání: do konce letního semestru 2016/2017

doc. Mgr. Petr Páta, Ph.D.  
vedoucí katedry



prof. Ing. Pavel Ripka, CSc.  
děkan

V Praze dne 19. 2. 2016



## **Poděkování**

Rád bych poděkoval vedoucímu práce Ing. Karlu Fliegelovi, Ph.D. za rady, tipy a doporučení ohledně koncepce celé práce a zejména za velkou a nutnou dávku trpělivosti. Dále děkuji kolektivu oddělení elektroakustiky a vedení firmy GZ Media za spolupráci při tvorbě zadání práce a zejména za umožnění využití dostupné techniky a přípravu měřicího pracoviště.

V neposlední řadě děkuji své rodině za morální podporu a toleranci stresových situací v průběhu studia a několika blízkým přátelům za cenné rady a zejména duševní podporu.



## **Čestné prohlášení**

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o etické přípravě vysokoškolských závěrečných prací.

V Praze dne 6.ledna 2017

.....  
(Martin Mohl - podpis)





## Abstrakt

Tato práce se zabývá problematikou měděných originálů mechanického záznamu zvuku, jakožto i stručným popisem samotné technologie mechanického záznamu a postupu jeho výroby. Hlavní část práce pak je věnována výzkumu možností optického snímání těchto originálů. Teoretická část se zabývá současně používanými technikami optického čtení mechanického záznamu a omezeními, které nám přináší kombinace velmi jemného mechanického zápisu s technickými omezeními snímacích zařízení. Cílem praktické části pak je na základě výzkumu současných technik navrhnout vlastní metodu sejmutí drážky a následná její implementace včetně tvorby algoritmu, který umožní co nejkvalitnější rekonstrukci zvukové stopy pro účel automatizované detekce chyb a problémů záznamu. Tato práce je vypsána ve spolupráci s GZ Media, a.s.<sup>1</sup>

### Klíčová slova:

Gramofonová deska, extrakce zvuku, mikrofotografie, detekce hran, analýza obrazu

## Abstract

The thesis deals with the issue of copper master discs of mechanical sound recording and briefly describes the technology itself and production process of vinyl records. The main part of the thesis is dedicated to options of optical reading of the discs. The theoretical part of the thesis gives a literature survey of methods that were implemented and tested in the past and restrictions given by the combination of very fine mechanical recording with technical restrictions of the reading devices. The goal of practical part is to design suitable method based on the survey. The developed method is implemented so it can reconstruct audio from series of images represented short part of a record. The topic is in collaboration with GZ Media, a.s.

### Keywords:

Vinyl record, sound extraction, microphotography, edge detection, image analysis

---

<sup>1</sup> <http://www.gzvinyl.com/> (cit. 6.1.2017)



# Obsah

1 Úvod .....	13
2 Popis technologie mechanického záznamu zvuku .....	15
2.1 Postup výroby gramofonové desky .....	15
2.2 Mechanický záznam .....	16
2.3 Vlastnosti drážky a mechanického záznamu .....	17
2.3.1 Záznamové charakteristiky.....	18
2.3.2 Princip stereozáznamu .....	19
2.4 Možné vady drážky a jejich důsledky .....	21
2.5 Metody čtení mechanického záznamu.....	21
3 Bezkontaktní snímání mechanického záznamu.....	23
3.1 Metody využívající laserového paprsku .....	23
3.2 Metody využívající snímání obrazu .....	25
3.3 Pokročilé metody využívající mikrofotografie .....	27
3.3.1 VisualAudio concept .....	29
4 Převod snímku drážky na zvukovou informaci.....	33
4.1 Rekonstrukce mono- a stereosložky signálu .....	33
4.2 Přímá rekonstrukce levého a pravého kanálu.....	34
4.3 Základní technická omezení při snímání drážky.....	34
4.3.1 Snímání rastrovou kamerou .....	35
4.3.2 Snímání řádkovou kamerou.....	36
4.3.3 Srovnání obou metod z hlediska parametrů.....	37
5 Ostatní teoretické principy .....	39
5.1 Principy detekce hran v obraze .....	39
5.1.1 K čemu je detekce hran vhodná a proč ji použít .....	39
5.1.2 Hrana z hlediska matematiky .....	40
5.1.3 Gradient .....	40
5.1.4 2D konvoluce .....	41
5.1.5 Algoritmy detekce hrany.....	41
5.2 Krokové motory a jejich řízení .....	43
5.2.1 Řízení krokového motoru pomocí Arduina a H-můstku.....	44
6 Obecný návrh experimentálního zařízení .....	47
6.1 Princip konstrukce externího pohonu talíře .....	47
6.2 Výpočet převodu pohonného mechanismu .....	48
6.3 Realizace převodu pohonného mechanismu.....	49
6.4. Posuv kamery.....	51
6.4.1 Páskový převod řízený krokovým motorem.....	52
6.4.2 Využití závitové tyče řízené krokovým motorem.....	52

6.4.3	Využití modulačního signálu gramofonové desky.....	53
6.5	Osvětlení drážky .....	53
6.5.1	Návrh osvětlovací aparatury .....	54
7	Sběr zdrojových obrazových dat .....	56
8	Detekce pozice drážky z obrazu .....	58
8.1	Určení překryvu snímků .....	59
8.2	Detekce polohy drážky .....	61
8.2.1	Virtuální sledování chybějící hrany .....	62
8.2.2	Detekce dotyků, prořezů a oprava kritických míst .....	63
9	Tvorba zvukového souboru.....	66
9.1	Převod množiny matic na dvojici vektorů .....	66
9.2	Další zpracování zvuku.....	68
9.2.1	Interpolace chybějících vzorků .....	68
9.2.2	Převzorkování .....	68
9.2.3	Frekvenční úpravy a finalizace .....	69
9.3	Vlastnosti zvukového výstupu .....	69
9.3.1	Spektrální vlastnosti .....	70
9.3.2	Zkreslení zvuku .....	71
10	Simulace snímání řádkovou kamerou .....	73
10.1	Frekvenční úprava audiostop.....	74
10.2	Simulace geometrických vlastností drážky .....	75
10.2.1	Obraz nasimulované drážky .....	76
10.2.2	Simulace degradace obrazu.....	76
10.2.3	Simulace čtení řádkovou kamerou a detekce hran .....	77
10.3	Rekonstrukce zvuku ze snímku drážky .....	78
10.4	Vyhodnocení rekonstruovaného zvuku .....	79
10.5	Vyhodnocení použité metody.....	79
11	Návrh na vylepšení metody a směr výzkumu .....	80
11.1	Specializované pracoviště optické kontroly .....	80
11.2	Synchronizované paralelní zpracování dat .....	81
11.3	Detekce vad.....	81
12	Závěr .....	82
13	Použitá literatura.....	85
	Seznam tabulek a obrázků.....	87
	Seznam zkratk a symbolů .....	89
	Příloha A – Obsah příloženého DVD.....	91

## Kapitola 1

# Úvod

V posledních letech velmi narůstá oblíbenost poslechu hudby prostřednictvím gramofonů. Ačkoliv technologie samotná je stará více než 100 let a princip mechanického záznamu jako takového se příliš nezměnil, výrazně se ovšem změnila potřeba konzumentů hudby z gramofonových desek. Za přežitím tohoto média totiž stojí zejména posluchači alternativní a nezávislé hudební scény. Pro gramofonovou desku, která dříve byla hojně používaná k záznamu analogových zvuků (klasická hudba, jazz), které se vyznačovaly vysokou dynamikou a malým obsahem harmonických frekvencí, to znamená nutnost pojmout zvuky tvořené elektronicky a s nízkým dynamickým rozsahem. Současně s trendem tzv. Loudness War zákazníci požadují vyšší hlasitost nahrávek. Pomocí složitých počítačových analýz se proto v dnešní době na deskách vydává hudba, která je na samotném okraji možností mechanického záznamu či dokonce až za ním. Záznam je díky tomu náchylný na zkreslení vzniklé při samotném zápisu i při přehrávání na méně kvalitní aparatuře (zejména co se týče kvality a možností přenosky).

I z tohoto důvodu je kromě pečlivých analýz a úprav signálu před výrobou originálního nosiče potřeba záznam důkladně kontrolovat i v různých fázích výroby. V některých případech není možné provést kontrolu přímo pomocí přehrávání přenoskou proto, že sama přenoska už mechanicky záznam nevratně nepatrně poškodí. Tento problém by řešilo optické snímání drážky pomocí mikroskopu a kamery. Cílem této diplomové práce je prozkoumat problematiku bezkontaktního čtení mechanických záznamových médií a analyzovat možnosti tohoto snímání.

V práci naleznete stručný úvod k problematice výroby gramofonové desky z hlediska principů a procesů, které umožňují ze znalosti mechaniky této desky zpětně rekonstruovat její zvukový obsah bez použití kontaktní přenosky (kapitoly 2 a 4). Dále práce obsahuje informace o způsobech, jakými v minulosti docházelo k experimentování s technologií bezkontaktní rekonstrukce zvuku ať už ve fázi teoretického výzkumu obecných principů, či kompletních funkčních projektů, které skutečně dokázaly převod provést na reálné desce a jako výstup poskytnout zvukový soubor se slyšitelnými a měřitelnými parametry (kapitola 3).

Cílem práce je shrnout veškeré potřebné teoretické znalosti související s touto problematikou a na základě těchto poznatků a poznatků z předešlých výzkumů pak navrhnout vlastní přístup k problematice.

Výstupem pak budou dvě navržené metody, z nichž jedna bude zpracována formou počítačové simulace bez použití reálné techniky a hardwaru, zatímco druhá metoda bude obsahovat praktickou demonstraci na reálném analogovém nosiči pomocí reálného optického systému (kapitoly 10, respektive 6-9).

Na závěr bude metoda vyhodnocena z hlediska kvality výstupu a rychlosti zpracování, tyto výsledky budou srovnány s předpoklady, které vznikly v rámci teoretického rozboru, a bude navrženo, jakým směrem by se měl případný budoucí výzkum ubírat, aby bylo možné v tomto oboru dospět k metodám, které budou použitelné i pro praktické podmínky při výrobě a kontrole gramofonových desek a nejen v oblasti teorie a předpokladů.

## Kapitola 2

# Popis technologie mechanického záznamu

Pro jednodušší pochopení cíle práce a skutečnosti, v které části celého výrobního procesu je optická kontrola užitečná, je vhodné nejprve stručně shrnout postup výroby gramofonové desky od začátku (tj. od zaslání hudebních podkladů) až po samotné sundání hotové desky z lisovacího stroje, popsat, na jakých principech je mechanický záznam postaven a jaké jsou jeho vlastnosti. Jelikož je cílem projektu snímání drážky, je důležité teoreticky popsat, čím je drážka tvořena, na čem závisí její tvar a další parametry a v neposlední řadě i to, jakými metodami se drážka tradičně snímá.

## 2.1 Postup výroby gramofonové desky

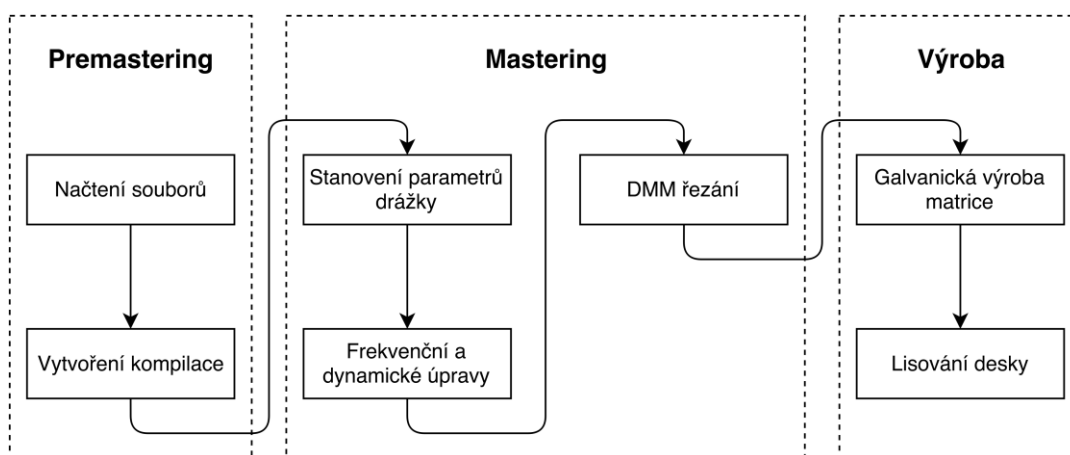
Informace v této kapitole popisují postup výroby přímo v zadávající firmě, kde osobně pracuji na pozici, na které přicházím se zmíněnými postupy do pravidelného kontaktu. Celý popsáný proces se odehrává ve specializovaném programu, který se používá pouze v rámci této jediné firmy.

V první fázi výroby operátor oddělení elektroakustiky načte do systému audionahrávky, které mohou být zaslány jak na CD či analogovém nosiči (pásce), tak prostřednictvím FTP jako sestava jednoho či více wav souborů. Pro zjednodušení se dále budeme zabývat možností, kde zákazník zaslal jednotlivé audio tracky ve zvláštních wav souborech. Tyto soubory operátor umístí do připraveného projektu do premasteringové části zmíněného softwaru a podle požadavků zákazníka mezi ně vloží či nevloží pauzy a zejména přechodové značky (VTM – visual track marker) – jedná se o obvykle cca 1 mm místa, ve kterém je pouze jedna drážka. Ve značce tedy může být nahraná hudba (například při záznamu koncertu či plynule navazujících skladbách), ale díky snížené hustotě drážek je značka opticky snadno viditelná.

Když je strana připravena tímto způsobem, pokračuje se v přípravě procesem masteringu. Zde operátor podle délky a charakteru podkladu určí vhodnou základní šířku drážky a jejich rozestup. Pokud to považuje za vhodné například z důvodu nedostatečné či naopak přílišné délky hudebního materiálu, může doporučit změnu rychlosti otáček z hodnoty požadované zákazníkem na jinou, více vhodnou z důvodu optimalizace zaplnění desky či přímo zvukovým vlastnostem. Více informací o otáčkách desky následuje v kapitole 2.3. Poté přicházejí na řadu ekvalizační a dynamické úpravy signálu. Pomocí zmíněného speciálního softwaru je v jeho části věnované masteringu připravené kompilace možné spočítat, jak bude vypadat reálná drážka a analyzovat, kde hrozí vznik

potenciálně problematických míst. Jedná se o příliš velké výchylky na vyšších kmitočtech, kde by došlo ke geometrickému zkreslení z důvodu příliš vysokých úhlů v zaznamenané drážce, místa, kde díky velké vertikální výchylce bude příliš mělká drážka, kde hrozí přeskokování přenosky a další nežádoucí jevy. Vhodným nastavením ekvalizéru, kompresoru a celkové hlasitosti je možné tyto jevy potlačit v dostatečné míře (ač za cenu ovlivnění zvuku nahrávky).

Následujícím krokem je řezání originálního nosiče na speciální aparatuře. Popisem řezání se zabývám v následující kapitole 2.2, zde se tedy omezím pouze na to, že výstupem je plech s drážkou. Jedná se o pozitiv, což znamená, že drážka je běžným způsobem přehratelná a – při zanedbání ztrát během další výroby – totožná s drážkou, která bude na vinylové desce. Z tohoto originálu se galvanickými procesy dále vyrábí negativní matrice, která je připevněna v lisovacím přístroji a za vysokých teplot použita k lisování desek. Celý tento postup je shrnut na diagramu 2.1.

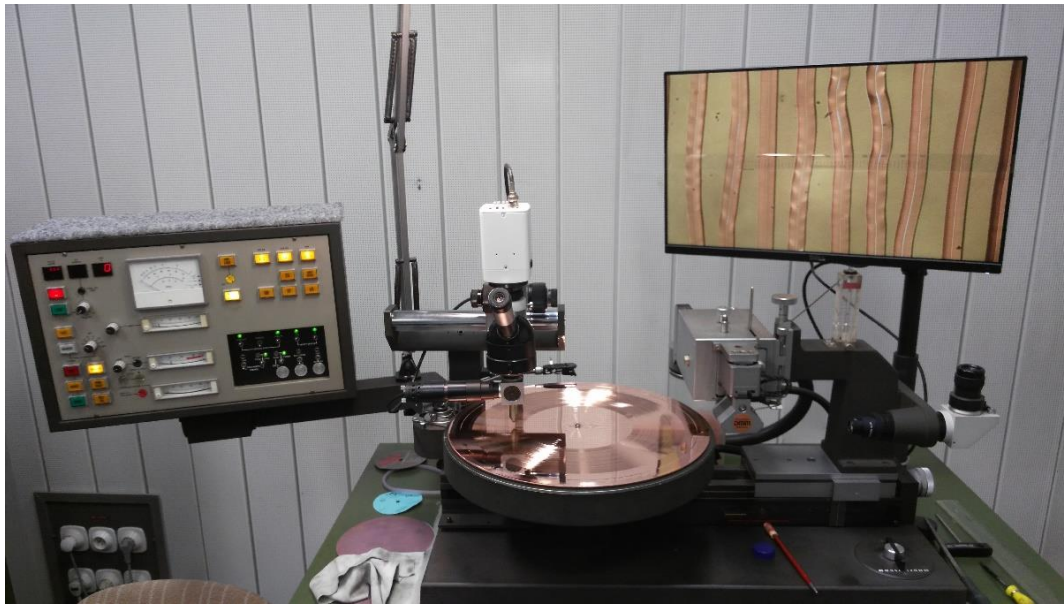


Obr. 2.1: Schéma procesu výroby gramofonové desky v GZ Media

## 2.2 Mechanický záznam

Záznam hudebního signálu do měděného originálu se v případě technologie DMM (Direct Metal Mastering) provádí na specializovaných řezacích aparaturách („řezačkách“). Podoba řezací aparatury včetně mikroskopu, kamery a kontrolního LCD monitoru je na obrázku 2.2 na následující straně. Jedná se o řezačky od německého výrobce Georg Neumann, které jsou v dnešní době ve světě již poměrně vzácné – počet vyrobených řezaček se počítá spíše v jednotkách kusů.





Obr. 2.2: DMM řezací aparatura

Základními stavebními kameny řezačky je otáčivý talíř, na který se vakuově připevní ocelový plech s galvanicky nanesenou vrstvou mědi (o tloušťce cca 100  $\mu\text{m}$ ) a řezací hlava. Jedná se o elektro-mechanický měnič, na jehož vstup je přiváděn elektrický signál, který je pomocí dvojice cívek a magnetů přiveden na řezací nůž s diamantovým hrotem, který následně provádí záznam spirálové drážky do měděné vrstvy. Záznam může být buď hloubkový, kde se nůž pohybuje ve směru kolmém k rovině desky nebo stranový, kde má drážka konstantní hloubku a výchylka je pouze horizontální. Na oba zmíněné typy záznamu by stačil řezací nůž pouze s jednou cívkou, nicméně neumožňují vícekanalový záznam. Proto je v dnešní době používán záznam kombinovaný, který umožňuje jak horizontální, tak vertikální výchylku, jejichž kombinací vzniká dvoukanalový (stereo) záznam. Princip zápisu více kanálů (stereo) bude podrobněji popsán v kapitole 2.3.2. a zároveň v [6].

## 2.3 Vlastnosti drážky a mechanického záznamu

Úhlová rychlost otáčení desky je konstantní, ale samotná drážka je spirálová. To vede k tomu, že vlnová délka zaznamenaného signálu se ke středu desky výrazně zkracuje. Pokud vezmeme v úvahu formáty desek, které se v dnešní době používají – tedy kombinace průměrů 30, 25 a 17 cm a rychlostí 33,3 a 45 otáček za minutu s tím, že maximální a minimální poloměr drážky činí 15, respektive 5 cm – vyjde nám podle vzorce (2.1), kde  $n$  je počet otáček za minutu a  $r_d$  poloměr drážky v cm že drážková rychlost se pohybuje v rozsahu 17,45 až 70,69 cm/s. Podle rovnice (2.2) pak vlnová délka signálu omezeného frekvencemi 20 Hz – 20 kHz může nabývat hodnot 8,72  $\mu\text{m}$  (pro srovnání: standardní šířka drážky je 40-50  $\mu\text{m}$ ) až 3,53 cm.

$$v_d = \frac{2\pi nr_d}{60} = \frac{\pi nr_d}{30} \text{ [cm/s]} \quad (2.1)$$

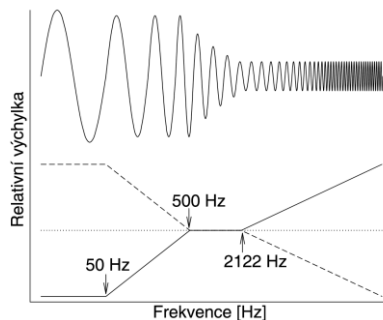
$$\lambda = \frac{v_d}{f} \text{ [cm]} \quad (2.2)$$

### 2.3.1 Záznamové charakteristiky

Běžné řezací hlavy pracují na elektromagnetickém principu, kde by při konstantní amplitudě signálu (elektrického) měla být konstantní i amplituda rychlosti rycí jehly. Jelikož platí, že při konstantní amplitudě musí s rostoucí frekvencí řezací hrot urazit mnohonásobně delší dráhu než při frekvenci nižší, vyplývá, že pro dodržení konstantní rychlosti musí být maximální výchylka řezacího hrotu (při konstantní amplitudě zvukového signálu) nepřímo úměrná frekvenci signálu. To nese hned dva problémy:

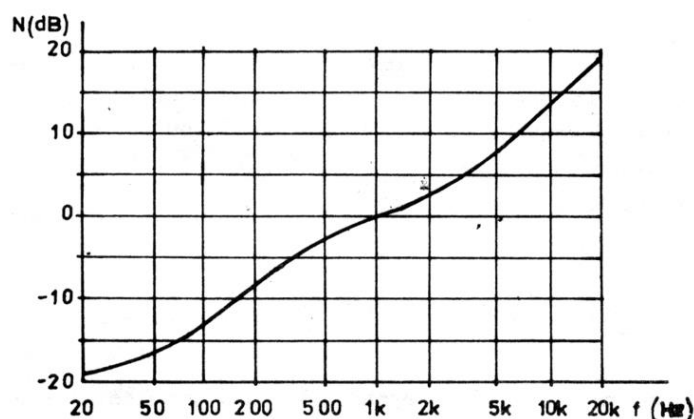
Pokud bychom chtěli tento požadavek dodržet a zároveň mít u vysokých kmitočtů dostatečnou výchylku z důvodu odstupu od šumu, musely by nízké kmitočty mít také velmi vysokou výchylku [2]. To by poté velmi významně omezilo kapacitu desky. Proto se frekvence pod 250 Hz v minulosti zapisovaly s konstantní max. výchylkou. Docílilo se toho díky snížení amplitudy signálu o 6 dB/oktávu oproti charakteristice zaručující konstantní rychlost [2].

Druhým problémem je, že i tak by se výchylka vysokých kmitočtů blížila velikosti zrn materiálu desky. Proto je naopak potřeba vyšší kmitočty patřičně zesílit (i za cenu zvýšení záznamové rychlosti, v důsledku pak i teploty systému). V roce 1953 byla přijata kombinovaná charakteristika (viz obrázek 2.3), která po úsecích obsahuje jak konstantní rychlost, tak konstantní výchylku [2]. Ta později byla převedena na dnes všeobecně uznávanou záznamová charakteristiku (tzv. RIAA - podle Recording Industry Association of America) uvedenou na obrázku 2.4, jejíž inverzní charakteristika je zabudována ve všech běžně používaných gramofonních předzesilovačích. V datasheetech těchto předzesilovačů je pak buď uvedena kompletní přehrávací charakteristika, nebo její odchylka právě od RIAA charakteristiky (která například u konkrétního výrobku TCC TC-750<sup>2</sup>, používaném ve firmě GZ Media tvoří maximálně ±0,5 dB v celém frekvenčním pásmu).



Obr. 2.3: kombinovaná záznamová charakteristika (převzato z [2])

<sup>2</sup> Informace dostupné na <http://www.phonopreamps.com/tc750pp.html> (cit. 5.1.2017)



Obr. 2.4: RIAA charakteristika (převzato z [2])

### 2.3.2 Princip stereozáznamu

Jak je naznačeno v kapitole 2.2, v dnešní době se pro tvorbu desek používá výhradně kombinovaný (stranový a hloubkový) záznam. Spočívá v tom, že řezací hlava používá dvě kolmé cívky (jednu pro každý kanál), které jsou pootočené o  $45^\circ$  vůči rovině desky. Na fotografii 2.5 je vidět detail řezací hlavy a poloha dvojice navzájem kolmých cívek.

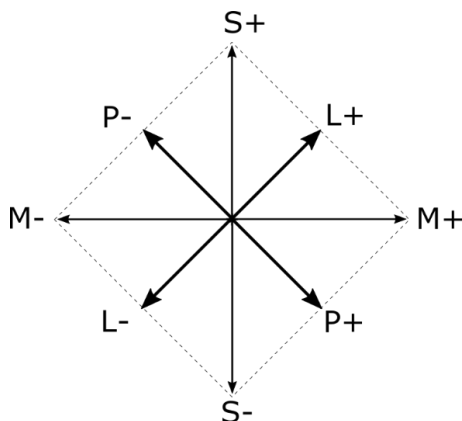


Obr. 2.5: Detail řezací hlavy (převzato z internetu<sup>3</sup>)

Aby pak záznam byl co nejsnadněji čitelný (což je záznam stranový, nikoliv hloubkový), mají cívky vůči sobě otočenou polaritu. Důsledkem této otočené polarity poté mono signál (tedy  $L=P$ ) vyústí pouze v čistě stranovou výchylku (vertikální složky výchylek se navzájem odečtou). Signály v přesné protifázi pak způsobí naopak čistě vertikální výchylku. Tento systém je v praxi vhodný zejména proto, že hudební signál má většinou mnohem větší podíl „soufázových“ úseků, na úkor částí „protifázových“.

<sup>3</sup> [www.gzvinyl.com](http://www.gzvinyl.com) (cit. 27.12.2016)

Ze schématu 2.6 je dobře patrné, jak jsou vůči sobě jednotlivé kanály a horizontální a vertikální výchylky řezacího hrotu postaveny.



Obr. 2.6: Schéma směru výchylky v závislosti na okamžitém signálu, způsob skládání monofonní a stereofonní složky signálu

Ze schématu pak lze snadno odvodit jednoduché vztahy mezi vektory (2.3, 2.4) a rozměry (2.5, 2.6) jednotlivých kanálů a zavést tak tzv. mono- a stereosložku (označeny M a S):

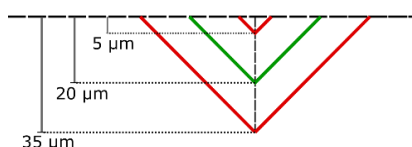
$$\mathbf{M} = \mathbf{L} + \mathbf{P} \quad (2.3)$$

$$\mathbf{S} = \mathbf{L} - \mathbf{P} \quad (2.4)$$

$$M = (L + P)/\sqrt{2} \quad (2.5)$$

$$S = (L - P)/\sqrt{2} \quad (2.6)$$

Vzhledem k vlastnostem mechanického záznamu a technologie DMM (tedy omezená vrstva mědi, diamantový nůž určitého průřezu), je zřejmé, že vertikální výchylka bude svým způsobem kritičtější než výchylka horizontální. Důvodem je to, že zatímco při horizontální výchylce dochází k posunu celé drážky směrem do strany, při výchylce vertikální se mění samotná hloubka drážky a důsledkem změny hloubky se stejným způsobem mění i šířka drážky. Na obrázku 2.7 je načrtnuta základní pozice drážky (nastavené na šířku 40  $\mu\text{m}$ ) a její průřez při extrémních vertikálních výchylkách (zde  $\pm 15 \mu\text{m}$ ). Větší výchylka by riskovala jednak příliš mělkou drážku, ve které by dobře nedržela snímací jehla, a v opačné půlvině pak drážku příliš hlubokou a z toho plynoucí riziko proříznutí vrstvy mědi a zničení řezacího hrotu. (Zde je nutno podotknout, že při dodržení správné tloušťky měděné vrstvy je toto riziko spíše minimální). Tyto limity nejsou ani tak čistě technické, jako spíše obecně používané dle zkušenosti z praxe.



Obr. 2.7: Průřez drážky bez vertikální výchylky a v obou „extrémních“

## 2.4 Možné vady drážky a jejich důsledky

Jedním z účelů projektu, kromě optického přehrávání, je také detekce kritických míst a poškození nosičů. Jedná se zejména o následující vady a problémy:

**Geometrické zkreslení** – Těž úhlové. Je způsobeno fyzickým tvarem a rozměry řezacího nože. Vzniká v místech, kde by měl nůž provést pohyb, se směrnicí odchýlenou od směru rotace o více než cca 40°. Toto zkreslení se více projevuje u středu desky, kde je drážková rychlost nejmenší. Projevuje se harmonickým zkreslením, proto je kritické jeho vzniku zabránit zejména v oblasti do 10 kHz, kde případné vyšší harmonické spadají do slyšitelné oblasti spektra.

**Prořezy** – Jedná se o vadu, kde v důsledku nevhodně nastavených přepisových parametrů dojde ke spojení dvou sousedních drážek. V tomto místě pak (zejména u špatně nastaveného gramofonu co do přítlaku a kompenzace odstředivé síly - antiskating) může dojít k přeskoku snímací jehly do sousední drážky.

**Mělká drážka** – Důsledkem většího podílu protifázového stereo signálu může drážka mít velký podíl vertikální výchylky, což znamená, že na desce budou místa, na kterých bude drážka velmi úzká a mělká. V kombinaci se stranovou výchylkou a aktuálních zrychlení čtecího hrotu může obdobně jako u prořezu dojít k přeskoku jehly do sousední drážky.

**Vady galvanoplastiky** – Jedná se o mechanické problémy drážky způsobené v procesu galvanoplastiky, kdy díky nedokonalostem v procesu nedošlo ke stoprocentnímu přenosu materiálu mezi jednotlivými fázemi výroby. Jedná se tak například o vytrhaná dna drážky, ztrhané hrany drážky apod. Důsledkem může být jednak klasický praskot při přehrávání desky, ale i zkreslení vlivem geometrických chyb takto vadně přenesené drážky, projevující se například nadbytečnou stereosložkou (jehla kmitá vertikálně i tam, kde nemá a tím způsobuje protifáze a jakési „rozpliznutí“ stereo obrazu).

## 2.5 Metody čtení mechanického záznamu

Jelikož měděné originály, jimiž se tato práce zabývá, jsou pozitivní, tedy nesou mechanicky stejnou informaci jako skutečná vyrobená deska, nebude tato kapitola příliš rozlišovat mezi metodami čtení používanými právě u těchto originálů a metodami čtení samotných desek. Zejména v případě mechanických a elektromechanických metod čtení v tom totiž není žádný rozdíl.

Metody čistě mechanické jsou již tak zastaralé, že nemá přílišný smysl věnovat jim vlastní kapitolu, stačí podotknout, že v případě mechanického záznamu jsou vibrace čtecí jehly zesíleny na mechanické bázi a následně převedeny na akustické vlny, které jsou opět zesíleny pomocí rozšiřujícího se

zvukovodu. Tato metoda čtení je použita na slavném logu vydavatelství „His Master’s Voice.“

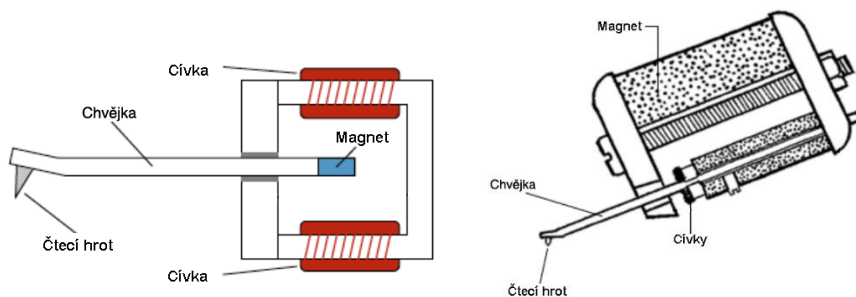


Obr. 2.8: Jedna z podob loga vydavatelství His Master’s Voice (převzato z internetu<sup>4</sup>)

Mnohem důležitější, kvalitnější a dodnes používané jsou metody, které zde nazývám jako elektromechanické. Jsou založeny na podobných principech jako většina elektroakustických měničů, tj. reciproký převod akustického (zde pouze mechanického) kmitání na elektrický signál pomocí principů indukce napětí na cívce v magnetickém poli. Rozlišují se dva systémy: MM (moving magnet) a MC (moving coil). Konstrukčně jsou si velmi podobné a liší se zejména v tom, který prvek měniče je připevněn k pohyblivé jehle (chvějce) a který je statický [5].

V případě MM systému je pohyblivá pouze čtecí jehla, na jejímž konci v pouzdře se nachází magnet. Statickým prvkem jsou pak dvě cívky indukující napětí. Pro každý kanál je tato cívka jedna. Obdobou je systém MIC (Moving iron cartridge) lišící se v tom, že místo magnetu je na konci jehly feromagnetický materiál. To umožňuje ještě větší snížení hmotnosti jehly.

Druhým systémem je pak MC (moving coil – systém s pohyblivou cívkou). Zde je magnet zabudovaný do pouzdra přenosky a k samotné jehle jsou připevněny (opět v páru) cívky. Jelikož jehla musí být schopna velice jemného a rychlého pohybu ke správnému sledování drážky v desce, jsou i tyto cívky nesmírně malých rozměrů a vyrobeny jsou z velice jemného vlákna. Důsledkem tohoto omezení je výstupní napětí v řádu zlomku milivoltů [5]. MC přenosky proto vyžadují speciální předzesilovače. Stavba obou systémů je na obrázku 2.9.



Obr. 2.9: Stavba přenosky MM (vlevo) a MC (vpravo) systému (převzato z [5])

<sup>4</sup> [www.hikehalfdome.com/hear-masters-voice](http://www.hikehalfdome.com/hear-masters-voice) (cit. 9.11.2016)

## Kapitola 3

# Bezkontaktní snímání mechanického záznamu

Vzhledem ke skutečnosti, že kontaktní snímání desky pomocí klasické čtecí jehly způsobuje její fyzické opotřebení a s ním postupnou ztrátu vyšších kmitočtů a vznik zkreslení vlivem „odřením“ nejjemnějších detailů drážky, v minulosti vznikla či byla rozpracována ve formě studie řada možných řešení bezkontaktního snímání. V těchto řešeních šlo jak o možnost bezkontaktního přehrání desky, tak zejména o možnost archivace starých či raritních nahrávek, které už jsou klasickými metodami nečitelné, ať už z důvodu rozsáhlých poškození drážky, nebo i celé desky (například desky neúplné či zlomené – tedy desky, u nichž nelze žádnými způsoby snímat drážku kontinuálně právě proto, že tato kontinuita již byla porušena) [6]. Příklad takovéto desky se nachází na obrázku 3.1 a jedná se o desku, jejíž obsah je stále v určité míře obnovitelný pomocí jedné z uvedených metod.

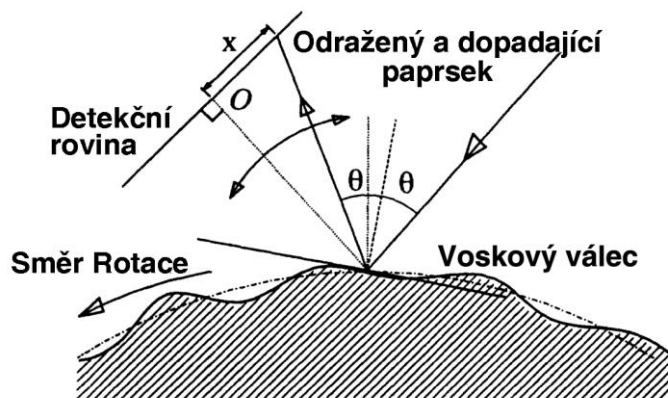


Obr. 3.1: Příklad poškozené archivní desky (převzato z [10])

## 3.1 Metody využívající laserového paprsku

Průkopníky v oblasti optického snímání byl Japonský kolektiv okolo T. Iwaie, který ve svém článku z roku 1986 popisuje využití laserového paprsku pro čtení záznamu z Edisonovského voskového válečku [6]. Záznam je na voskovém válečku uložen v podobě záznamu hloubkového (pouze monofonní). Laser je v tomto případě vyslán směrem k válci a od něj následně odražen. Snímání odraženého paprsku probíhá na rovině k němu kolmé. S otáčením válce se mění místo, kam paprsek dopadá a tato změna po ekvalizaci odpovídá generovanému zvukovému signálu. Z obrázku 3.2 je patrné, že čím vyšší frekvence bude

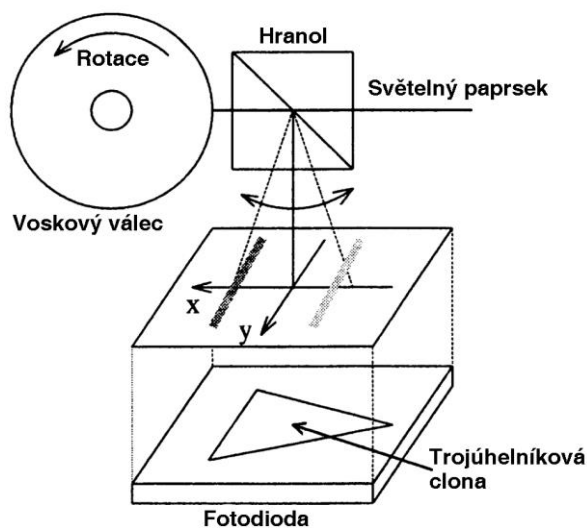
zaznamenána, tím více bude laser vychýlen ze své základní polohy, neboť okamžitá výchylka  $x$  závisí jak na hloubce čtené drážky, tak na úhlu drážky, který odpovídá směrnici analogového signálu.



Obr. 3.2: Laserové čtení Edisonovského válečku (převzato z [11])

Později v dalším článku [11] jeden z autorů navazuje návrhem několika různých vylepšení: Mezi nejdůležitější patří například možnost kompenzace chyby sledování drážky. Používají zde 2D detekční rovinu, přičemž pouze detekce ve správné ose má vliv na výsledný signál. Další vylepšení spočívá v zabudování ochrany proti vlivu parazitního světla v podobě vložení další čočky a clony s malým otvorem mezi záznamový váleček a snímací senzor.

Optický signál se na elektronický převádí pomocí fotodiody s trojúhelníkovou maskou. Na obrázku 3.3 dále vidíme, jak tento převod funguje a jak je zajištěno, že budou na výsledný signál mít vliv pouze výchylky ve směru osy  $x$ . Použití tohoto prvku má ovšem za důsledek nutnost použití jiného prvku pro detekci chyby sledování drážky, neboť tato konfigurace fotodiody není citlivá v ose  $y$ .



Obr. 3.3: Diagram použití fotodiody jako snímacího prvku (převzato z [11])



Další dvě úpravy spočívaly ve „zkontaktnění“ této bezkontaktní metody použitím stylusu a optického vlákna. První takto obohacená metoda použila optické vlákno jako vodič světelného paprsku a čtecí hrot byl připevněn k němu tak, že pohybem válečku se pohyboval konec vlákna tak, že světelný paprsek byl směřován na detekční rovinu bez nutnosti odrazu od válečku. Jednalo se tak vlastně o převod mechanického pohybu na pohyb optického paprsku prostorem. Přestože metoda nebyla plně bezkontaktní, působil stylus na povrch drážky mnohem menší silou než u tradičního mechanického čtení a nedocházelo k tak rychlé degradaci materiálu.

Druhá úprava kombinovala obě předešlé varianty tak, že jako stylus bylo využito přímo optické vlákno. Mechanicky tedy došlo k vyloučení chyby trackování drážky a pro detekci byl opět využit odraz od povrchu válečku a detekční senzor.

Druhá metoda laserového čtení desky je určena přímo pro přehrávání klasických gramofonních desek a jedná se přímo o laserový gramofon, využívající principů objevených v roce 1986 Stanfordským studentem Robertem E. Stoddardem, a sestavený v roce 2003 firmou ELP<sup>5</sup> [7]. Tento laserový systém obsahuje 5 laserových paprsků, z nichž dva jsou použity přímo pro snímání stěn drážky a potažmo tak audiosignálu, zatímco tři zbývající slouží k určení správné polohy celé drážky. Kromě bezkontaktnosti řešení nemá tato metoda proti klasickému kontaktnímu přehrávání téměř žádné výhody. Veškeré nečistoty (prach, apod.) způsobí nekvalitní reprodukci – je tedy před každým použitím potřeba desku zbavit prachu. Systém není schopen přehrát desky, které jsou příliš ohnuté a samozřejmě pak ani desky fyzicky poškozené, což dělá systém prakticky nepoužitelný pro digitalizaci a archivaci starých záznamů. Systém je velmi drahý (řádově stovky tisíc Kč<sup>6</sup>).

## 3.2 Metody využívající snímání obrazu

Jelikož laserová metoda se ukázala být příliš drahá a pro mnoho účelů nepoužitelná, bylo třeba přijít s metodami, které využívají snímání obrazu pomocí kamery/fotoaparátu a následné zpracování nasnímaného materiálu.

V roce 2002 přišel Ofer Springer s myšlenkou tzv. Digitální jehly ve virtuálním gramofonu. Práce byla založena na myšlence naskenování obrázku desky a z něj následně pomocí principů mechanického zápisu dekodovat obsah

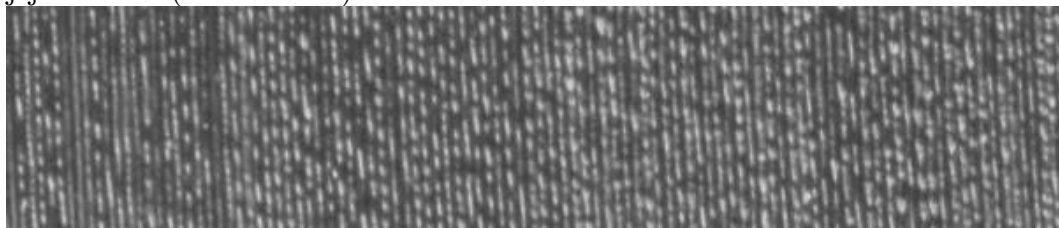
---

<sup>5</sup> Oficiální web společnosti ELP (<http://www.elpj.com/>) na stránkách cenu neuvádí.

Z pochopitelných důvodů veškerá komunikace s (potenciálními) zákazníky, včetně tedy cenové kalkulace, probíhá velmi individuálně prostřednictvím e-mailu. (cit. 5.1.2016)

<sup>6</sup> Dle článku ze 7.5.2015 se jedná o rozpětí 14-18,5 tisíce amerických dolarů. Dostupné z: <http://www.cultofmac.com/321881/14000-turntable-wont-wear-out-your-vinyl-what-a-deal> (cit. 27.12.2016)

nahrávek [6] [9]. Z informací na webové stránce [9] projektu vyplývá, že tyto principy spíše odhadl, navíc chybně. Z kapitoly 2.3.2 již víme, že jednotlivé kanály jsou modulované do drážky pod úhlem  $45^\circ$  a tak horizontální a vertikální výchylka tvoří monofonní, respektive stereofonní složku signálu. Springer se domníval, že horizontální a vertikální výchylka už jsou přímo jednotlivé kanály. Jeho virtuální jehla pak spočívala v tom, že byl určen střed desky a jehla se podobně jako skutečná jehla pohybovala po kružnici, na které četla detaily drážky. Výsledky jsou vzhledem k velmi malému rozlišení snímků značně neuspokojivé a nedostatečné (k poslechu na [9]). Celý projekt je nicméně obdivuhodný už z toho důvodu, že Springerova práce ani neměla ambice stát se odborným článkem, ale spíše se jednalo o odpolední projekt studenta bez detailní znalosti problematiky. Zajímavé na tomto projektu je to, že pro získání co největších detailů se Springer nepustil do analýzy stranové výchylky (pro jejíž čtení měl příliš malé rozlišení obrazu) nýbrž do čtení výchylky hloubkové, která je díky kolmému osvětlení i na nekvalitním snímku mnohem lépe patrná (obrázek 3.4). Nerekonstruoval tak tedy jeden z kanálů stereozáznamu, nýbrž jejich rozdíl (stereoslůžku).



Obr. 3.4: Část desky v podobě, v jaké s ní pracoval Springer (převzato z [9])

Bohužel již nevíme, jestli obrázek na webové stránce je v plném rozlišení a barevnosti, a tak nejsme schopni přesně určit snímací parametry systému.

Práce nicméně o rok později inspirovala kolektiv švédských studentů okolo P. Olssona, kteří snímání vylepšili o kvalitnější zpracování signálu [7] použitím Wienerovy filtrace, spektrálních operací apod. Obě metody nicméně v důsledku použití běžného skeneru (max 2400 pixelů na palec, odpovídající přibližně  $10 \mu\text{m}/\text{pixel}$ ) mají omezené rozlišovací schopnosti ústící v příliš vysokou hodnotu kvantizačního šumu, jelikož maximum horizontální výchylky bývá typicky do  $150 \mu\text{m}$  (Principu omezeného kvantování vycházejícího z pixelového rozlišení snímaného obrazu se budu věnovat v kapitole 4.3.1). Na další problémy použití tohoto typu skeneru upozorňuje Stotzer v práci [10], které se budu věnovat později v kapitole 3.3.1. Jedná se o nutnost desku skenovat na více částí a následně skládat dohromady a také problémy spojené s transformací obdélníkových fotografií s čtvercovou mřížkou pixelů do polárních souřadnic. Dále pak poukazuje na zmíněné nedostatečné rozlišení a stanovuje nutnost rozlišení alespoň  $1 \mu\text{m}/\text{pixel}$ .

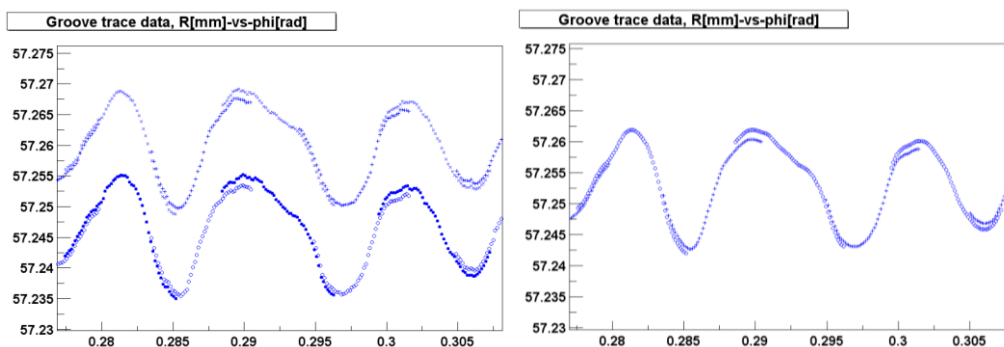
Většina dalších pokusů pak už jen navazovala na tento princip s tím, že se zabývala jednotlivými částmi zpracování a ty pak vylepšovala.

### 3.3 Pokročilé metody využívající mikrofotografie

Pánové Fadeyev a Haber v roce 2003 vyvinuli metodu, které pro převod na zvuk využívala snímání dna drážky [6], umožňující tak poměrně nenáročnou rekonstrukci monofonního signálu. Jejich vylepšení spočívalo v použití mikrofotografie, mající za důsledek (kromě nutnosti mnohem složitějšího zpracování) zejména vyšší rozlišení obrazu a tím i větší kvalitu rekonstruovaného zvukového signálu. Zabývají se rekonstrukcí zvuku na archivních deskách s rychlostí 78 otáček za minutu. Jednak proto, že nejstarší desky, u kterých je větší pravděpodobnost, že jsou raritní a poškozené, používaly tento standard a jednak proto, že tyto desky bývají často monofonní a mají širší drážku (až 160  $\mu\text{m}$  oproti typickým dnešním 45 u moderních dlouhohrajících desek). Použili již profesionální skenovací zařízení s jemným X-Y posunem a mikroskopem, pomocí kterého dokázali na jeden snímek o rozlišení  $768 \times 494$  pixelů umístit výřez obrazu o rozměrech až  $260 \times 200$  mikronů (při maximálním zvětšení).

Dle tabulky vypočtené v jejich práci [12] je rozumný úhel záběru  $700 \times 200 \mu\text{m}$  a takovýchto polí se na užitém plochu 10“ desky vejde  $10^5$ - $10^6$  v závislosti na efektivitě snímání. Při předpokládané datové náročnosti přibližně 1 MB na černobílý snímek lze stanovit, že pro nasnímání jedné desky bude potřeba datový prostor v řádu stovek GB. V případě reálnového snímání by se systém dostal na datový tok v rozmezí 0.5-5 Gb/s, přičemž vhodnou optimalizací procesu může být redukován. Jako vhodnou optimalizaci autor navrhuje použití rychlého signálového procesoru, který data předzpracuje a zmenší jejich velikost před samotným přenosem do počítače, kde dojde k dalšímu zpracování. Monofonní signál rekonstruovali pomocí detekce hran dna drážky, které byly zvýrazněny vhodně zvoleným osvětlením. Podél drážky byly dvojice bodů samplovány každých 8  $\mu\text{m}$ , což na použitém formátu desky odpovídá vzorkovací frekvenci 61,3 kHz na poloměru 60 mm. Pokud byly konkrétní body nečitelné (například z důvodu prachu) byly přeskočeny a interpolovány, čímž se snadno eliminoval šum, praskot a další elementy, které by při přehrávání mechanickou přenoskou mohly působit rušivě. Tyto dvojice byly následně zprůměrovány.

Data byla převedena do polárních souřadnic, kde každý bod obsahoval informaci o úhlu natočení a vzdálenosti od středu desky. Na obrázku 3.5 je patrná podoba dat v polárních souřadnicích před zprůměrováním na jednotlivé body a po něm – za povšimnutí stojí místa, kde se překrývají data pocházející z překrývajících se snímků. Následně byla vyfiltrována problematická data a potlačeny vlivy nepřesně určeného středu a přirozené spirálovitosti drážky. Data byla numericky vyhlazená polynomem 4. řádu a resamplována do standardního CD formátu.



Obr. 3.5: Překrývající se data před a po zprůměrování (převzato z [12])

Jelikož systém není navržen specificky za účelem skenování desek, trvalo tak naskenování 1 vteřiny záznamu přibližně 50 minut, nicméně autoři si jsou jistí, že dedikované zařízení čas „dramaticky“ zkrátí. V závěru autoři srovnávají použití optického čtení, čtení přenoskou a nový CD master těžce skladby. Výsledkem je dle očekávání minimalizace praskotu v optické cestě, jelikož zrnka prachu, které praskot při mechanickém čtení způsobují, byly při optickém zpracování úspěšně eliminovány. Autoři bohužel neuvádí žádnou informaci o výsledné hodnotě šumu v signálu, omezují se pouze na tvrzení, že „v rekonstruovaném signálu byl přítomen souvislý šum pozadí, lehce modulovaný signálem o frekvenci přibližně 4 Hz“ [12].

Dedikované zařízení by se pak lišilo zejména v možnosti pohybu přímo po kružnici, nikoliv pak v pravoúhlém rastru.

Baozhang Tian s Johnem L. Barronem pak v roce 2006 navrhli systém snímající 3D obraz celé drážky [7], což je zbytečně složité, neboť pro rekonstrukci signálu není potřeba zdaleka tolik informací. Kolektiv nicméně věřil tomu, že se ubírají správným směrem, protože ostatní 2D metody (včetně těch, kterými se zabývám v této práci já) jsou přímo závislé na hranách, ať už se jedná o dno drážky či rozhraní plocha desky – drážka. Myšlenkou práce pak byla skutečnost, že velké množství informace lze získat z orientace povrchu stěny drážky a tím pádem pro rekonstrukci signálu bude dostupné maximální množství informace a bude tak odolnější vůči abnormalitám, jako je například prach usazený v kritických místech. Systém (stejně jako všechny ostatní s výjimkou komerčního laserového gramofonu) nepracuje v reálném čase kvůli limitům na straně kamer a výpočetního výkonu. Jak v práci sami uvádí, dvouvteřinový klip pro svoje zpracování vyžaduje vygenerování 1390 skupin po 36 snímcích [6]. Celý projekt se tak nachází ve formě návrhu k další práci a přes použití velmi komplikovaných algoritmů a systémů nepřináší příliš praktických výsledků.

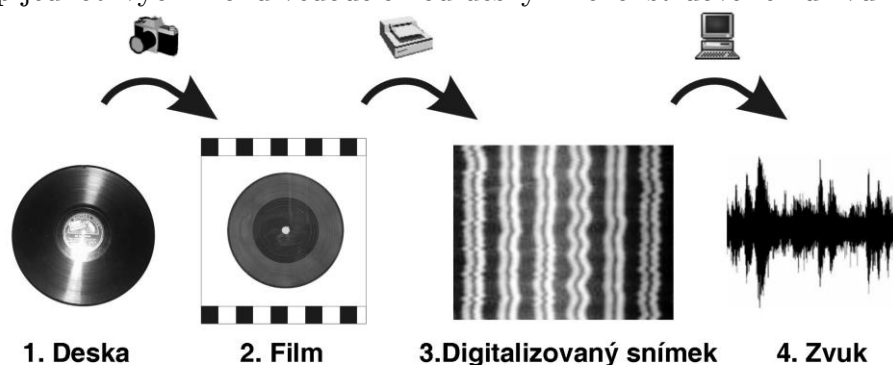
Kombinaci 2D a 3D technologií zvolil v roce 2007 francouzský kolektiv v čele s Luisem Laborellim [8]. Jejich práce spočívala na principu osvětlení desky svazkem barevných paprsků a následný tvar drážky vyhodnotili dle barvy světla zaznamenané kamerou. Dá se tedy říci, že částečně navázali na práci [9], který vyhodnocoval výchylku na základě změny jasu drážky osvětlené stolním skenerem. Čtecí systém je možné naklonit o 45° a tím použít ke čtení

stereozáznamu. Systém dosahuje poměrně slušných hodnot při výsledném měření kvality. Harmonické zkreslení (THD)<sup>7</sup> se pohybuje mezi 10-13 procenty, po nelineární korekci dokonce mezi 1-2 procenty. Odstup od šumu na čistých deskách bez viditelného prachu dosahuje hodnot až 45 dB.

### 3.3.1 VisualAudio concept

V roce 2006 také Sylvain Stotzer publikoval rozsáhlou disertační práci [10], ve které se zabývá principem extrakce zvukové nahrávky z desky optickou metodou tak, že je nejprve vizuální obsah desky ve skutečné velikosti 1:1 přenesen na světlocitlivý film a z něj následně snímán řádkovou kamerou. Mezi hlavní výhody systému autor považuje právě mezikrok spočívající v analogové fotografii desky. Zaprvé umožňuje tak zachovat velké množství desk velmi rychle a následně digitální a počítačové zpracování tak provádět až ve chvíli, kdy je skutečně potřebné či vyžadované a za druhé má fotografie mnohem větší hloubku ostrosti než mikroskop, pomocí kterého jsou data následně čtena. Umožňuje proto mít v rovině ostrosti jak okraje drážky, tak její dno.

V této práci jsou poměrně detailně popsány všechny kroky vedoucí od samotné desky k restaurované zvukové nahrávce. Na schématu 3.6 je naznačen princip jednotlivých kroků vedoucích od desky k rekonstruovanému zvuku.



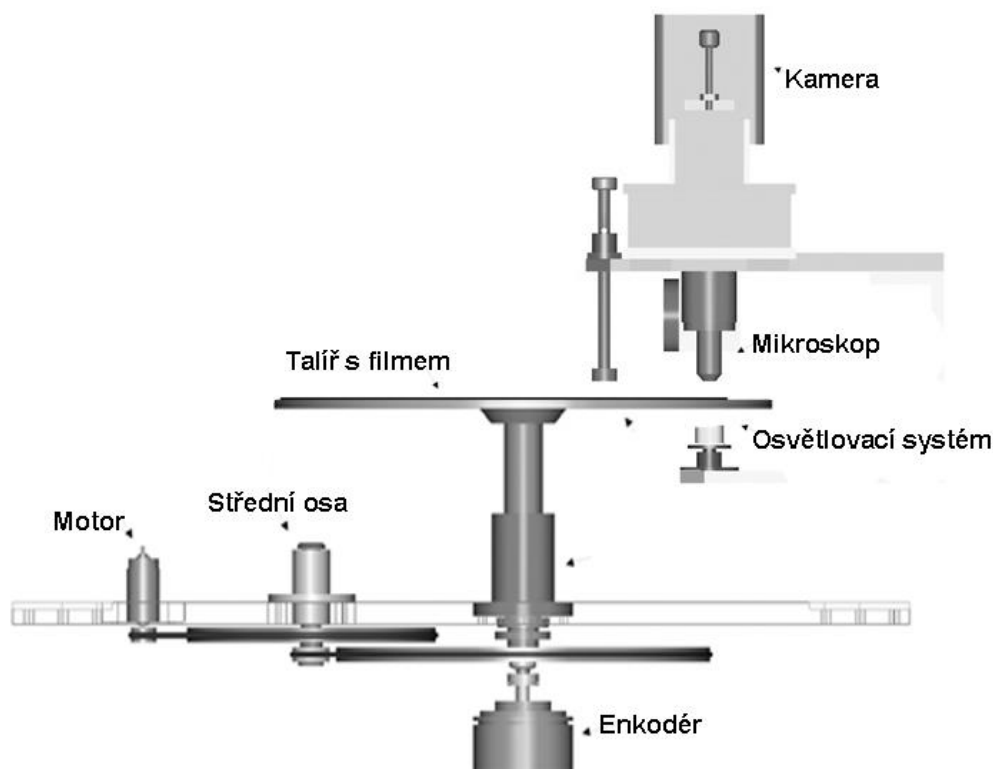
Obr. 3.6: Jednotlivé kroky VisualAudio Conceptu (převzato z [10])

Zabývá se výběrem vhodného filmu, který musí být co nejostřejší a mít pokud možná co nejmenší velikost zrn, aby dokázal kvalitně reprezentovat detaily gramofonové drážky. Dalšími požadavky byly například vysoký kontrast, časová stabilita materiálu a obrazu, rychlost vyvolávání a samozřejmě rychlá dostupnost v dostatečně velkém formátu a cena. Kvůli dříve jmenovanému problému s hloubkou ostrosti byl zvolen film černobílý, neboť obsahuje pouze jednu vrstvu světlocitlivého materiálu a tedy při skenování stačí menší hloubka ostrosti, než jakou by vyžadoval film barevný. Jelikož se jedná o velmi specifickou aplikaci,

<sup>7</sup> Harmonické zkreslení desky při čtení klasickou přenoskou záleží na celém řetězci, nicméně v závislosti na kvalitě a typu přenosky a úrovni signálu se může pohybovat mezi 1 až 10 procenty [1].

nenášel autor práce žádný komerčně dostupný fotoaparát, který by dokázal splnit veškeré technické požadavky, byl proto pro tento účel sestrojen fotoaparát nový s rozměry cca 2 metry na výšku a metr na šířku. Ohnisková vzdálenost použité čočky pak byla rovna 42 cm. Ve výšce čočky byl umístěn osvětlovací systém tvořený spirálou modrého zdroje světla s vlnovou délkou 463 nm umístěnou za difuzním sklem, aby chování světla bylo co možná nejvíce homogenní. Vnější průměr spirály byl 83 cm. Expoziční doba takového fotoaparátu se pro správné výsledky pohybuje okolo 30 sekund. Při navýšení času na jednu minutu se pak díky difuznímu halo efektu zužuje obraz drážky z původních reálných 120 na cca 80  $\mu\text{m}$ .

Stěžejní mechanickou částí celého projektu pak je skenovací systém, který je konstrukčně podobný jako samotná řezací aparatura, která je používána v GZ Media. Původní návrh z roku 2002 obsahoval přímý náhon osy rotačního talíře, nicméně se ukázalo, že přenáší příliš mnoho vibrací, které jsou následně slyšitelné v rekonstruovaném zvuku v podobě šumu. Byl proto navržen systém, kde byla rotace asynchronního motoru přenášena dvoustupňovým páskovým převodem v poměru 1:100. Náhon poháněl skleněný talíř, na kterém byl připevněný film s obrazem desky. Celý pohonný i čtecí mechanismus je na obrázku 3.7.



Obr. 3.7: Mechanismus pohonu a čtení VisualAudio Projectu (převzato z [10])

Z obrázku však není patrné, že snímací a osvětlovací část se pohybují radiálně společně za pomoci závitové tyče, která umožňuje dostatečnou přesnost v rozsahu průměrů 9-40 cm, což stačí na snímání jakéhokoliv formátu desky.

Srdce snímače tvoří řádková CCD kamera s rozlišením 2048 pixelů a optika, která na tento rozsah převede přibližně 5 mm reálného obrazu. To tak dává k dispozici rozlišení 2,44  $\mu\text{m}/\text{pix}$ . Kamera byla nastavena na snímkovací frekvenci 13 kHz a synchronizace probíhala podle reálných rotací dekodéru umístěného souose s rotačním talířem tak, že při kolísání rychlosti rotace kamera snímala stále ve stejných úhlových, nikoliv časových intervalech.

Vzhledem ke kontinuální rotaci talíře bylo třeba potýkat se s pohybem, který snímek vykonal v době aktivní integrace světla na čipu kamery, jelikož tato doba v reálu nemohla být nekonečně malá. Doba integrace se musí pohybovat mezi 20  $\mu\text{s}$  (což je minimum, s jakým konkrétní použitá kamera dokáže pracovat) a 76  $\mu\text{s}$ , což je převrácená hodnota snímkovací frekvence. Jelikož se s průměrem desky snižuje obvodová rychlost, snižuje se zároveň délka drážky, o kterou se talíř otočí během integrační doby a tím se snižuje i množství pohybového rozmazání.

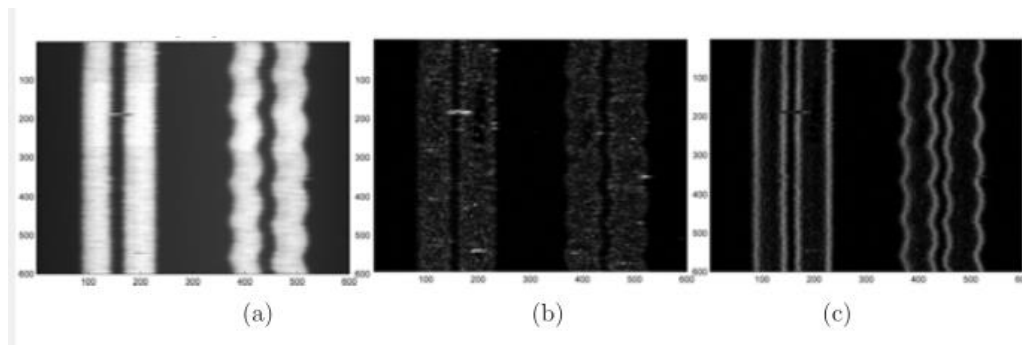
V celém systému je celá řada možných zdrojů rozmazání, takže autoři stanovili celkové minimální rozlišení systému jako odmocninu součtu kvadrátů jednotlivých zdrojů dle rovnice (3.1), kde  $B_1$ - $B_5$  jsou postupně rozostření fotoaparátu, filmu, skenování, pohybu a samplování.

$$B = \sqrt{B_1^2 + B_2^2 + B_3^2 + B_4^2 + B_5^2} = \quad (3.1)$$

$$= \sqrt{23,7^2 + 1,7^2 + 4,8^2 + 3,3^2 + 2,5^2} = 24,6 \mu\text{m}$$

V závislosti na rovinnosti filmu, modulaci drážky a skutečnosti, že rozostření vlivem pohybu (motion blur) a samplování se projevují ve směrech navzájem kolmých je pravděpodobné, že reálná hodnota se bude pohybovat ještě níže. Největší příspěvek k celkové rozostřenosti je ovšem samotný fotoaparát a jeho hloubka ostrosti. Pokud by byl disk plošší, než se předpokládá, je možné celkové rozostření stanovit na přibližně 17,4  $\mu\text{m}$ .

Jelikož zdrojové fotografie byly černobílé, a stěny drážky se vyznačovaly velmi rozdílným jasnem než zbytek desky a dno drážky, byl pro hledání hran využit tečný (vertikální) a kolmý (horizontální) gradient snímku. Jak lze předpokládat, tečný gradient dokáže odhalit škrábance a vady povrchu, ale je mnohem citlivější na šum, neboť ve vertikálním směru nedochází k tak výrazným změnám jasu (obrázek 3.8). Pro hledání hran je tedy výrazně vhodnější gradient kolmý, nicméně – jak je patrné z obrázku – díky rozostření fotografie je šířka maxima tohoto gradientu příliš široká.



Obr. 3.8: Srovnání zdrojového snímku a jejího gradientu v obou směrech (převzato z [10])

Jelikož skenování v této práci probíhalo pomocí řádkové, nikoliv rastrové kamery, jednotlivými „snímky“ tedy byly pouze 1D vektory neboli jasové profily daného řádku. Pomocí vyhlazených aproximací derivací tohoto profilu, hledání lokálních minim a maxim, lineárního vyhlazování apod. byly najity tzv. kandidující body, což jsou body, které by se mohly stát hranou drážky na daném řádku. Pomocí vyhodnocení v kontextu s dalšími body je teprve určeno, jestli daný bod pravděpodobně leží na hraně, nebo se jedná o bod, jehož poloha leží mimo všechny ostatní body, je pravděpodobně způsoben chybou algoritmu či samotné desky a je tak z vyhodnocení vyřazen.

Zpracování signálu dále obsahuje celou řadu algoritmů pro odstraňování chyb, ať už náhodných, či majících původ ve vadě materiálu, či kteréhokoliv kroku z celkového procesu.

Na závěr jsou získaná audiodata převedena do některého ze standardních audioformátů (typicky 44,1 kHz, 16 či 24 bitů) a změřena kvalita audia. Jelikož úroveň šumu záleží zejména na rozlišení zdrojových snímků, ale měřena byla až na výsledném souboru, byla zvolena metodika postupného snižování bitové hloubky a měření SNR. V hodnotách mezi 24 a 10 bity se SNR neměnilo, k postupnému snižování došlo až mezi 9-8 bity, lze tedy říci, že kvalita měřena tímto kvantifikátorem odpovídá tedy cca 512 rozlišitelným úrovním. Změřené celkové SNR systému bylo 16,71 dB.



## Kapitola 4

# Převod snímku drážky na zvukovou informaci

Základní myšlenkou převodu obrazu na zvuk je skutečnost, že průřez drážky je tvořen pravoúhlým trojúhelníkem, jehož pravý úhel se nachází na dně drážky. K získání informace o tvaru drážky nám tedy postačí znát jen některé její geometrické vlastnosti a tyto jsme následně schopni převést na informaci o audiosignálu, který byl zapotřebí pro její vytvoření a který z ní tedy lze zpětně rekonstruovat.

## 4.1 Rekonstrukce mono- a stereosložky signálu

Jedním z možných přístupů je získání informace o poloze středu drážky (dna) a informace o celkové šířce drážky. Z těchto informací pak můžeme okamžitou hodnotu monofonní a stereofonní složky signálu spočítat pomocí rovnic (4.1) a (4.2), kde  $m(t)$  je okamžitá poloha středu drážky,  $m_0$  rovnovážná poloha,  $w(t)$  okamžitá šířka drážky a  $w_0$  základní šířka drážky (tedy například zmíněných 40  $\mu\text{m}$ ).

$$M(t) = (m(t) - m_0) \cdot K \quad (4.1)$$

$$S(t) = (w(t) - w_0) \cdot K \quad (4.2)$$

$K$  je převodní konstanta pro přepočítání mezi metrickými rozměry a rozměry snímaného obrazu v pixelech, případně pro normování signálu, jeho zesílení apod. Nastavuje se individuálně, nicméně je důležité podotknout, že pro oba dva směry výchylky musí být shodná, aby byla zaručena kolmost signálů jednotlivých kanálů.

Pokud známe mono a stereo signály, tak je již velmi snadné upravit rovnice (2.3 až 2.6) do podoby (4.3 a 4.4) a z nich zpětně určit informaci náležící levému i pravému kanálu.

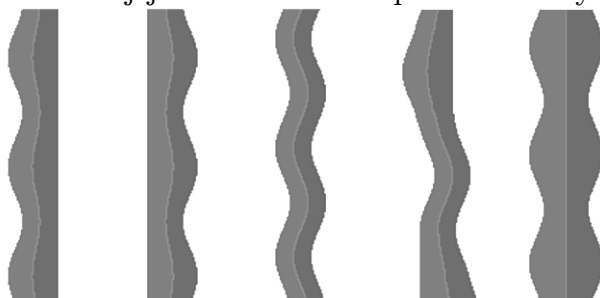
$$L = \frac{M + S}{\sqrt{2}} \quad (4.3)$$

$$P = \frac{M - S}{\sqrt{2}} \quad (4.4)$$

Tento přístup je vhodný například v situacích, kdy je snazší nalézt polohu dna drážky či nám stačí informace pouze o monofonním signálu.

## 4.2 Přímá rekonstrukce levého a pravého kanálu

Druhým možným (a při rekonstrukci stereofonního signálu i vhodnějším) přístupem je pak rekonstrukce přímo jednotlivých kanálů signálu. Jak je patrné ze schématu 2.6 a z konstrukce řezacího nože, směr řezu v jednotlivých kanálech je sice pootočen o 45° vůči rovině desky, nicméně audiosignál, který se nachází pouze v jednom kanále, vyústí v drážku, jejíž jedna hrana bude naprosto rovná a druhá bude reprezentovat signál z daného kanálu. Na obrázku 4.1 je naznačen vliv jednotlivých kanálů a jejich kombinací na podobu drážky.



Obr. 4.1: Drážka v případech, že je signál (zleva): Pouze v pravém kanále, pouze v levém kanále, v obou stejný (mono), v každém kanále jiný (stereo), v kanálech vzájemně v protifázi (převzato z internetu<sup>8</sup>)

V případech, kdy jsme schopni snadněji detekovat polohu hran drážky, nikoliv jejího dna (jako například v praktické demonstraci v rámci této práce) je proto tento postup výhodnější proto, že z tvaru jednotlivých hran lze přímo stanovit signál náležící jednotlivým kanálům. Pozor: Záměrně nepíšu, že tvar drážky přímo reprezentuje tvar signálu, neboť toto tvrzení není pravdivé. Jak je naznačeno v kapitole 2.3.1 a dále pak znovu v praktické části práce v kapitole 9.2 o implementaci převodu informace o poloze drážky na zvuk, okamžitá výchylka elektromagnetického ústrojí (tedy řezací jehly) neodpovídá okamžité hodnotě signálu. Tomu odpovídá okamžitá rychlost jehly. Pro zpětnou rekonstrukci signálu je tedy potřeba informaci o drážce zderivovat.

## 4.3 Základní technická omezení při snímání drážky

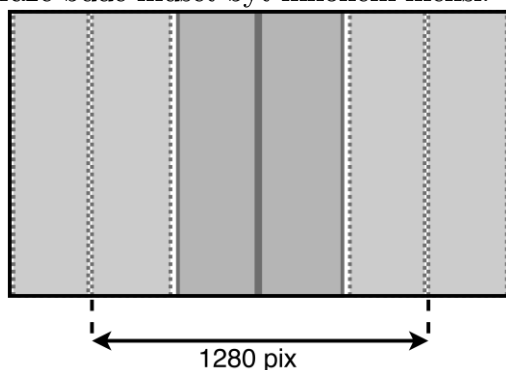
Ke snímání drážky kamerou lze přistoupit několika technikami, lišícími se způsobem snímání, použitým hardwarem a tím pádem i technickými omezeními té které metody.

---

<sup>8</sup> [www.vinylrecorder.com/stereo.html](http://www.vinylrecorder.com/stereo.html) (cit. 30.12.2016)

### 4.3.1 Snímání rastrovou kamerou

Mějme k dispozici standardní kameru s rozlišením 1920 na 1080 pixelů (Full HD). Berme v úvahu nejjednodušší možnost převodu a to snímání jedné drážky na celou šířku záběru. Pro další zjednodušení zatím stačí brát v úvahu jednokanálový signál. Jestliže standardní drážka bude mít základní šířku  $45\mu\text{m}$  a maximální výchylka bude dosahovat stejné hodnoty, znamená to, že na plnou šířku jednoho snímku potřebujeme umístit  $135\mu\text{m}$  reálné šířky, přičemž pouze dvě třetiny obrazu, tedy 1280 pixelů mohou být využity k výpočtu horizontální výchylky. To při přepočtu na bitovou hloubku odpovídá pouze hodnotě  $\log_2(1280) = 10,32$  bitů. Což je hodnota možná dostačující na hledání problémů a chyb, nikoliv pak na reprodukci v kvalitě srovnatelné s tradičními metodami. Po přibližném přepočtu sice vyjde dynamický rozsah cca 60 dB, je ovšem potřeba počítat s tím, že tohoto rozlišení zdaleka nedosáhneme na všech frekvencích rovnoměrně. Zároveň lze současně tvrdit, že toto maximální zvětšení (naznačené na obrázku 4.2) zdaleka nebude použitelné univerzálně a reálná velikost drážky v obraze bude muset být mnohem menší.



Obr. 4.2: Naznačení maximální možné horizontální výchylky ve FullHD obraze

Obdobným způsobem je možné určit maximální přenositelnou frekvenci při tomto způsobu snímání. Již jsme naznačili, že maximální převodní konstanta při tomto způsobu snímání bude  $K = \frac{1920\text{px}}{135\mu\text{m}} = 14,22 \text{ px}\mu\text{m}^{-1}$ . Již v kapitole 2.3 jsme si ukázali, že nejmenší možná vlnová délka (tedy při kmitočtu 20 kHz u středu drážky desky s rychlostí 33 ot/min) dosahuje hodnoty  $8,72\mu\text{m}$ , po přepočtu tedy více než 124 pixelů, což znamená, že odpovídá vzorkovací frekvenci přes 2,48 MHz. Je tedy zřejmé, že problém nedostatečného vzorkování v časové ose skutečně nehrozí. Další dvě omezení se týkají snímkové frekvence a rychlosti snímání. Snímkové frekvence by byla větším problémem zejména u okraje desky, kde má drážka největší „obvod“ (uvozovky proto, že se jedná o spirálu, nikoliv kružnici), přičemž rozměr snímku zůstává stejný. Z dřívějších výpočtů známe jak maximální obvod desky a možné rychlosti otáček (Z tohoto hlediska pro nás je horší varianta se 45 otáčkami, proto nadále budu počítat s tou), tak výšku záznamu převedenou na mikrometry.

Počet potřebných snímků při překrytí 0% lze spočítat podle následující rovnice (4.3), kde  $N$  je celkový počet snímků,  $t$  doba jedné otočky,  $n$  počet otáček za minutu a  $r_d$  poloměr desky:

$$f_{min} = \frac{N}{t} = \frac{\frac{0}{1080}}{\frac{K}{60}} = \frac{2\pi n r_d \cdot 10000 \cdot K}{1080 \cdot 60} = \quad (4.3)$$

$$= \frac{2\pi \cdot 45 \cdot 15 \cdot 10000 \cdot 14,22}{1080 \cdot 60} = 9307 \text{ fps}$$

Pokud by tedy technologie měla být použitelná pro realtime přehrávání hudby, musela by být použita skutečně vysokorychlostní kamera a to neberu v úvahu nutnost velmi rychlých výpočtů.

Obdobný problém při snímání rotující desky nastává ohledně rychlosti závěrky. Pokud bychom se chtěli vyhnout rozmazání úplně (tedy aby za dobu otevření závěrky došlo k rotaci desky o maximálně 1 pixel, musela by doba závěrky být rovna maximálně  $\frac{1}{f_{min} \cdot 1080} = 99,5 \text{ ns}$ .

Možným řešením těchto problémů je použití menšího přiblížení, kdy s dvojnásobkem zobrazené délky a šířky (tedy čtyřnásobek plochy) se nám o polovinu sníží rozlišovací schopnost v amplitudové i časové souřadnici, ale zároveň se na polovinu sníží i nutná snímkovací frekvence a dvojnásobně prodlouží doba závěrky. V jedné z metod, kterými se tato práce bude zabývat, je přiblížení dokonce ještě výrazně nižší, nicméně stále to neumožňuje použití stávající dostupné techniky pro kontinuální snímání drážky.

### 4.3.2 Snímání řádkovou kamerou

Druhou – vhodnější – metodou je použití řádkové CCD kamery. Jak název napovídá, jedná se o takový typ kamery, která obsahuje pouze jeden řádek fotocitlivých pixelů [4]. Jsou vhodné v průmyslu jako senzory k měření šířky (například různých mezer), či dalším jednodimenzionálním aplikacím, případně v rámci postupného skenování i aplikacím plošným.

Jednou ze základních výhod řádkových kamer pro danou aplikaci je skutečnost, že díky specifické konstrukci, kde namísto celé matice pixelů obsahují pouze jeden řádek, mohou tyto snímací buňky být větší než u kamer plošných, což v důsledku znamená více dopadajícího světla a tím pádem snazší snímání.

Mezi nejpodstatnější parametry těchto kamer patří rozlišení (počet pixelů), které v případě snímání drážky má přímý vliv na bitovou hloubku rekonstruovaného zvukového signálu (viz předchozí kapitola). Dá se tedy říct, že co do snímání rozměru, kolmého ke směru pohybu desky, fungují oba typy kamer totožně. Rozdíl je ovšem ve snímání v časové oblasti. Zatímco plošná kamera v tomto směru má určité rozlišení (v příkladu se jedná o 1080 pixelů), které v ideálním případě „vyfotí“ současně a je potřeba následně řešit skládání

jednotlivých snímků dohromady (i přes zanedbání tzv. rolling shutter efektu, kde ve skutečnosti není celý snímek sejmuto současně – například u kamery použité pro praktickou demonstraci práce se tento efekt projevuje velmi výrazně), v případě řádkové kamery se bude jednat spíše o klasické vzorkování v časové oblasti.

Druhým podstatným parametrem je tedy maximální řádková frekvence kamery, která lze spočítat pomocí vzorce (4.4), kde  $f_P$  je frekvence pixelů,  $N$  počet pixelů a  $N_p$  pak počet pasivních pixelů [4].

$$f_L = \frac{f_P}{N + N_p} \text{ [Hz]} \quad (4.4)$$

Jak nám říká Nyquistův (vzorkovací) teorém, platí, že maximální rekonstruovatelná frekvence navzorkovaného signálu se rovná polovině vzorkovacího kmitočtu. Pro nás to pak znamená, že pokud chceme chyby detekovat v reálném čase či dokonce kameru použít pro přehrání zvuku (ač nám v tom bude bránit omezená bitová hloubka), je potřeba zvolit kameru s řádkovou frekvencí nejméně 40 kHz. V ideálním případě pak přímo 44,1 kHz a více. Rizikem takového vzorkování je skutečnost, že nemáme (v časové oblasti) povrch desky sejmuto tak detailně, jako v případě plošné kamery a proto, je možné, že některé drobné mechanické vady nosiče tak zůstanou optickou kontrolou nezachyceny a neidentifikovány. Na základě vzorců (2.1 a 2.2) v kapitole 2.3, že při realtime vzorkování rychlostí 44100 Hz se fyzická vzdálenost jednotlivých vzorků může pohybovat mezi 3,96  $\mu\text{m}$  u středu desky s rychlostí 33,3 otáček a cca 16  $\mu\text{m}$  na okraji desky s rychlostí 45 otáček za minutu.

### 4.3.3 Srovnání obou metod z hlediska parametrů

Jak je zřejmé z rozebraného popisu obou zařízení, každý typ snímače poskytuje výhody a nevýhody v jiné oblasti a je schopen dosáhnout jiných parametrů. Rastrová kamera zachytí celý obraz, ale je natolik pomalá, že není možnost skenovat desku v reálném čase přímo při rotaci. Rychlý řádkový snímač tuto možnost poskytnout dokáže, ale pracuje s jednotlivými řádky zvlášť bez kontextu zbytku obrazu. V následující tabulce 1 jsou shrnuty některé maximální možnosti dvou přístupů, se kterými je v práci dále nakládáno. Prvním z nich je použití rastrové kamery s fullHD rozlišením a zoomem nastaveným tak, že 1920 pixelům odpovídá cca 720  $\mu\text{m}$  a snímá se postupně snímek po snímku. Druhý pak naznačuje možnosti použití řádkového senzoru o rozlišení 2048 pixelů, na kterém je umístěn rozsah cca 150  $\mu\text{m}$ . V obou případech se počítá s maximální výhylkou odpovídající 100  $\mu\text{m}$ .

Tabulka 1: Předpokládané parametry rekonstruovaného zvuku:

	Rastrová kamera	Řádková kamera
Rozlišení [px, $\mu\text{m}$ ]	1920, 720	2048, 150
Převodová konstanta [pix/ $\mu\text{m}$ ]	2,67	13,65
Maximální výchylka [px]	266,7	1365,3
Počet bitů na výchylku [bit]	8,06	10,41
Možné SNR [dB]	50,27	64,46
Minimální vzorkovací frekvence (u středu desky s rychlostí 33,3 ot/min) [Hz]	464 258	44 100

## Kapitola 5

# Ostatní teoretické principy

Tato kapitola shrnuje další teoretické poznatky a principy, které sice přímo nesouvisí s technologií gramofonové desky, nicméně je potřeba je znát, neboť jsou využívány pro samotné měření dat, či jejich následné vyhodnocování. Zejména se jedná o teorii detekce hran a princip použití krokových motorů.

## 5.1 Principy detekce hran v obraze

V předchozích kapitolách věnovaných principu mechanického záznamu a podobně drážek gramofonové desky se píše, že pro obnovu zvukové informace z tohoto mechanického principu není třeba znát každý mikrometr záznamového média, ale že slušně postačí vědět, kde má drážka své okraje (dokonce ani tedy není třeba vědět, kde se nachází její dno). Již ze slova „okraje“ je tak zřejmé, že při optické detekci se tak uplatní možnosti hledání hran a další operace právě s hranami. V této kapitole se proto budu věnovat tomu, co to hrany vlastně jsou z hlediska čísel v rastrovém obraze, proč jsou obecně mnohem významnější než zbytek obrazu (například než jednobarevné plochy), jak je lze vyhledávat a co to samotné vyhledávání znamená opět z hlediska čísel, potažmo tedy matematiky, které je implementována v softwaru určeném pro zpracování obrazu (např. Adobe Photoshop) či obecném softwaru určeném pro veškeré maticové operace (Matlab).

### 5.1.1 K čemu je detekce hran vhodná a proč ji použít

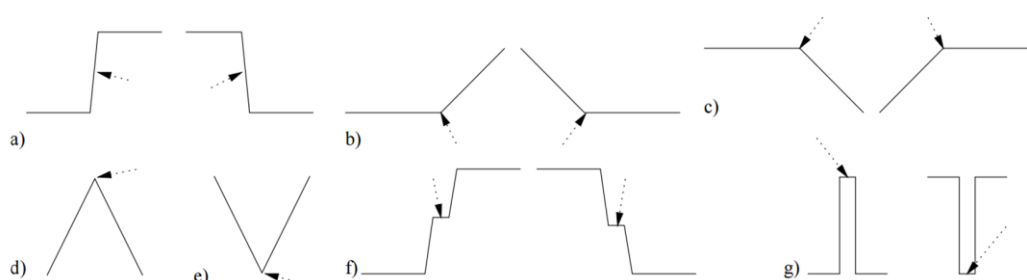
Významným nosičem informací při zpracování obrázků a vyhledávání objektů jsou tak hrany objektů. Zjednodušeně řečeno se jedná o místa, kde se skokově (či obecně prudčeji) mění číselná hodnota jasů pixelů v obraze. Základní výhodou tak je, že v určitých mezích je detekce hran relativně nezávislá na spoustě proměnných elementů scény, kterými je například osvětlení či v případě drážky odrazy světla od vnitřních ploch drážky v závislosti na úhlu natočení. Jas jednotlivých částí scény se tedy může zcela změnit, nicméně výstupem hranového detektoru stále bude tatáž hrana, neboť není podstatné, jakým způsobem se tento jas mění.

Využití detekci hran je tedy vhodné zejména při hledání objektů (či přímo samotných hran a čar) v obraze, pokud není podstatný samotný zbytek obrazu. „*Hrany vznikají díky nespojitostem v normále k povrchu, hloubce, odrazivosti povrchu (barvě), odleskům nebo nespojitostech v osvětlení (stínům)*“ [15]. V praxi

se pak uplatní například v umělé inteligenci, typicky se jedná o různé roboty, které potřebují ke svému fungování informace o svém okolí, o objektech a překážkách.

### 5.1.2 Hrana z hlediska matematiky

Co je to hrana každý člověk samozřejmě ví ze zkušenosti a je to pro něj přirozené. V oblasti zpracování obrazu nicméně počítač žádnou takovou zkušenost nemá a je proto třeba mu tuto znalost vnutit matematicky (stejně jako veškeré zpracování bude matematicky probíhat). Jelikož se jedná o změnu úrovně jasu, jde tedy o relativně prudký nárůst či pokles hodnoty. Ve speciálních případech, jako je například hrana střešová pak o kombinaci obojího.



Obr. 5.1: Různé typy hran – idealizované: a) skoková hrana, b), c) konvexní a konkávní svahová h., d) střešová h., e) údolní h., f) schodišťová h., g.) špičková h. (převzato z [15])

V případě těchto idealizovaných hran ve spojitém 1D prostoru (tedy pokud se zabýváme pouze jedním směrem, ve kterém se mění hodnota jasu, nikoliv celým dvourozměrným obrázkem, který je navíc diskrétní) není problém hranu najít pomocí operace derivace (například v případě a) v místě hrany první derivace dosáhne svého maxima, zatímco derivace ve všech ostatních bodech bude nulová). Zde už jsem částečně naznačil problematiku hledání hran v 2D obrázcích, kde na rozdíl od 1D vektoru je třeba určit kromě skutečnosti, že se v tomto místě hrana nachází také směr hrany (na který je tak kolmý její gradient).

### 5.1.3 Gradient

Gradient obecně lze vyjádřit jako směr růstu. Ve vektorovém poli je matematicky definován jako vektor parciálních derivací (vzorec (5.1))

$$\nabla f(x_1, x_2, \dots, x_n) = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right) \quad (5.1)$$

V případě zpracování obrázků se jedná o pole dvourozměrné, tedy lze upravit na podobu  $\nabla I(x, y) = \left( \frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right)$ , kde I je obrázek (image).



Velikost gradientu hrany a jeho směr pak spočítám podle vzorců (5.2 a 5.3).

$$|\nabla I(x, y)| = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} = \sqrt{G_x^2 + G_y^2} \quad (5.2)$$

$$\theta = \tan^{-1}\left(\frac{G_y}{G_x}\right) \quad (5.3)$$

### 5.1.4 2D konvoluce

Kromě gradientu se v algoritmech detekce hran uplatňuje také konvoluce. V rámci zpracování obrazu se jedná konkrétně o diskretní 2D konvoluci. Aby bylo možné definovat operaci konvoluce, je nutné předem definovat pojmy jednotkový impuls a impulzní odezvu systému.

Jednotkový impuls ( $\delta[]$ ) je v diskretních 1D systémech určen jako funkce, jejíž první hodnota je rovna 1 a ostatní pak 0. Impulzní odezva ( $h[n]$ ) je pak výstupní signál lineární soustavy, na jejímž vstupu se nachází právě jednotkový impuls. Konvoluci v diskretním čase pak lze definovat pomocí impulzní odezvy vzorcem (5.4), její 2D obdobu pak vzorcem (5.5).

$$y[n] = x[n] * h[n] = \sum_{k=-\infty}^{\infty} (x[n-k]h[k]) \quad (5.4)$$

$$y[n_1, n_2] = \sum_{k_1=-\infty}^{\infty} \left( \sum_{k_2=-\infty}^{\infty} (x[n_1 - k_1, n_2 - k_2]h[k_1, k_2]) \right) \quad (5.5)$$

2D impulzní odezva  $h[n_1, n_2]$  (ve 2D nazývaná PSF – point spread function) se zde pak nazývá *konvoluční jádro* či *konvoluční maska*. Obdobně jako ve zpracování zvuku je možné konvolucí řešit filtraci signálu, ať už se jedná o filtry hornopropustní (zostření) či dolnopropustní (rozmazání) [20]. 2D konvoluce se využívá pro úpravu snímku před samotnou detekcí hrany (odstranění šumu či naopak zostření hran).

### 5.1.5 Algoritmy detekce hrany

Většina hranových detektorů proto pracuje na třech základních principech [15]:

1. Hledání maxima první derivace
2. Hledání bodu průchodu druhé derivace nulou
3. Lokální aproximace určitým modelem (Vhodné, když víme, co hledat)

V praxi digitálního zpracování obrazu ovšem pracujeme zásadně s diskretními funkcemi, a jelikož jednou z obecně známých základních podmínek derivace je spojitost signálu, je třeba učinit určitá opatření a tuto spojitost zajistit. Možnosti jsou v zásadě dvě, vedoucí k podobným výsledkům: Můžeme

buď rekonstruovat spojitou funkci a z ní pak spočítat derivaci nebo můžeme spočítat jednotlivé diference a nadále pracovat s nimi.

Druhým základním problémem použití derivace je její citlivost na šum (tzn., že v místech kde oko a lidský mozek vidí zrnitou plochu, počítač vidí spoustu skokových změn hodnoty, což znamená špiček v derivaci (či diferenci v případě diskrétních systémů), a proto je třeba vektor nejprve vyhladit pomocí konvoluce a následně teprve počítat derivaci či diferenci. Díky linearitě operace konvoluce a vlastnostem obou operací (komutativita a asociativita) je lze shrnout do jediného operátoru konvoluce, kde konvoluční jádro již bude zderivované [15], dle vzorce (5.6), kde  $h$  je konvoluční jádro vyhlazovací funkce a  $f$  samotný vektor, kde vyhledáváme hranu.

$$\frac{d}{dx}(h * f) = \frac{dh}{dx} * f \quad (5.6)$$

Jednotlivé detekční metody, implementované např. v Matlabu ve funkci *edge()* se pak liší zejména podobou konvolučního jádra, které metodu předurčuje k úspěchu či neúspěchu pro určité typy hran. Typická konvoluční jádra mají rozměr  $3 \times 3$  a každá metoda obsahuje 2 základní konvoluční jádra pro horizontální a vertikální směr. Následující vzorce ukazují dvě varianty – horizontální a vertikální operátoru Prewittové (5.7) a Kirschova (5.8) [20].

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, G_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad (5.7)$$

$$G_x = \begin{bmatrix} -5 & 3 & 3 \\ -5 & 0 & 3 \\ -5 & 3 & 3 \end{bmatrix}, G_y = \begin{bmatrix} 3 & 3 & 3 \\ 3 & 0 & 3 \\ -5 & -5 & -5 \end{bmatrix} \quad (5.8)$$

Tato konvoluční jádra patří spolu s metodami Sobel a Canny mezi tzv. detektory založené na gradientu neboli detektory dle první derivace. V předchozím výčtu se tedy jedná o skupinu 1. Mezi detektory dle druhé derivace (hledání průchodu nulou) pak patří Laplacián Gaussiánu (LoG – Laplacian of Gaussian). Laplacián lze definovat dle vzorce (5.9) [17] – Vidíme tedy, že se jedná o druhou derivaci nezávislou na směru (jedno číslo pro 2D matici).

$$\nabla^2 I = \left( \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \right) \quad (5.9)$$

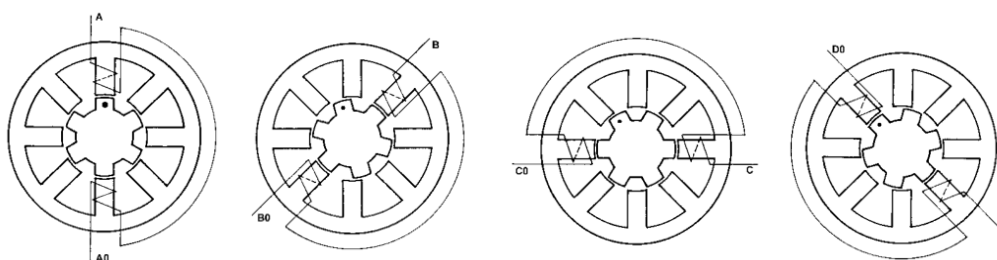
V praktické části práce používám pro detekci hran metodu Prewittové ve vertikálním směru, neboť je cílem najít hrany drážky, které jsou takto orientované. Pro vyhledávání prasklin a dalších vad pak naopak může být vhodnější detektor horizontální.

## 5.2 Krokové motory a jejich řízení

Jelikož pro snímání desky pro ověření funkčnosti algoritmu je zvolena metoda, kdy snímací kamera je statická a pohybuje se samotným otočným talířem s originálem budoucí desky, je potřeba zajistit jeho rotaci s parametry, co nejhodnějšími pro snímání. Já jsem pro tento úkol zvolil krokový motor. V praktické části se jím pak dále zabývám přímo z hlediska návrhu konstrukce měřicí aparatury a považuji ho za uzavřenou jednotku s pevně danými vlastnostmi, proto by bylo užitečné zde probrat, proč jsem zvolil tento systém pohonu a jak motor funguje.

V zásadě existují tři přijatelné varianty motorů, což je klasický stejnosměrný elektromotor, servomotor a právě krokový motor. Pro výběr motoru vhodného právě pro konkrétní aplikaci tak je potřeba určit si, jaké požadavky na systém máme a jaké jsou vlastnosti jednotlivých motorů. Pro aplikaci, jíž se zabývám v této práci, nejsou potřeba vysoké otáčky, které by ostatně dokázal zajistit vestavěný motor s variantami 33 či 45 otáček za minutu s možností volitelného přepnutí na rychlost poloviční. Naopak pro snímání jednotlivých snímků vyžadujeme možnost posunu o malý úhel (či vzdálenost v tečném směru) a vyžadujeme možnost zastavení rotace ve vhodném okamžiku. Tedy je zjevné, že krokový motor bude nejhodnější varianta.

Krokové motory fungují v principu tak, že součástí statoru je (v našem případě) 8 pólových nástavců s navinutými cívkami, kterými prochází elektrický proud a součástí rotoru pak o 2 menší počet železných pólových nástavců. Protilehlé cívký statoru tvoří vždy jednu fázi. Postupným přiváděním proudu do jednotlivých fází pak dojde k postupnému skokovému otáčení rotoru [13]. Princip pro tento modelový motor je uveden na následujícím obrázku 5.2.



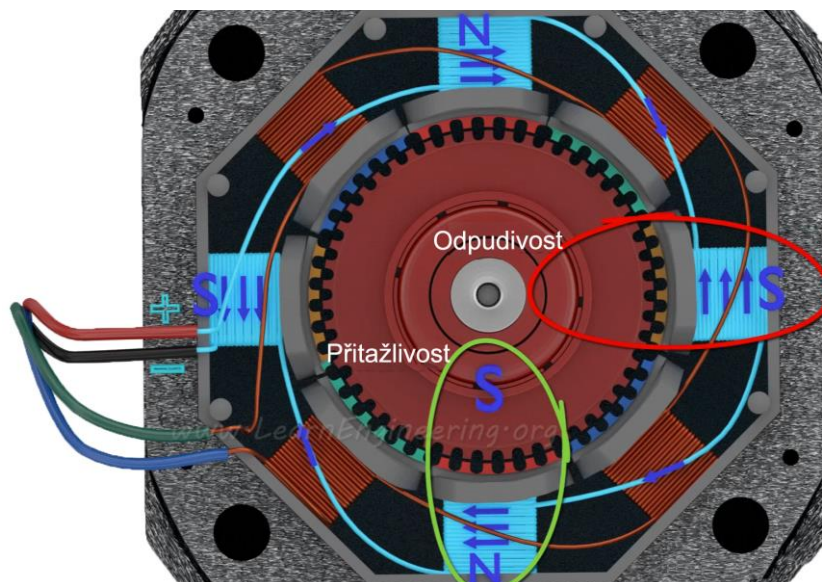
Obr. 5.2: Princip postupného buzení fází a krokové rotace (převzato z [13])

Z obrázku je patrné, že tento modelový motor dokáže vykonat otočku o úhel  $\frac{360}{8} = 45^\circ$  ve 3 krocích, tedy znamená to, že plnou otáčku motor dokáže rozdělit do 24 kroků (krok je tedy roven  $15^\circ$ ). Většina dnešních komerčně prodávaných motorů – tedy i ten, který v praktické části sám používám) je konstrukčně řešena tak, že při zachování 8 cívek jsou tyto zapojeny obvykle pouze do dvou fází s tím, že počet pólových nástavců na cívkách statoru a na rotoru je volen takovým způsobem, aby stejná posloupnost budících proudů způsobila menší rotaci

motoru (sousední rovnovážné polohy jsou si tak mnohem blíže). Typický krokový motor tak disponuje krokem o velikosti  $1,8^\circ$ , který jde dále zmenšovat pomocí inteligentního systému řízení, tzv. mikrokrokování – viz následující kapitola.

### 5.2.1 Řízení krokového motoru pomocí Arduina a H-můstku

Zmíněné motory, které se používají v běžné praxi, tedy i v této práci mají rotor tvořený stálým magnetem s nasazenými koncovkami o 50 pólových nástavcích na každé straně, přičemž nástavce na opačných pólech jsou od sebe pootočené tak, aby zub na jednom pólu odpovídal mezeře na pólu druhém. Cívky jsou zapojeny tak, že každá čtveřice cívek odpovídá jedné fázi, nicméně vždy dvojice a dvojice má opačnou polaritu. Postupným přiváděním proudu na jednotlivé fáze tak umožníme pootáčení rotoru o zmíněných  $1,8^\circ$  [14]. Stavba motoru a nákres cívek a pólových nástavců je na obrázku 5.3.



Obr. 5.3: Stavba typického krokového motoru s rozlišením 200 kroků (převzato z [14])

Pokud nehodláme využít možnosti mikrokrokování, existují v podstatě 3 základní způsoby, jak krokový motor řídit: jednofázový, dvoufázový a kombinovaný (half-stepping). Jednofázový režim přivádí impulzy na jednotlivé fáze tak, že při daném okamžiku je aktivní současně vždy jen jedna, což ústí v jednoduchý postupný pohyb s rovnoměrným krouticím momentem během celé rotace. Dvoufázový režim zapojuje vždy sousední dvojici cívek tak, že rovnovážná poloha je oproti jednofázovému režimu pootočena o polovinu kroku. Tento režim se vyznačuje dvojnásobným odběrem proudu, ale také dvojnásobným krouticím momentem. Posledním režimem je režim kombinovaný, kde dochází ke střídavému aktivování jedné a dvou cívek, tedy získáme možnost polovičního kroku, ale musíme počítat s tím, že jednotlivé polohy se nevyznačují stejným krouticím momentem – je to dáno tím, že je stator

a rotor přitahován a odpuzován střídavě jednou a dvěma fázemi a tedy tato magnetická síla není konstantní. Na následující tabulce 2 je ukázáno, v jakém pořadí a jakým způsobem jsou jednotlivé cívky buzeny – fáze, kterou aktuálně protéká elektrický proud je vyznačena symbolem „1“:

Tabulka 2: Logické schéma fází krokového motoru:

Fáze	Jednofázový r.				Dvoufázový r.				Kombinovaný režim								
A	1	-	-	-	1	-	-	1	1	-	-	-	-	-	-	1	1
B	-	1	-	-	1	1	-	-	1	1	1	-	-	-	-	-	-
!A	-	-	1	-	-	1	1	-	-	-	1	1	1	-	-	-	-
!B	-	-	-	1	-	-	1	1	-	-	-	-	1	1	1	1	-

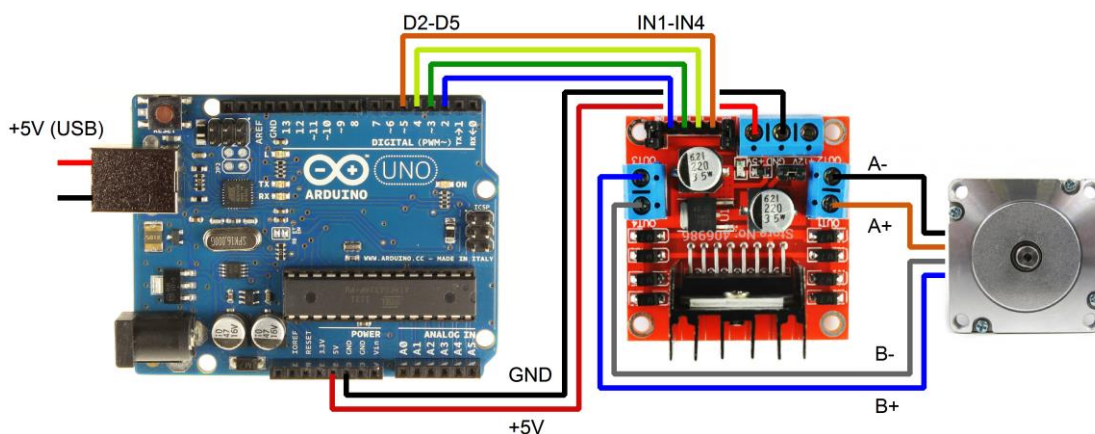
Mikrokrokování je pak speciální režim řízení motoru, kde je využita PWM modulace, pomocí které je možné umístit rovnovážnou polohu i mezi standardní polohy, které umožňuje výše zmíněné řízení. V praxi se používá spíše pro dosažení plynulejší rotace, než přímo pro využití mezipoloh, jakožto přesně daného úhlového natočení. Více se jím v této práci nebudu zabývat, neboť pro účel otáčení talíře není mikrokrokování potřeba a jelikož je stejně nutné použít převodový mechanismus, bude lepší přesnost zvyšovat pomocí něj (více v praktické části práce).

Jelikož od motoru nevyžadujeme žádné složité řízení, jako je třeba nastavení konkrétního úhlu, mikrokrokování a podobné, zvolil jsem jako řídicí platformu Arduino<sup>9</sup>, což je notoricky známý mikropočítač, založený na mikrokontroleru ATmega. Vyznačuje se tím, že je to snadno a levně dostupná platforma se speciálně vyvinutým programovacím prostředím. Díky tomu, že Arduino UNO obsahuje 14 digitálních vstupně-výstupních pinů (z toho 6 s podporou pulzní šířkové modulace) a 6 analogových vstupů, bývá často používáno jako základní stavební kámen projektů, kde je potřeba načíst přítomnost či hodnotu signálu z nějakého senzoru a po vyhodnocení provést nějakou operaci. Během studia jsem osobně Arduino využil 2x, jednou se jednalo o koncept navigační přilby pro zrakově a koordinačně postižené, která pomocí ultrazvukových senzorů a akcelerometru vyhodnocovala možná rizika v okolí a svého nositele na ně upozornila akustickým signálem. Podruhé se jednalo o infračervený snímač vzdálenosti.

Pro řízení motoru tedy stačí vyhradit čtyři digitální výstupy a přivádět na ně signál dle tabulky (5.1) výše. Výhodou řízení pomocí Arduina je skutečnost, že je možné mezi jednotlivé kroky vložit libovolné zpoždění a tak například zajistit synchronizaci s rychlostí kamery, zajistit dobu potřebnou na ustálení talíře vlivem setrvačnosti po samotném kroku, apod.

<sup>9</sup> Více informací na oficiálním webu projektu <https://www.arduino.cc/> (cit. 28.12.2016)

Jelikož maximální výstupní proud z jednoho portu Arduina je 40 mA, nelze na jeho výstupy připojit přímo vinutí cívek. Je třeba využít nějakého mezistupně, který dokáže dodat potřebný proud. Tento obvod lze samozřejmě navrhnout a zkonstruovat z tranzistorů, nicméně já jsem zvolil již hotový modul duálního H-můstku okolo integrovaného obvodu L298<sup>10</sup>. Na rozdíl od různých inteligentních driverů je tak třeba veškeré logické posloupnosti vyslat už z Arduina. Na obrázku 5.4 je naznačeno schéma zapojení Arduina, použitého H-můstku a krokového motoru. Napájení obvodu je realizováno 5V, které dokáže dodat Arduino napájené externím zdrojem v podobě USB powerbanky. Druhou testovanou možností napájení bylo použití spínaného zdroje 12V zapojeného do driveru s tím, že jeho druhý vstup je použit jako zdroj 5V pro Arduino. Toto zapojení sice disponuje větším krouticím momentem, ale způsobuje poměrně velké zahřívání celého systému.



Obr. 5.4: Schéma zapojení krokového motoru do Arduina přes H-můstek

<sup>10</sup> Jedná se o modul určený jako zdroj dostatečného proudu pro krokové motory a DC motory, řízený logickými úrovněmi TTL standardu.

## Kapitola 6

# Obecný návrh experimentálního zařízení

Cílem práce je praktická demonstrace navržené metody snímání nosiče a převod obrazové informace o podobě drážky na informaci zvukovou. To se samozřejmě kromě programu, který data zpracuje, neobejde bez hardwarového řešení, které daná data nasbírá v podobě co nejlépe čitelné a zpracovatelné. Tato kapitola se tedy bude věnovat návrhu snímací aparatury a jejím parametrům na základě ostatních prací, které se tématem zabývaly v minulosti, vlastních výpočtů na základě požadovaných parametrů a v neposlední řadě pak v závislosti na dostupné technice a technologii.

## 6.1 Princip konstrukce externího pohonu talíře

Pokud bychom navzdory omezením uvedeným v kapitole 4.3.1 chtěli přesto použít kameru rastrovou i za cenu ztráty možnosti realtime čtení signálu (které by poskytovalo základní výhodu v tom, že by bylo možné optickou stopu signálu načíst již v procesu řezání měděného originálu), je vhodné jednotlivé snímky načítat postupně v čase, který bude řádově delší než je hrací doba desky. Mezi jednoznačné výhody tohoto řešení patří možnost pořízené snímky ihned zpracovávat a šetřit tak značně místo na pevném disku, mezi nevýhody pak samozřejmě ještě větší časová náročnost (Nebylo by možné během zpracování snímků hned načítat snímky další desky).

Toto řešení přesahuje rámec standardního vybavení řezací aparatury, nicméně umožňuje využít její mechanický základ a být tak jen jakousi neinvazivní nadstavbou stroje. Cílem návrhu je externí krokový pohon talíře, umožňující otočení talíře s deskou o specifický úhel, který umožní pořizovat jednotlivé statické snímky s právě takovým posunem, aby jejich překryv byl dostatečný pro rozumné zpracování snímků s minimálním množstvím redundantní informace. Základem řešení se samozřejmě stává krokový elektromotor, jehož fungování a řízení je společně s důvody jeho zvolení popsáno dříve v teoretické části. Jelikož tyto motory mají počet kroků v jedné otáčce roven typicky hodnotě 200, je zřejmé, že bude potřeba využít nějaký převodový systém. Jak kvůli tomu, že otočka o  $1,8^\circ$  je vůči rozměrům drážky a zornému poli optické soustavy kamera-mikroskop, příliš veliká, tak i proto, že konstrukce řezací aparatury neumožňuje výměnu pohonu, aby bylo možné deskou pohybovat napřímo.

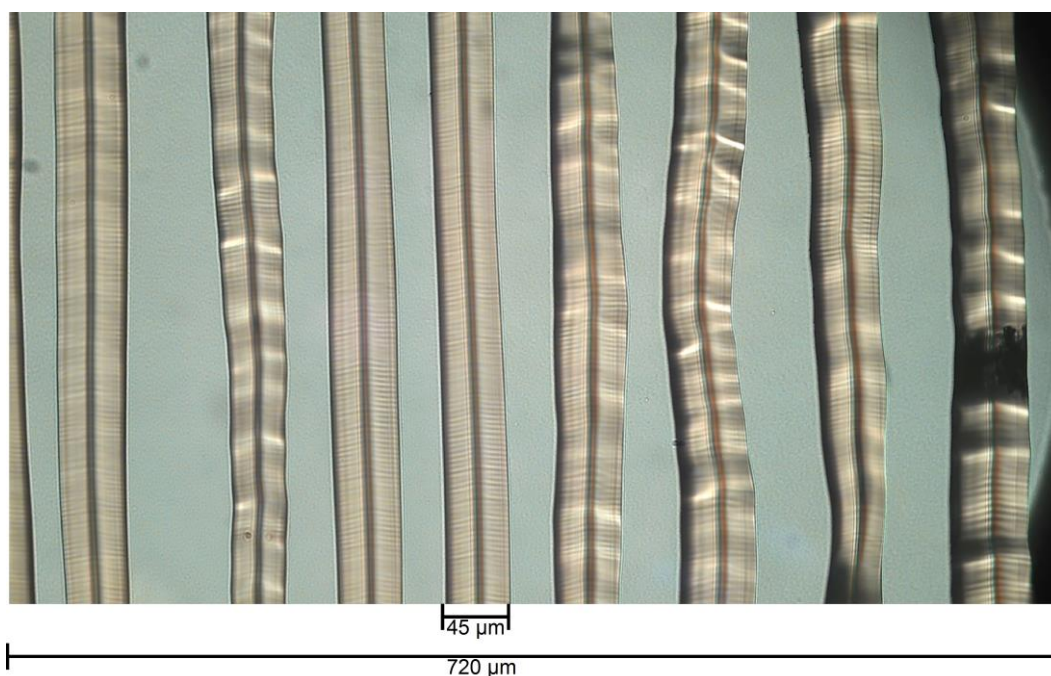
Externí pohon pak využívá skutečnost, že pokud je na stroji vypnuta funkce otáčení talíře, tak se tento může volně otáčet a tedy s ním může být volně



otáčeno jakoukoliv vnější silou (ať už se jedná o obsluhujícího operátora nebo zcela samostatný stroj). Fotografie aparatury se nachází v kapitole 2.2 na fotografii 2.2.

## 6.2 Výpočet převodu pohonného mechanismu

V předchozí kapitole jsme si stanovili, že při použití rastrové kamery se budeme zabývat takovým způsobem snímání obrazu, kde se jednotlivé snímky budou překrývat a bude tedy pro jejich další zpracování nutné aplikovat pokročilejší algoritmy pro detekci překryvů a jejich následné vyhodnocování. Je tedy nutné předem spočítat minimální možnou převodovou konstantu pohonného systému tak, aby i na krajích desky došlo k dostatečnému překryvu snímků právě z důvodu správného určení překryvu a následného spojení snímků. Jelikož tato konstanta samozřejmě závisí na zorném úhlu sestavy kamery-mikroskop (dále jen „kamera“), může se velice lišit pro demonstrační realizaci a pro následné finální řešení, použitelné ve výrobním procesu. Zorný úhel je samozřejmě analyticky vypočitatelný z parametrů optické soustavy (které povětšinou neznám, neboť větší část technického vybavení včetně mikroskopu a kamery je zapůjčená od GZ Media, kde neslouží k měření konkrétních parametrů), nicméně pro demonstraci metody postačí odvodit přibližnou hodnotu z pořízených zkušebních snímků (obrázek 6.1).



Obr. 6.1: Určení reálného rozlišení snímku v  $\mu\text{m}$

Z rozlišení snímku 1080p a známé základní šířky drážky 45  $\mu\text{m}$  byla změřena šířka drážky v pixelech přímo ve snímku rovna 120 px. Z toho lze snadno trojčlenkou odvodit, že do jednoho snímku lze při otočení kamery o  $90^\circ$  umístit



současně až  $\frac{1920}{120} \times 45 = 720 \mu\text{m}$  drážky. Bez otočení (tedy při snímání menšího kusu více drážek – tak jak ukazuje snímek výše) to potom bude  $405 \mu\text{m}$ . Pokud pak vhodný a bezpečný překryv stanovíme na necelou třetinu snímku, vyjde nám, že jednotlivý snímek by měl obsahovat nejméně  $480 \mu\text{m}$ , respektive  $270 \mu\text{m}$  jedinečné informace. Dále je třeba spočítat celkový obvod desky na jejím okraji, neboť motor zajistí konstantní úhlový posun, tedy směrem ke středu, kde se obvod drážky zmenšuje, bude docházet k nárůstu překryvu snímků. Jak je uvedeno v kapitole 2.3, nájezdová drážka desky začíná na průměru  $30 \text{ cm}$ , tedy dle vzorce pro obvod kružnice je tento roven cca  $94,3 \text{ cm}$ .

Znalostí ekvivalentního rozměru snímku v metrech a obvodu desky taktéž v metrech je pak snadné spočítat, že po obvodu první drážky při splnění všech podmínek pro všechny uvedené případy potřebujeme minimálně  $N = \frac{0,943}{2,7 \times 10^{-4}} \doteq 3493$  snímků pro snímání „na šířku“, které nakonec bylo využito.

Převodová konstanta celého mechanismu  $K$  tedy musí být určena jako poměr počtu požadovaných kroků na jednu otočku (3493) ku počtu kroků, které motor sám o sobě umožňuje (200). V tomto případě lze tedy stanovit konstantu  $K = \frac{3493}{200} = 17,46$

## 6.3 Realizace převodu pohonného mechanismu

Jak je uvedeno výše, mechanismus pohonu talíře aparatury by měl být pokud možno odnímatelný a zcela neinvazivní vůči samotné rezačce. V praxi pak připadají v úvahu dvě možnosti řešení: První z nich je přímý převod s invertujícím směrem, kde by talíř byl přímo poháněn motorem s převodním členem s povrchovou úpravou, zlepšující tření mezi dvěma systémy – například pogumovaný kotouč. Toto řešení může být velice elegantní pro převodové poměry v řádu jednotek, kde vzhledem k lineární závislosti rychlosti otáčení na obvodu a tedy i průměru převodového členu by hnací kotouč mohl mít rozměry v řádu centimetrů, nikoliv milimetrů. Jednoznačnou nevýhodou je ovšem velmi malá kontaktní plocha a z toho plynoucí riziko nepřilíš přesného přenosu rotace.

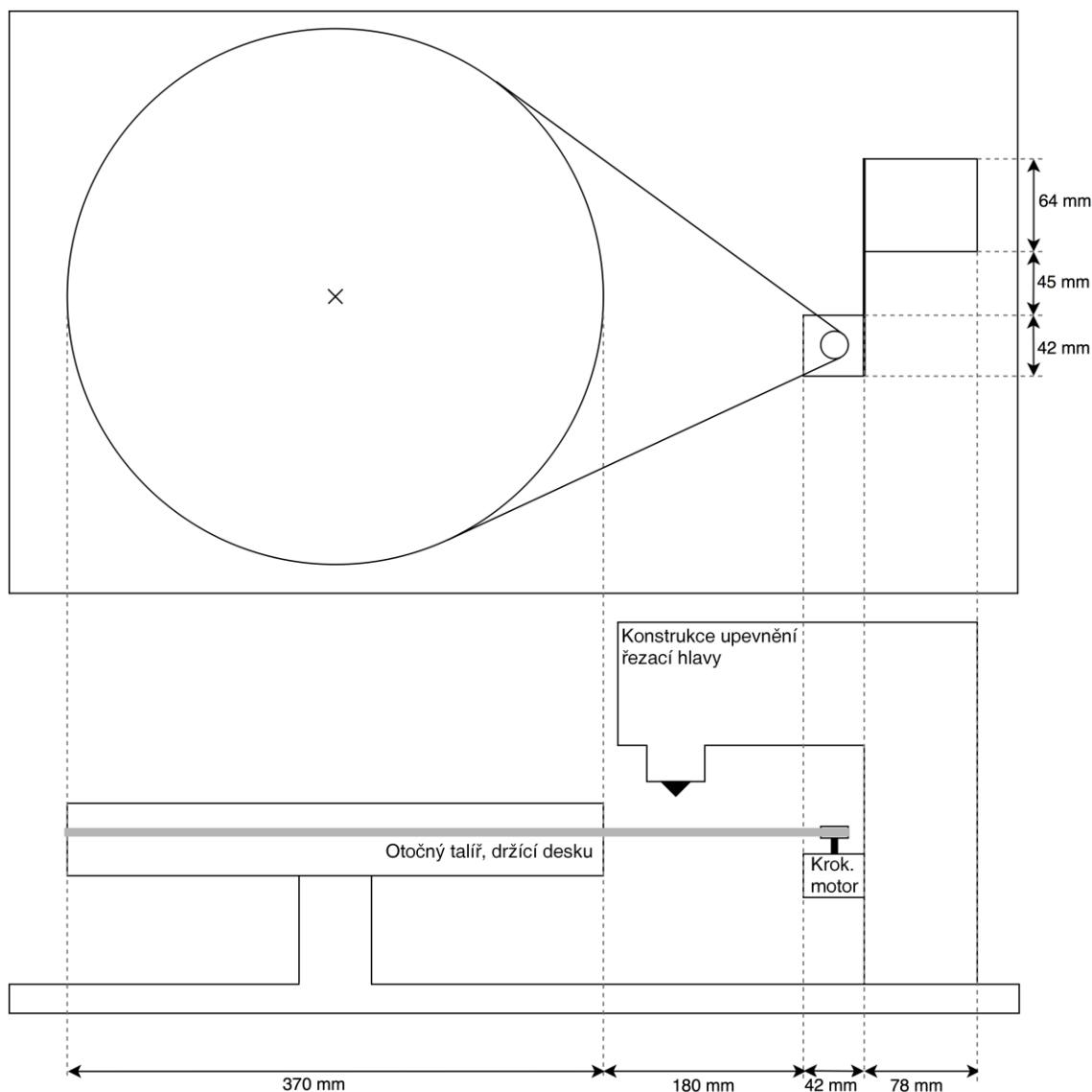
Druhá možnost přenosu rotační energie vychází z řemínkového náhonu známého z klasických gramofonů. Je zjevné, že tato metoda bude přenášet energii v opačném směru než náhon přímý, nicméně rozměry jednotlivých prvků zůstávají stejné (tj při jednoprvkovém převodu by opět poměr průměrů primární a sekundární řemenice musel být roven poměru požadovaných úhlů rotace). V případě potřeby velmi vysokých převodových poměrů by tak hrozilo, že primární řemenice by musela mít velmi malý průměr, což opět snižuje přesnost přenosu. V těchto případech by pak bylo více než vhodné použít vícestupňový převod, kde součin jednotlivých převodových poměrů bude roven požadovanému výslednému poměru.

S obdobným systémem se ostatně zabývá práce [10] s tím rozdílem, že v řešení zmíněného článku šlo především o to, jak výrazně zpomalit kontinuální rotaci synchronního motoru, která, jak autor uvádí, bohužel přenášela vibrace a ústila v konstantní šum do zvukového výstupu systému. V případě použití krokového motoru je možné toto riziko eliminovat inteligentním softwarovým řízením, kde by byl snímek pořízen během zastaveného pohonu. Bohužel se dá předpokládat, že nutnost častého zastavení a obnovení pohonu pro každý pořízený snímek (či rovnou i jeho dílčí nebo úplné zpracování) celou metodu výrazně zpomalí.

Pro realizaci metody v praxi byl nakonec zvolen převod řemínkový a to ozubený pomocí shodně ozubené řemenice připevněné ke krokovému motoru. Zvolen byl řemen typu T2.5, který disponuje rozestupem zubů 2,5 mm. Měřením bylo zjištěno, že průměr otočného talíře řezací aparatury je roven přesně 370 mm, což po přepočtu přes konstantu  $\pi$  znamená obvod 1162,4 mm. Ekvivalentní počet zubů, kdyby se jednalo o řemenici tak je roven 465. (Zde je potřeba upozornit na drobnou odchylku, které jsem se dopustil skutečností, že talíř zuby nemá a řemen do něj nezapadne tak, jako by zapadl do řemenice. Nicméně tato odchylka může způsobit větší překryv snímků nikoliv menší, není s ní tedy problém.)

Jelikož dostupné primární řemenice byly pouze s 16, 18 či 25 zuby, jsme se schopni dostat na převodové poměry 29,26, respektive 18,6 [-]. Všechny tři tak splňují požadavky na minimální převodový poměr a proto jsem zvolil variantu s 25 zuby.

Zpětným výpočtem lze určit, že při převodové konstantě 18,6 a rozlišení 200 kroků na primárním pohonu dosáhneme 3720 jedinečných snímků na jednu otáčku talíře. Během zpracování se ukázalo, že těch snímků je ve skutečnosti dokonce 3746. Odchylka tak dělá cca 0,7%, což je hodnota která může být způsobena lehkým protažením řemenu v průběhu měření, jeho nedokonalým spojením do uzavřené smyčky či právě zmíněnou absencí zubů otočného talíře. Obě varianty nicméně překračují stanovené bezpečné minimum a proto nehrozí, že v nějaké části desky dojde k nedostatečnému, či dokonce žádnému překryvu snímků, znemožňujícím zpětné složení snímků dohromady. Náskres pohonu v kontextu se zbytkem řezačky s vyznačenými rozměry, které bylo potřeba znát přímo pro konstrukci pohonu, je na obrázku 6.2.



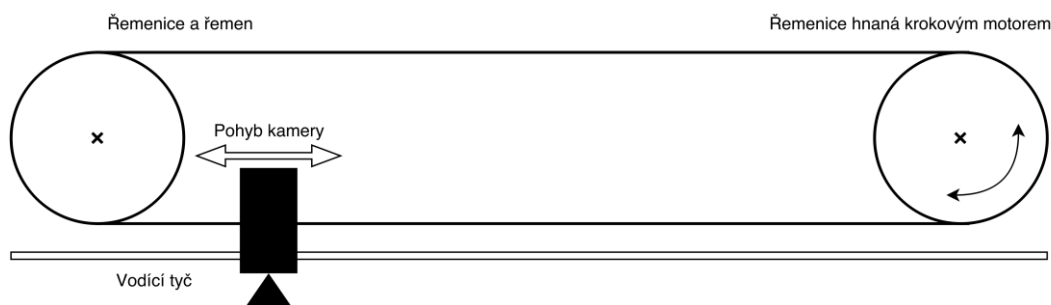
Obr. 6.2: Náskres konstrukce externího pohonu

## 6.4 Posuv kamery

Z rozměrů desky je ovšem zřejmé, že externí pohon otočného talíře není jediným pohyblivým prvkem, který je potřeba vyřešit; druhým takovým je pak pohyb samotného mikroskopu a kamery od okraje desky směrem k jejímu středu. Jak je vypočteno v kapitole 6.2, delší straně snímaného obrazu (1920 px) odpovídá cca 720  $\mu\text{m}$  reálné vzdálenosti. I v případě naprosto tiché drážky s minimálními rozměry (40  $\mu\text{m}$  drážky oddělené 10  $\mu\text{m}$  mezerami) lze tvrdit, že na jeden snímek, orientovaný tak, že kratší strana je rovnoběžná s drážkami (tato metoda snímání ostatně byla nakonec zvolena), lze umístit maximálně 14 drážek. Jak je zřejmé z testovacího snímku výše, v případě reálného záznamu se může jednat spíše tak o 8 drážek, z nichž u krajních dvou hrozí riziko výchytky větší, než dokáže kamera zaznamenat a tudíž nemožnost rekonstrukce signálu či výpadek.

V případě opačné orientace kamery lze mluvit o přibližně polovině, tedy 3-4 drážkách. To je v případě pomalejší varianty (tedy 33 otáček za minutu) tedy pouhých 5,4 vteřin záznamu. Z tohoto je již jasné, že pro nasnímání celé desky bude potřeba kamerou velice „často“ hýbat v tomto směru. Tento pohyb je řešitelný celou řadou možností. Ty, které připadají v úvahu, jsou popsány v kapitolách 6.4.1 až 6.4.3.

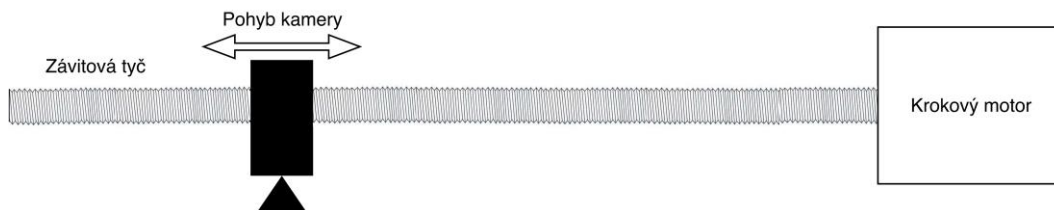
### 6.4.1 Páskový posuv řízený krokovým motorem



Obr. 6.3: Princip páskového posuvu

Tento systém využívá jednoduchého mechanismu, kde na koncích systému jsou dvě řemenice, z nichž jedna využívá pohonu krokovým motorem. K řemenu, který je mezi nimi napnut, je připevněn samotný mikroskop s kamerou. Velikost řemenic opět závisí na požadovaném rozlišení pohybu. Jestliže se budeme držet spočítané konstanty, kde 1920 px odpovídá 720  $\mu\text{m}$  a budeme chtít použít i možnost otočení kamery o 90°, pak lze přibližně stanovit, že bude třeba kameru posouvat s krokem maximálně cca polovinu kratší strany obrazu, tedy 540 px. Této hodnotě odpovídá reálná vzdálenost 202,5  $\mu\text{m}$ . Aby bylo možné tento mechanismus prakticky využít, potřebovali bychom řemenici o takovém rozměru, kde spočítaná vzdálenost vynásobená počtem kroků motoru (200) bude rovna obvodu řemenice. Ten bude tím pádem roven 40,5 mm. Po vydělení  $\pi$  pak získáme přímo průměr řemenice o rozměru 12,9 mm.

### 6.4.2 Využití závitové tyče řízené krokovým motorem



Obr. 6.4: Princip posuvu pomocí závitové tyče

Oproti jednoduchému pásku je toto řešení mechanicky složitější, nicméně výrazně přesnější. Rozlišení posunu lze určit jako poměr stoupání závitů použité tyče a počtu kroků samotného motoru. Při použití závitové tyče se stoupáním například 1,5 mm a motoru s 200 kroky pak získáme rozlišení 7,5  $\mu\text{m}$ , což několikanásobně splňuje potřebné rozlišení pro bezpečné překrytí snímků.

### 6.4.3 Využití modulačního signálu gramofonové desky

Ačkoliv jsem se tomu v teoretickém popisu výroby desek nevěnoval, neboť to pro účel práce nebylo nezbytně nutné, kromě samotného cílového audiosignálu vstupuje do řezačky ještě signál tzv. modulační. Již v době řezání z magnetického pásku obsahoval magnetofon zpoždovací smyčku o délce odpovídající polovině hrací doby desky. Sloužila k tomu, aby si stroj změřil charakteristiky signálu, který teprve bude následovat a přizpůsobil tomu parametry řezání, zejména pak rozstup drážek. Tedy pokud následující půlotáčku bude tichá pasáž, je možné posadit následující drážku mnohem blíže k drážce aktuální, než pokud bude následovat hlasitější pasáž s většími výchyly.

Dnes v době digitální přípravy celé desky, kdy lze pomocí analýz určit, jak bude celá deska vypadat mikrometr po mikrometru, je i tento modulační signál spočítán počítačem. Na řezačce pak velice přesně řídí pohyb řezací hlavy (jejíž kmity jsou pak prováděny vůči poloze určené právě tímto signálem). Pohyb řezací hlavy je pak velmi jemně řízen právě závitovou tyčí.

Teoreticky by proto mělo být možné podobu tohoto modulačního signálu, či přímo záznam o průběhu rychlosti otáček tyče, uložit do souboru a ten následně využít pro rekonstrukci tohoto pohybu na jiném stroji obdobné konstrukce, který by místo řezací hlavy obsahoval snímací kameru.

Tato metoda by tedy dosahovala zdaleka nejvyšší přesnosti, nejnadhnějšího sledování drážky a nejjednoduššího následného zpracování, nicméně sama o sobě je konstrukčně náročná a připadala by v úvahu až při konstrukci specializovaného kontrolního zařízení, pracujícího na optickém principu.

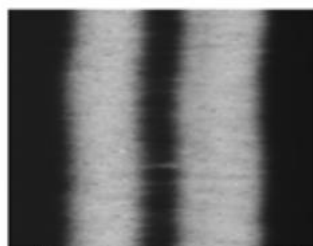
## 6.5 Osvětlení drážky

V současnosti používané mikroskopy, které jsou součástí řezací aparatury, již obsahují vlastní osvětlovací systém. Ten je zabudován přímo v mikroskopu v podobě běžné přímo žhavené žárovky, jejíž světlo se skrze polopropustné zrcadlo a vlastní optický systém soustřeďuje až na měděný nosič, kde osvětluje plochu odpovídající přibližně zornému úhlu soustavy mikroskop-kamera. Tento systém byl ovšem navržen za zcela jiným účelem, než je optické snímání desky kamerou a následné počítačové vyhodnocení – byl navržen proto, aby operátor stroje dostal co nejrealističtější obraz povrchu desky a snadněji na ní zachytil

různé vady materiálu, nečistoty, vady samotného řezu apod. Z tohoto důvodu je využita poměrně širokopásmová žárovka.

Širokopásmovost osvětlení však může přinášet i řadu nevýhod, jakými je například možnost vzniku chromatických aberací na optických prvcích (čočky, zrcadla) z důvodu rozdílného úhlu lomu pro jednotlivé vlnové délky světla. To ve výsledku může vést k neostře zobrazené hraně, přestože mikroskop byl na hranu zaostřen.

Jelikož v oblasti optického zpracování jde zejména právě o co nejlepší zobrazení hran pro jejich co možná nejsnazší a nejúčinnější detekci, je potřeba uzpůsobit tomu osvětlení a k externímu pohonnému mechanismu sestrojít i externí osvětlovací prvek. S. Stotzer v [10] navrhl pro svůj systém jako zdroj světla pro fotografii homogenní mezikruží s vnějším průměrem 84 cm (tedy více než dvojnásobek rozměru desky) tak, aby díky geometrickým vlastnostem drážky došlo k odrazu světla směrem k objektivu pouze od vodorovných ploch – tedy od povrchu desky a dna drážky. Ty se pak na výsledné negativní fotografii projeví jako tmavé plochy. Plochy šikmé (tedy stěna drážky, která v tomto konceptu nenese žádnou důležitou informaci) světlo odrážely mimo objektiv a tak byly zobrazeny jako plochy tmavé, viz obrázek 6.5, na kterém se nachází výřez jediné drážky.



Obr. 6.5: Vliv osvětlení na obraz drážky (převzato z [10])

Je tedy patrné, že volba vhodné geometrie může zásadně ovlivnit srozumitelnost a oddělitelnost hran a identifikovatelnost drážky od samotného povrchu desky. V tuto chvíli lze namítnout, že na výše uvedeném obrázku rozhodně nemůžeme mluvit o správně zaostřených hranách, nicméně je třeba vzít v úvahu, že v [10] se jedná analogovou fotografii celé desky na jeden jediný film.

### 6.5.1 Návrh osvětlovací aparatury

Tato práce si ovšem neklade za cíl osvětlit celou desku současně. Jednak by to nebylo praktické, jelikož mikroskopem chceme snímat desku po jednotlivých malých segmentech a jednak ani možné, neboť na řezací aparatuře se nachází několik dalších zařízení (řezací hlava a samotný mikroskop), které by světlu překážely. Při porovnání s [10] v této kapitole tedy navrhujeme osvětlení nikoliv

ve fázi fotografie, nýbrž ve fázi skenování, kde měl Stotzer výhodu spočívající v průsvitnosti fotografického filmu a mohl tedy svítit přímo skrze materiál z druhé strany, než kde měl umístěn mikroskop s kamerou. Ocelový plech žádnou takovou výhodu bohužel nenabízí a je třeba se spolehnout na osvětlení snímané strany a zároveň se potýkat s odrazivostí mědi, která v případě vysoce rovného povrchu, jakým galvanicky nanesená měď (povrch plechu) a měď, řezaná diamantovým hrotem (drážky) dosahuje vysokých hodnot.

Původní myšlenkou tak bylo sestrojít osvětlení ve tvaru mezikruží, které by se nacházelo okolo mikroskopu bezprostředně blízko snímané desky a tak poskytlo pokud možno rovnoměrné osvětlení snímaného povrchu. Pro test konceptu byla zvolena čtveřice LED červené barvy nad plastovým povrchem, který měl způsobit dostatečnou difuzi světla. Červená byla zvolena čistě pragmaticky proto, že jsem měl tyto LED k okamžité dispozici. Z viditelného spektra má sice nejdelší vlnovou délku a tak by mohlo hrozit nedostatečně přesné pokrytí všech detailů a hran, nicméně potřebné rozlišení se pohybuje řádově výš, než je vlnová délka tohoto světla.

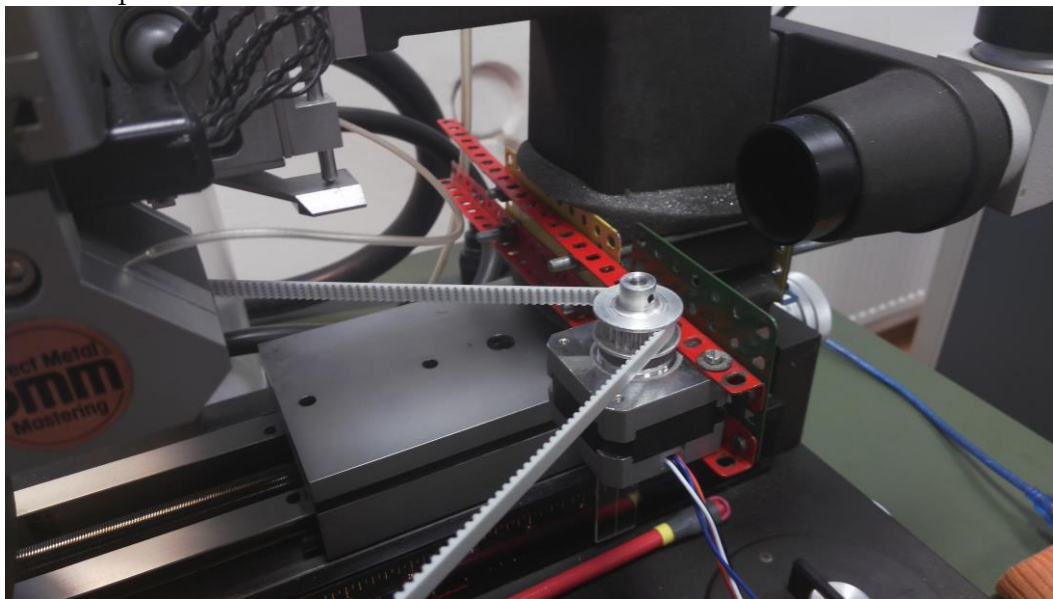
Při praktickém testu se však ukázalo, že se jednalo o volbu naprosto nevhodnou a bohužel nebylo v mých možnostech stejný systém otestovat s jinak barevnými diodami (například zelená barva, která má na Bayerově barvodělicí masce dvojnásobné zastoupení než červená a modrá, by nejspíš byla mnohem lepší volbou – práce sice pracuje s černobílými snímky, nicméně dostupná kamera byla barevná). V praktické demonstraci metody tak nakonec využívám světlo přímo z mikroskopu, neboť testovaný světelný zdroj zdaleka nedosahoval dostatečné srozumitelnosti a záznam pořízený s tímto osvětlením by byl naprosto nepoužitelný.

## Kapitola 7

# Sběr zdrojových obrazových dat

Snímání dat probíhalo nesynchronně se zbytkem zpracování. Důvodem je fakt, že na rozdíl od zpracování probíhalo přímo na pracovišti GZ Media a nebylo by žádoucí obsadit stroj na zbytečně delší dobu než je nezbytně nutné. Zároveň také proto, že zpracování dat probíhalo v Matlabu, které na daném pracovišti není k dispozici, zatímco sběr dat pomocí programu Media Express fy. BlackMagic<sup>11</sup> a interní grabovací karty DeckLink stejného výrobce, kterou jsem zase neměl k dispozici já na vlastním počítači. Možnostem, jak provádět všechny operace synchronně, se budu věnovat v závěrečné kapitole, kde bude návrh, jak by mohlo vypadat specializované pracoviště do budoucna.

Sběr dat probíhal na řezací aparatuře doplněné o externí pohon vlastního návrhu (kapitola 6.3) bezprostředně po řezání plechu s testovacím záznamem. To je důležité z důvodu, že sundání a následovně opětovné nasazení plechu typicky vyústí v odlišné umístění středu rotace a bylo by tak výrazně složitější drážku správně lokalizovat.



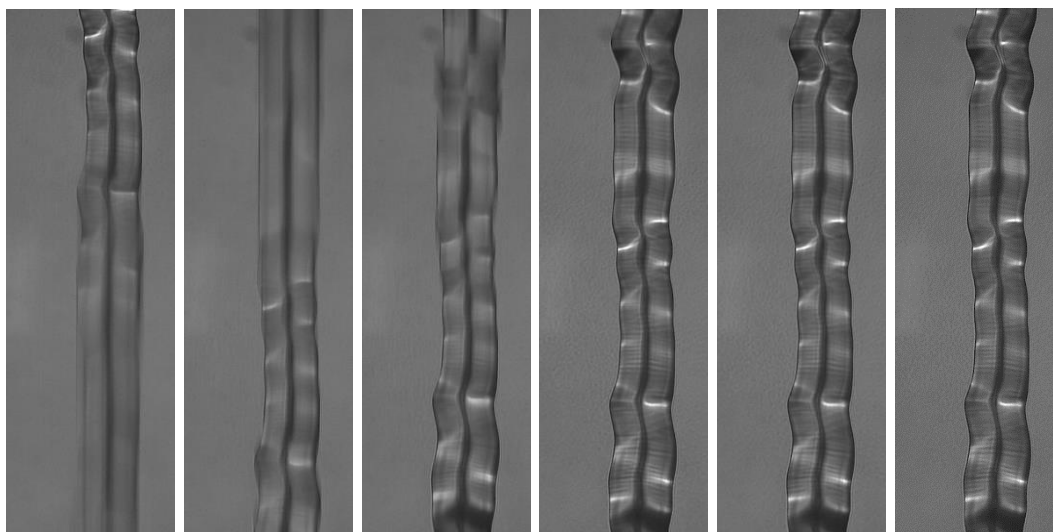
Obr. 7.1: Detail připevněného krokového motoru na řezací aparatuře

Při sběru dat jsem se snažil použitím krokového externího pohonu dle kapitol 6.2-6.3 minimalizovat pohybové rozmazání, ke kterému by došlo při použití spojitého pohonu. Jelikož ale natáčení probíhalo kontinuálně, zaznamenaný materiál obsahuje množství nadbytečných dat, která je třeba eliminovat. Záznam videa probíhal rychlostí 25 snímků za vteřinu. Znamená to tedy, že je praktické nastavit krokování motoru na násobky 40 ms v závislosti na rychlosti ustálení pozice talíře po jednotlivých krocích a na samotné rychlosti

<sup>11</sup> <https://www.blackmagicdesign.com/products/intensity/mediaexpress> (cit. 6.1.2017)

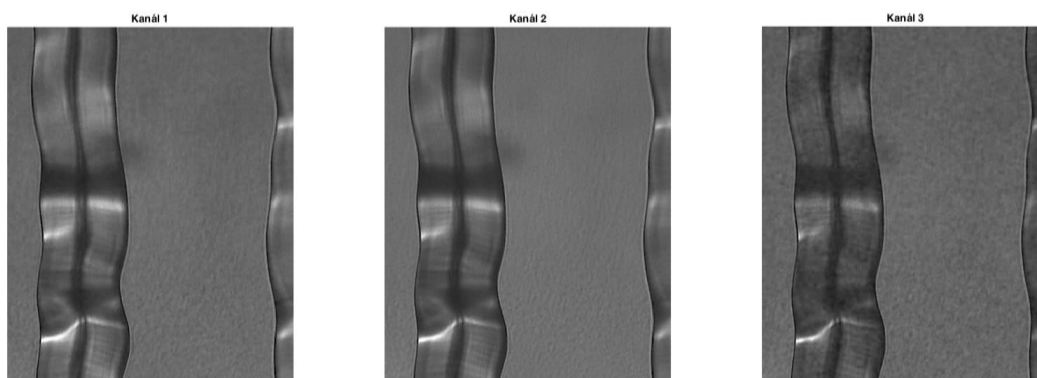


pohybu. Experimentálně jsem zkoušel jednotlivé násobky a při hodnotě zpoždění o 240 ms mezi jednotlivými kroky jsem zjistil, že vždy 3-4 snímky jsou rozmazané pohybem a zbývající 2-3 jsou ostré v celé ploše.<sup>12</sup>



Obr. 7.2: Postupné střídání ostrých a pohybem rozmazaných snímků

Přestože by bylo možné jednotlivé snímky videa nechat ve formě vícerozměrné struktury, dal jsem přednost jejich uložení do BMP obrazových souborů. Jelikož bylo naprosto zbytečné zachovat všechny snímky a všechny jejich RGB kanály, zachycením několika snímků a jejich manuální kontrolou jsem vybral, kolikátý z každé šestice snímků je nejostřejší a zároveň který z kanálů poskytuje pokud možno největší kontrast hrany, pro co možná nejsnazší budoucí vyhodnocení. Vzhledem k nejmenšímu množství šumu jsem zvolil kanál 2 (tedy zelený).



Obr. 7.3: Porovnání jednotlivých kanálů barevného snímku

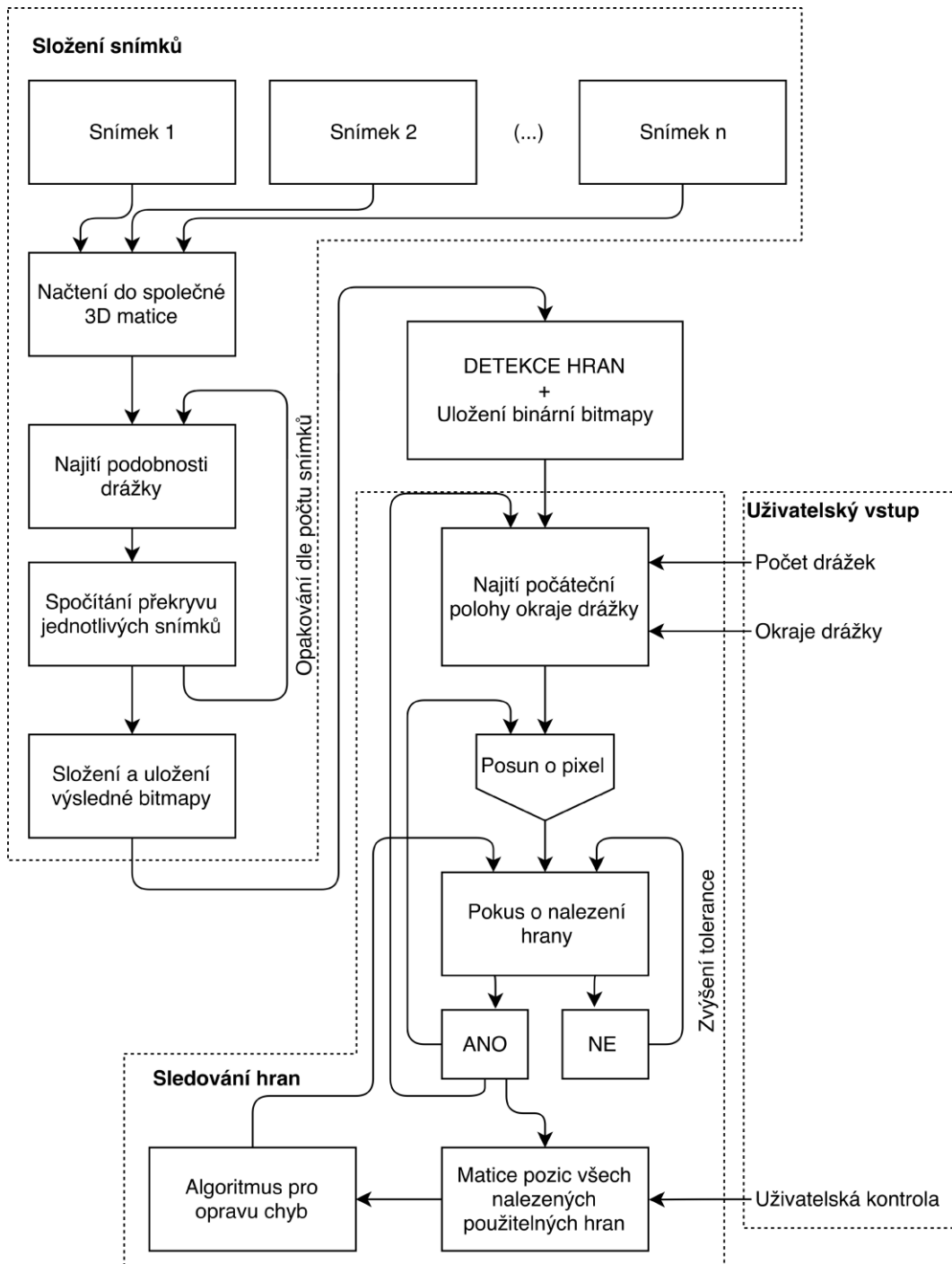
---

<sup>12</sup> Zároveň lze na nasnímaných záběrech snadno demonstrovat vliv tzv. „rolling shutter“ u CMOS snímačů.

## Kapitola 8

# Detekce pozice drážky z obrazu

Tato kapitola je věnovaná přechodu od sady černobílých snímků až po několik málo (konkrétně 14) vektorů, obsahujících informace o pozici okrajů jednotlivých drážek. Celý systém se řídí algoritmem uvedeným na následujícím schématu 8.1.

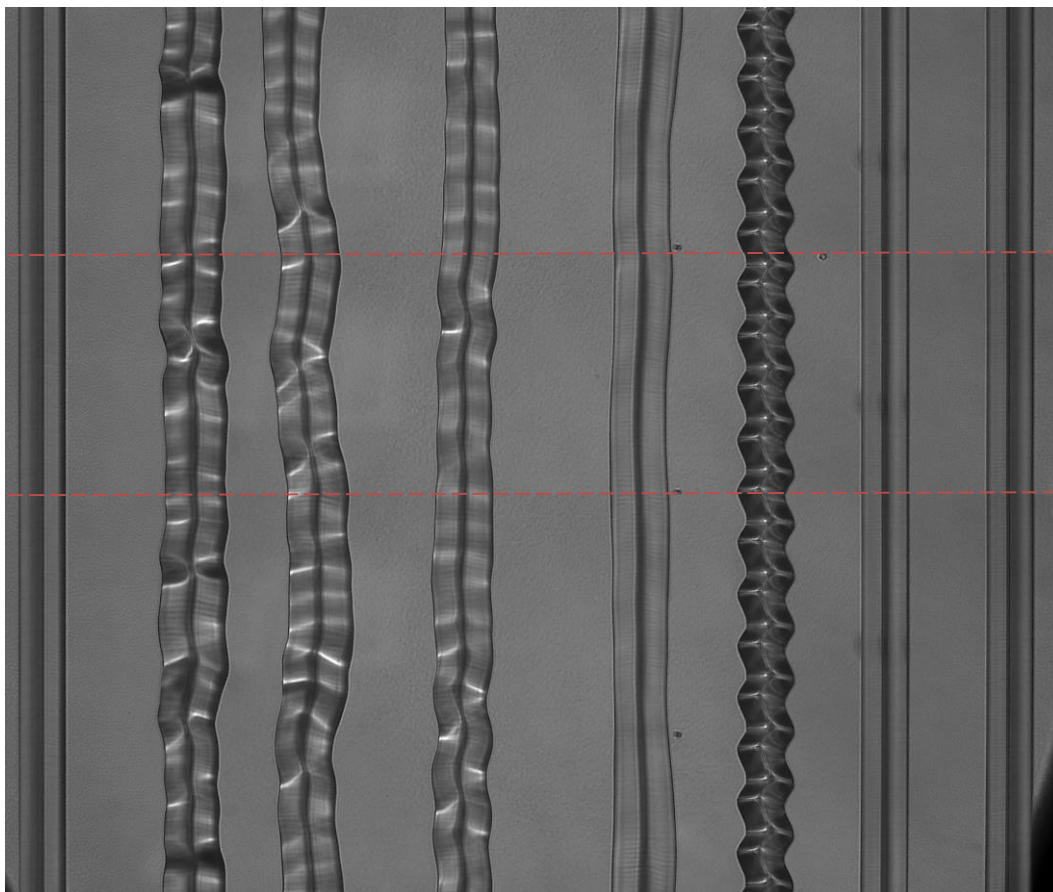


Obr. 8.1: Algoritmus spojení snímků a detekce hran

## 8.1 Určení překryvu snímků

Jelikož frekvence pořízených snímků a jejich překryv není dokonale konstatní, ale liší se vždy v rozmezí pixelů, je třeba sestrojít algoritmus na přesnou detekci překryvu snímků, aby bylo možné z množství malých snímků vytvořit snímek delší, v rámci kterého budeme detekovat drážku.

Přístupů je v principu několik: Skládat přímo obrázky a následně na větším snímku provést převod na zvuk, ten vygenerovat a následně skládat dohromady menší množství delších zvukových intervalů. Nevýhodou je nutnost delších operací s 2D vektory (maticemi), které jsou výpočetně náročnější než operace s vektory klasickými (což je již zmíněný zvuk). Druhá možnost je zpracovat snímky zcela jednotlivě, převést na zvukovou informaci a teprve následně tyto zvukové vektory skládat dál za sebe, což je operace, která bude muset být tak jako tak vykonána. V případě automatizovaného měření, které bude navrženo jako možný směr vývoje do budoucna, by snímání probíhalo řízeně tak, že by byl pořízen snímek teprve ve chvíli, kdy byl předchozí snímek zpracován. Výhodou by bylo, že by odpadla nutnost skladování většího množství dat, protože by byla okamžitě zredukována na zvuková data nebo alespoň data převeditelná na zvuk. Bavíme se zde tedy pouze o datech nalezené polohy drážky, což jsou jediná data, která nás v tuto chvíli zajímají. Ukládání obrazových dat s sebou ovšem přináší spoustu naprosto nepotřebných informací, jako je například šum kamery na homogenních plochách měděného plechu v místech bez drážky (land) či stínování samotné drážky. Tento šum a další proměnné artefakty pak jednak zabírají zbytečně moc místa na datovém úložišti a navíc díky nim hrozí, že detekční algoritmy nedokáží jednoznačně správně určit polohu drážky. Ostatně během zpracování snímků se několikrát skutečně stalo, že program mylně sledoval nějaký odlesk v drážce či okraj zrnka prachu, namísto skutečné hrany.



Obr. 8.2: Spojený obrázek s naznačením místa spojů („švy“)

Nakonec jsem se s ohledem na určitou míru ruční práce při zpracování dat rozhodl obvod desky rozdělit na 8 úseků, které jsou zpracovány postupně zvlášť a následně složeny dohromady. Každý z těchto 8 úseků obsahuje jako vstup cca 500 obrázků ve FullHD rozlišení (1920x1080 obrazových bodů).

Jelikož kamera byla seřízena tak, aby k pohybu obrazu docházelo pouze po jedné ose, jako detektor byla zvolena metoda srovnání úseků obrazu pomocí kritéria MSE<sup>13</sup> (mean square error). První z dvojice snímků byl zvolen jako referenční a z tohoto snímku byl pevně určen řádek, o kterém se dá předpokládat, že na snímku následujícím nalezneme jeho ekvivalent. Na snímku druhém byl v cyklu vybírán vždy konkrétní řádek, který byl pomocí MSE srovnáván s řádkem referenčním. Tam, kde byla podobnost nejvyšší, bylo stanoveno, že se jedná o odpovídající úseky desky a snímky na sebe byly napojeny. V cyklu se tak dohromady složilo všech cca 500 snímků. Pro zamezení rizika, že daný řádek sice bude shodný, ale bude se jednat o pouhou náhodu je pomocí MSE vyhodnocen nejen konkrétní řádek, ale i jeho blízké okolí (konkrétně 40 pixelů) na každou stranu, díky čemuž už lze téměř s jistotou tvrdit, že se jedná o totožný úsek desky, nejen nějakou náhodnou podobnost drážky.

---

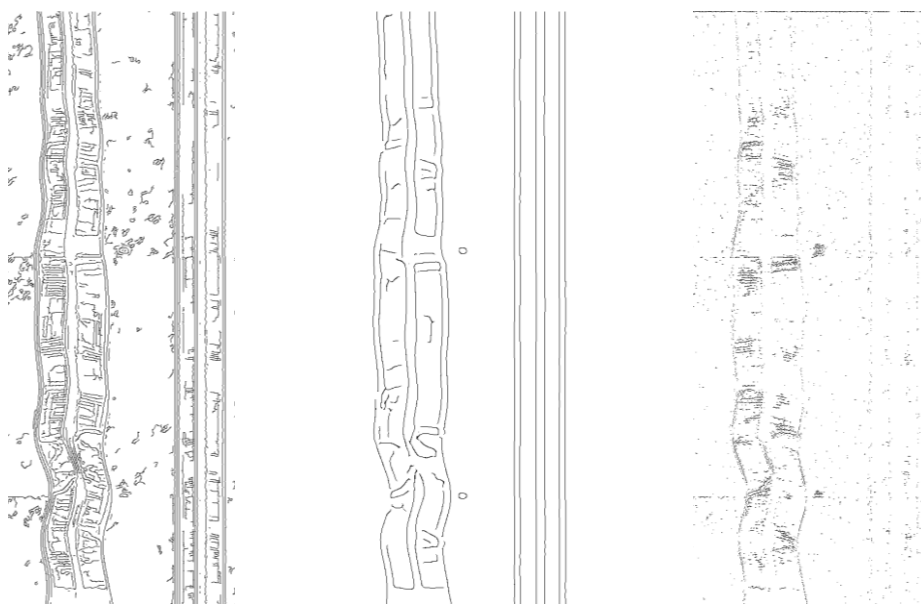
<sup>13</sup> MSE je střední kvadratická chyba. V tomto konkrétním případě se tak jedná o průměrný rozdíl hodnoty jasu odpovídajících pixelů. Tam kde je tento rozdíl nejmenší jsme ochotni tvrdit, že se jedná o nejpodobnější pás obrázku.

Výstupem této části programu bylo 8 snímků o různých rozlišeních v rozmezí od  $1920 \times 85$  418 pixelů až po  $1920 \times 247$  071 pixelů.

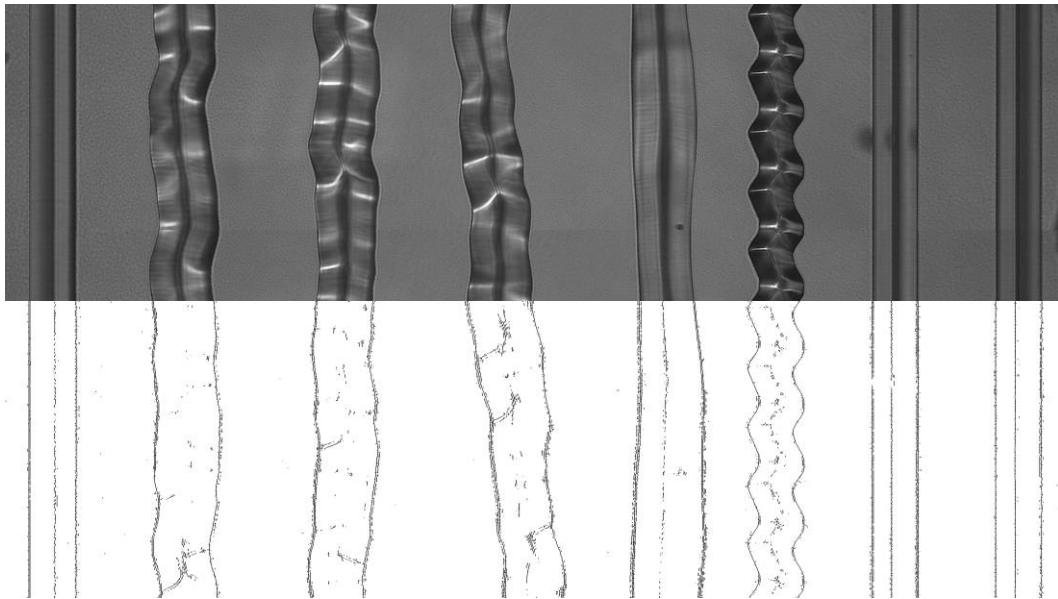
Tento rozdíl byl způsoben tím, že během záznamu došlo po 13 minutách k vypnutí osvětlení mikroskopu díky pojistce proti přehřátí. Záznam je tak tvořen dvěma částmi. Původně jsem chtěl obě videa spojit do jednoho a pracovat s celým videem v kuse, nicméně se během zpracování ukázalo, že v místě, kde by dle tvaru drážky videa měla navazovat, jsou posunuta o cca 15 pixelů horizontálně, a proto jsem se rozhodl zpracovat je zvlášť. Rozdílná délka spojených videí je proto způsobena tím, že nejkratší z nich obsahuje závěr prvního videa a nejdelší pak celé video druhé.

## 8.2 Detekce polohy drážky

Pro detekci drážek byla na základě teoretického rozboru zvolena metoda detekce hran. Porovnáno bylo několik algoritmů a následně jako nejvhodnější (tzn., že nejlépe zvýrazní drážku a nejvíce potlačí okolí) zvolena metoda Prewittové ve vertikálním směru. Ukázka jiných metod je na obrázku 8.3 a srovnání vybrané metody s originálním snímkem pak 8.4. Pro účely práce jsou snímky upraveny na negativ.



Obr. 8.3: Ukázka jiných detektorů hran. (zleva: Cannyho metoda v defaultním nastavení, Cannyho metoda se zvýšeným prahováním, metoda Prewittové v horizontálním směru)



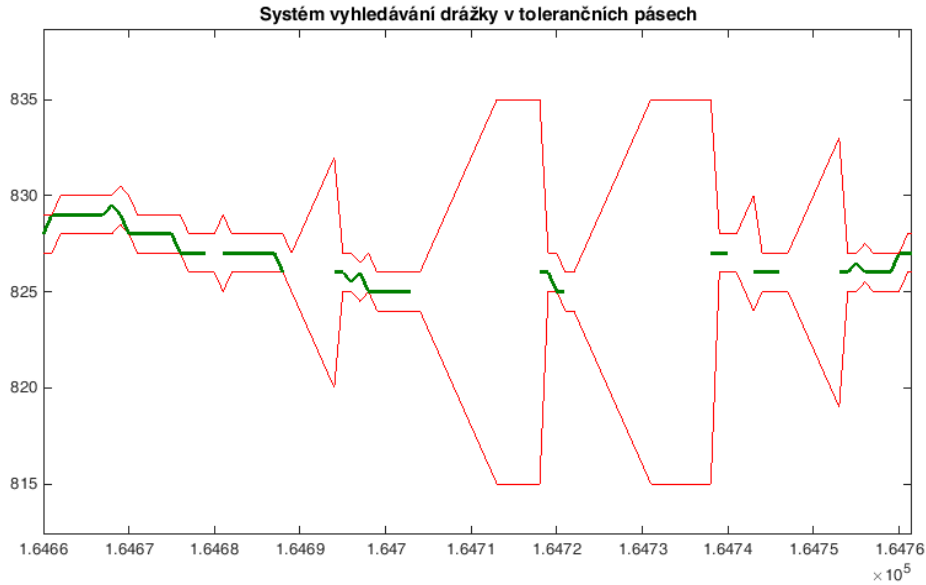
Obr. 8.4: Srovnání bitmapy v ostínech šedi a výsledné detekované hrany (v negativu)

Po detekci hran je potřeba využít spolupráci uživatele, který na zobrazeném obrázku postupně musí kliknout na hrany, které jsou skutečně hranami drážky, aby se zamezilo tomu, že program bude sledovat například dno drážky či nějaký výraznější vzorec šumu. Je tedy potřeba mu pomocí funkce *ginput()* „vnutit“, kde má začít se sledováním drážky.

Navržený algoritmus sledování drážky počítá se skutečností, že maximální úhel drážky bývá  $45^\circ$  a proto v prvním řádku považuje za hranu přesné souřadnice vybrané uživatelem a následně v dalších řádcích vyhledává detekovanou hranu (tedy hodnotu 1 v matici formátu logic) na téže souřadnici. Pokud nenajde (=drážka není svislá), přidá ještě hledání o 1 pixel vlevo a vpravo. Výsledná souřadnice drážky je spočítána jako průměr všech nalezených „jedniček“.

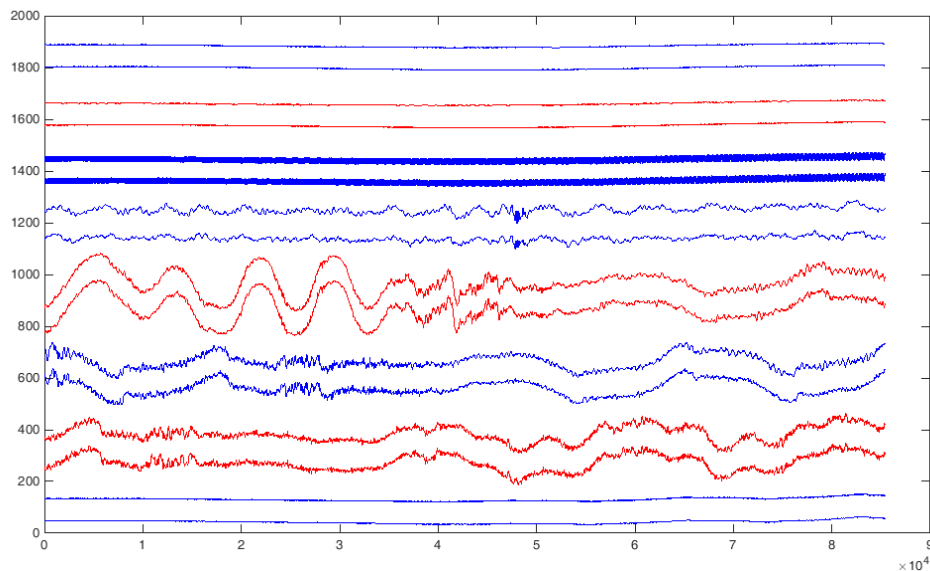
### 8.2.1 Virtuální sledování chybějící hrany

Jelikož detekce hran ani snímání není úplně ideální, program obsahuje algoritmus zpracování chybějícího kusu drážky, kde podle zmíněného kritéria pokračuje s hledáním drážky i v případě, že v daném řádku žádnou nenajde. Zde pak s každým pixelem v jedné ose zvětší oblast vyhledávání až do maximální hodnoty nastavené na 10 pixelů – tato mez je stanovena proto, aby se minimalizovalo riziko, že v tomto kritickém místě program začne sledovat jinou drážku než je žádoucí. Princip tohoto sledování i v případě nenalezené drážky je naznačen na obrázku 8.5, kde je zeleně označena správně nalezená drážka v místech výpadku a červeně pak aktuální toleranční pásmo v rámci kterého je drážka dále hledána.



Obr. 8.5: Princip vyhledávání drážky v tolerančních pásech

Pokud nenastanou žádné chyby ve sledování či například spojení svou sousedních drážek, chybné sledování šumu, apod., výstupem sledovacího algoritmu je matice obsahující polohu všech nalezených hran ve všech pixelech. V praxi se tedy jedná o 14-16 vektorů v závislosti na množství drážek v obraze. Grafickým znázorněním bezchybného výstupu tak může být například obrázek 8.6.



Obr. 8.6: Grafické znázornění detekovaného průběhu drážky. Kromě rozdílného měřítka v horizontální a vertikální ose je analogické k bitmapovému obrazu drážky

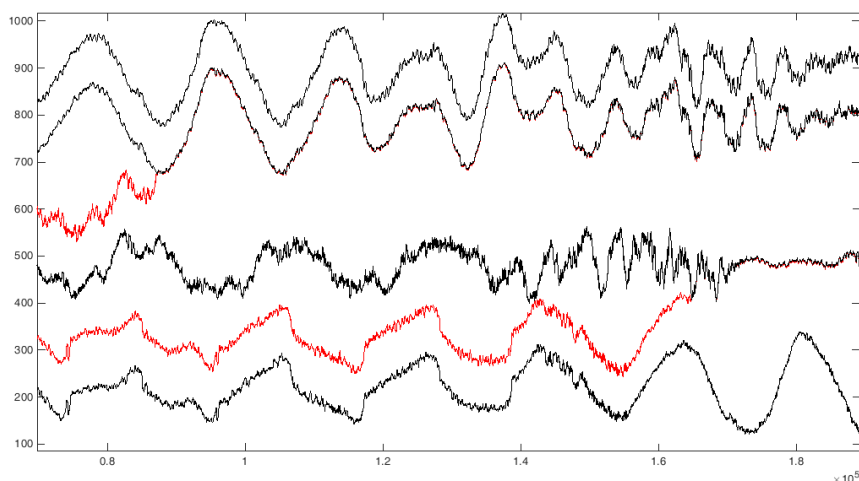
## 8.2.2 Detekce dotyků, prořezů a oprava kritických míst

Často se však může stát, že je již z grafického výstupu zřejmé, že se na desce vyskytlo místo, které automatický detektor nedokáže bezchybně projít bez

uživatelské asistence. Typicky se jedná o dotyk sousedních drážek, kdy existuje řádek, v rámci kterého získají dvě sousední drážky totožnou číselnou hodnotu. Jelikož program vyhledává další hranu v definovaném okolí této hodnoty, je zřejmé, že od tohoto řádku dále bude vyhledávání fungovat pro obě drážky shodně<sup>14</sup>.

Pro tyto případy je tedy v programu implementován opravný algoritmus, který uživateli umožní najít a opravit kritická místa (i za cenu ztráty detailů v tomto místě). Systém tak vlastně do jisté míry nepřímo, prostřednictvím neschopnosti daný snímek zpracovat automatizovaně, získává schopnost nalézt místa, kde se nachází nějaký problém. Jak je naznačeno v diagramu 8.1, popisujícím postupný proces zpracování dat, je tato část programu zařazena na úplný závěr vyhledávání drážky a má schopnost zopakovat příslušnou část procesu vyhledávání, pokud v jeho rámci nastaly nějaké problémy.

Funguje tak, že poté, co je zobrazen finální graf polohy hran drážek (viz obrázek 7.5), uživatel napíše indexy hran, o kterých se domnívá, že nejsou v pořádku. Například na ukázce níže (obrázek 7.6) je zřejmé, že došlo k chybné detekci drážek s indexy 2 a 4 (odspoda)



Obr. 8.7: Ukázka chybně detekovaných – spojených – drážek

Uživatel proto v tuto chvíli do programu vepíše, že si přeje opravit hrany 2 a 4 a spustí příslušnou část programu, kde uživatel opět kurzorem pomocí funkce *ginput()* klikne do místa, od kterého došlo k problému s drážkou (pokud je chybných drážek více, postupuje se podle indexu odspoda nahoru). Následně je zobrazen výřez drážky s chybou (část obrázku 8.8) tak, aby uživatel mohl chybu odstranit klepnutím na její začátek a konec. Hodnoty mezi těmito body se

<sup>14</sup> Pozn: To nakonec není úplně pravda, algoritmus má implementované podmíněčné zaokrouhlování tak, že vždy u levé a pravé drážky zaokrouhluje vypočtené indexy nahoru, respektive dolů právě proto, aby se v případě, že dotyk drážek není úplný, snažil drážky od sebe oddělit. V testu se však ukázalo, že toto opatření na opravu chyb nemá prakticky žádný vliv.



dopočítají pomocí interpolace, detekce od tohoto místa postupuje dále a uživateli se zobrazí nový grafický průběh. Postup se opakuje vícekrát, pokud je to nutné.



Obr. 8.8: Ukázka detekovaných průřezů

## Kapitola 9

# Tvorba zvukového souboru

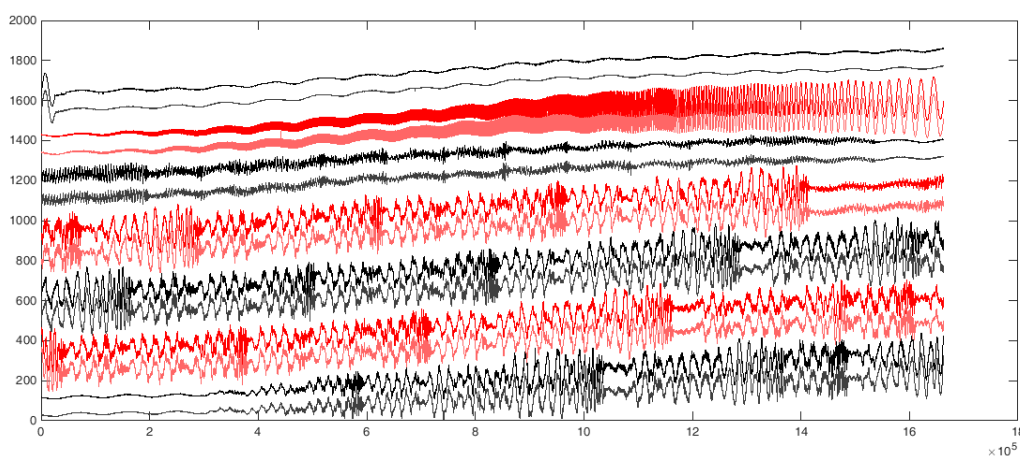
Tato kapitola obsahuje popis procesů, úprav a výpočtů, pomocí kterých bylo dosaženo toho, že z matice reprezentující polohu všech drážek na desce vznikl přehratelný audiosignál.

## 9.1 Převod množiny matic na dvojici vektorů

Z předchozí kapitoly je zřejmé, že do tohoto kroku zpracování vstupují data v podobě osmi matic, které lze graficky znázornit způsobem, jaký je ukázán na obrázku 8.6. Každá z těchto matic ale obsahuje informace o 7-8 drážkách, které na sebe nenavazují. Proto je třeba spojení signálu provést postupně ve 2 krocích:

1. Najít podobnost celých matic a spojit je za sebe
2. Rozdělit výslednou matici na jednotlivé drážky a ty opět správně spojit

První úloha je poměrně jednoduchá a jedná se o analogii úlohy z kapitoly 8.1, proto se jí už nebudu detailněji zabývat, snad jen s tím rozdílem, že nyní se nejedná o grafická obrazová data, nýbrž souřadnice polohy drážky, a díky tomu je kritérium MSE mnohem přesnější a proto nám stačí vyšetřovat mnohem menší okolí (použil jsem rozmezí 4 pixelů na obě strany). Druhou výhodou pak je skutečnost, že šířka matice je pouze 14 či 16, místo předchozích 1920, takže je výpočet značně rychlejší. Grafické znázornění výstupu prvního kroku je na obrázku 9.1.



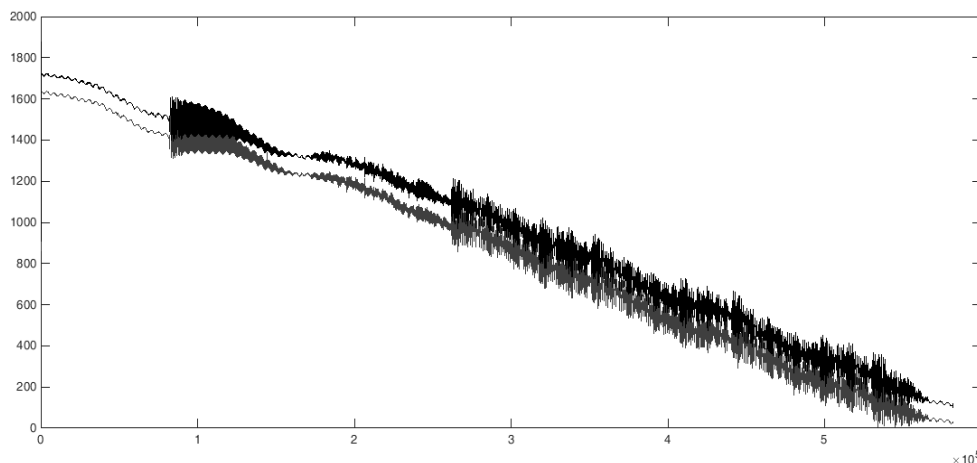
Obr. 9.1: Grafické znázornění složené informace o všech 8 drážkách

Z výše uvedeného znázornění je patrných hned několik skutečností:

- Díky první, tiché, drážce jsme schopni říct, že ke spojení snímků došlo správně.
- Celý signál je modulován nízkofrekvenčním signálem – Na jednu otáčku jsme schopni napočítat cca 20 period, jedná se tedy o frekvenci  $f_p = \frac{20}{1,8} \cong 11,1 \text{ Hz}$ . Nachází se tedy bezpečně mimo slyšitelné pásmo. Spekulace nad příčinou této parazitní frekvence se nachází v kapitole 9.3.1.
- Na rozdíl od jednotlivých snímků už zde začíná být velice patrné, že se jedná o spirálové, drážky a že tedy výsledek bude nutné nějakým způsobem korigovat

Druhá úloha je v principu ještě jednodušší než první, protože délka všech vektorů je již upravená tak, aby na sebe navazovaly. Bohužel se ukázalo, že mikroskop, který je umístěný na hladké tyči, pro co nejmenší tření, měl tendenci v době natáčení záznamu změnit svoji polohu o přibližně 15 pixelů a díky tomu spojené snímky vykazují skokovou změnu, která by mohla vyústit v praskot a lupance v místě spojů (tedy periodicky každých 1,8 vteřin). Bylo tak tedy třeba vektory napojovat metodou, kdy se vždy spočítal rozdíl mezi posledním pixelem prvního vektoru a prvním pixelem druhého vektoru a o jejich rozdíl pak celý spoj kompenzovat.

Jelikož mikroskop snímá odspodu, zatímco indexování veškerých snímků probíhá odshora, je třeba ještě celý vektor následně zrcadlit (funkcí *flipud()*). Výsledná reprezentace je na snímku 9.2 níže:



Obr. 9.2: Kompletně spojený vektor reprezentující 7 drážek z 8 sérií snímků

## 9.2 Další zpracování zvuku

Nyní jsme tedy ve fázi, že známe informace o kompletní relativní poloze hran drážky v obraze. Zde už je velmi patrný vliv spirálovitého tvaru drážky. Intuitivním řešením by nyní bylo signálem proložit přímkou, spočítat její parametry a následně pomocí ní signál korigovat („srovnat“). Jelikož je však amplituda zvukového signálu úměrná rychlosti nože, nikoliv přímo mechanické výchylce, je třeba získanou informaci o vzdálenosti převést na informaci o rychlosti pomocí první derivace, jelikož víme, že rychlost je derivace vzdálenosti podle času. Díky tomu, že se nacházíme v oblasti diskretních signálů, bude se tak jednat o numerickou diferenci mezi jednotlivými vzorky. Spirálovitost drážky proto není třeba kompenzovat, jelikož je při diferenci vliv nejnižších frekvencí (tedy i celkového zakřivení) eliminován.

Dále je potřeba signál ekvalizovat podle inverzní RIAA charakteristiky a vhodně převzorkovat. Vzhledem ke způsobu zpracování dat, čtení souřadnic hran apod, je zároveň patrná nutnost signál vyhladit.

### 9.2.1 Interpolace chybějících vzorků

Informace o poloze drážky se během sledování hran ukládaly současně do dvou matic. Tato metoda byla zvolena kvůli možnosti sledování i chybějící drážky dle kapitoly 8.2.1, proto, že se algoritmus odkazoval vždy na předchozí polohu a pokud by žádná nebyla uložena, nedokázal by nadále pokračovat. Bylo by samozřejmě možné použít tyto pomocné matice, ale jelikož obsahují skokové změny hodnot, v důsledku by se v signálu projevil praskot, lupance, apod. (jelikož difference skokové hodnoty (tedy směrnice spojnice dvou bodů) je rovna výšce skoku) a bylo by třeba se s těmito nadále potýkat. Z tohoto důvodu jsou použita data, obsahující poměrně velké procento vynechaných hodnot a do spojitosti jsou následně interpolována pomocí kubické interpolace (metoda *pchip*).

### 9.2.2 Převzorkování

Vektory dat obsahují informace o poloze drážky v jednotlivých pixelech a je tedy jejich vzorkování výrazně vyšší než je pro audio nutné a standardní. Jelikož neznáme přesný údaj o ekvivalentní velikosti pixelů v  $\mu\text{m}$ , ze které bychom následně byli schopni vzorkování spočítat, bude nejněsnější vyjít ze skutečnosti, že testovací data vyústila ve vektory o délce 11 640 328 a obsahují záznam 7 otoček gramofonové desky, jejíž doba rotace při použitím standardu je

rovna 1,8 vteřiny. Výsledný poměr převzorkování na CD formát, tedy 44100 Hz, poté spočítáme dle rovnice (9.1).

$$Q = \frac{n_1}{n_2} = \frac{n_1}{n_0 t_0 f_s} = \frac{11640328}{7 \times 1,8 \times 44100} = 20,9487 [-] \quad (9.1)$$

Výhoda převzorkování přímo v Matlabu spočívá v tom, že použitá funkce v sobě obsahuje antialiasingový filtr (dolní propust), není proto potřeba před samotným převzorkováním provádět žádné vyhlazení dat. Naopak dodatečné vyhlazení až po převzorkování by mělo za důsledek ztrátu detailů ve vyšších kmitočtech.

### 9.2.3 Frekvenční úpravy a finalizace

Posledními úpravami je pak aplikace diference, kdy z informace o poloze hrany získáme informaci o rychlosti záznamového nože či přenosky při přehrávání. V této fázi tedy přehráním vektoru získáme frekvenčně obdobný signál, jako kdybychom skladbu přehráli přímo na gramofonu bez patřičného předzesilovače. Z tohoto důvodu je třeba ještě aplikovat inverzní RIAA charakteristiku. V Matlabu se jedná o funkci *diff()*, která spočítá směrnici spojnic sousedních vzorků vektoru a jelikož se jedná o ekvidistantně rozmístěné hodnoty (na ose x tedy), je výstupem vektor rozdílů vždy sousedních dvou vzorků – výsledný vektor je tedy o 1 vzorek kratší. Druhou použitou funkcí je *filter()*, do které jako parametry vstupuje vektor dat a vektory koeficientů, převzatých z [3]. Výměnou pořadí koeficientů *a* a *b* pak určíme, jestli se bude jednat o filtraci záznamovou či přehrávací. Délka vektoru se v tomto případě nemění.

Jediným problémem u funkce *resample()* bývají okrajové efekty, neboť funkce předpokládá, že hodnoty před vektorem a po něm jsou nulové, což má za vliv kmitání na začátku a na konci signálu – v tomto případě poměrně výrazné, neboť se přesamploval signál o relativně vysokých hodnotách. Jedním z řešení je ošetřit tyto okraje doplněním dostatečného počtu hodnot, odpovídajících průběhu funkce, druhým je pak (po diferenci) problémové vzorky na začátku a na konci vektoru vynulovat, neboť v tomto místě nepřijdeme o žádnou podstatnou informaci.

Závěrečným krokem je normalizace signálu na maximální úroveň 0,99 a export zvukového wavu.

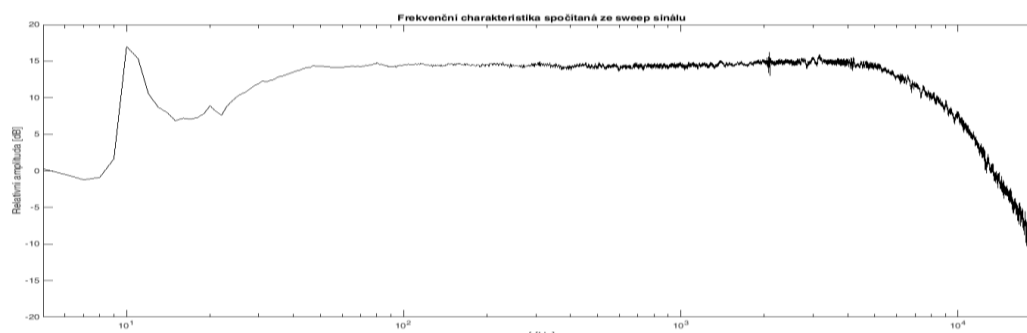
## 9.3 Vlastnosti zvukového výstupu

S vyhodnocením výstupu systému může být trochu problém, neboť nebyly použity exaktní měřicí signály nýbrž signál hudební – tedy úsek písničky. Pro účely vyhodnocení signál sice začíná (kromě tiché drážky) přeladovaným

harmonickým signálem pro alespoň hrubou možnost určení impulzní odezvy, nicméně i ten byl v průběhu přípravy před řezáním frekvenčně upraven, neboť by zejména v části s vyššími kmitočty došlo k vysokému úhlovému zkreslení a proudy do řezací hlavy by byly výrazně vyšší, než je její limit. Jedná se proto o charakteristiky nikoliv samotné rekonstrukce, nýbrž celého řetězce masteringu, řezání a následného snímání.

### 9.3.1 Spektrální vlastnosti

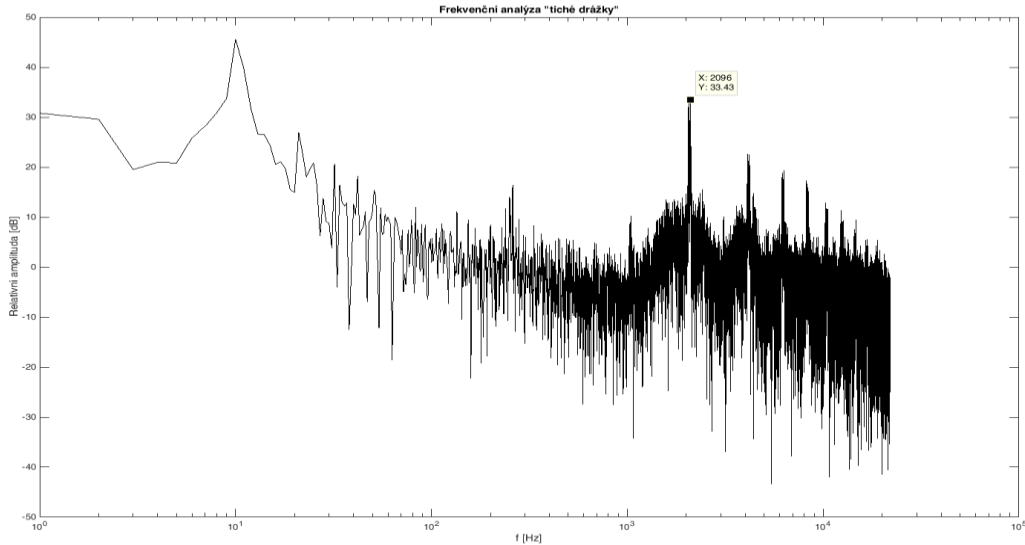
Přesto je z frekvenční charakteristiky na obrázku 9.3 patrné několik důležitých informací.



Obr. 9.3: Frekvenční charakteristika spočítaná ze sweep signálu

První zřetelná je parazitní frekvence okolo 10 Hz. Jedná se o frekvenci mimo slyšitelné pásmo. Pravděpodobně se dá říct, že je způsobená externím pohonem, neboť po přepočtu perioda tohoto parazitního signálu odpovídá jedné otáčce krokového motoru (10 Hz znamená, že do jedné otáčky desky se vejde 18 period, což odpovídá převodní konstantě pohonu). Je tedy možné, že tento problém je způsoben tím, že primární řemenice pohonu nemá dokonale vystředěný upínací otvor a díky tomu v periodických intervalech mění rychlost posunu mechanického talíře a tyto síly následně ovlivňují i polohu mikroskopu.

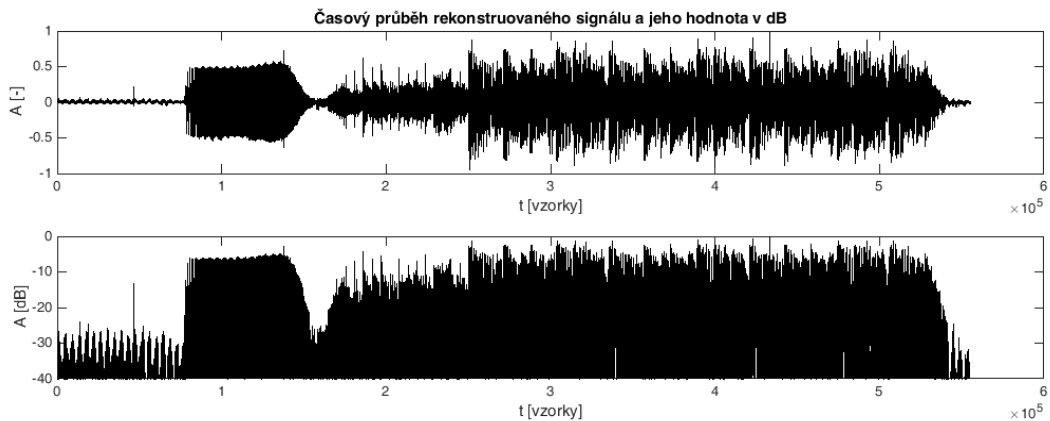
Druhá parazitní frekvence není tak patrná z celkové analýzy sweepu, ale o to patrnější je při počítání spektra z tiché drážky (obrázek 9.4).



Obr. 9.4: Frekvenční charakteristika spočítaná z tiché drážky

Jedná se o frekvence v rozsahu cca 2060 až 2100 Hz a její vyšší harmonické frekvence. Zde už se jedná o problém, neboť se jedná o frekvence slyšitelné. Jejich zdrojem pravděpodobně může být spojování snímků, neboť na jednu otáčku připadá 3746 snímků (tedy po vydělení 1,8 vteřinami je ekvivalentní snímková frekvence rovna 2081 Hz). Tato frekvence nebyla konstantní, měnila se periodicky (s ekvivalentní frekvencí 10 Hz – viz výše), proto je i tato parazitní frekvence rozptřena do okolí.

Dalším důležitým parametrem je odstup signálu od šumu, respektive úroveň šumu samotného. Pro to je opět možné snadno využít zobrazení tichého úseku drážky (9.5) – vidíme, že v tichých úsecích šum dosahuje hodnoty okolo -30dB.



Obr. 9.5: Rekonstruovaný signál v absolutních hodnotách a v dB stupnici

### 9.3.2 Zkreslení zvuku

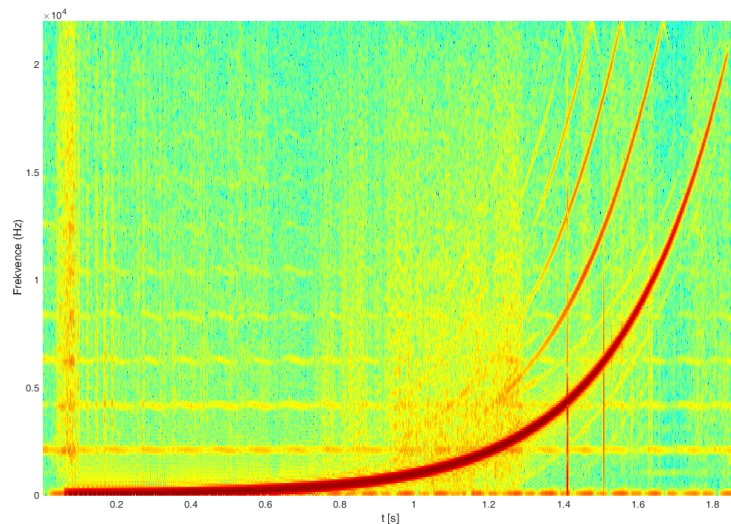
Jednou z dalších kvantifikovatelných vlastností zvukových systémů je zkreslení. To se dělí na zkreslení harmonické a ostatní. Harmonické zkreslení je dané jako

poměr výkonu vyšších harmonických frekvencí vůči frekvenci základní, neboli fundamentální [18]. Vypočítat ho lze podle vzorce (9.2).

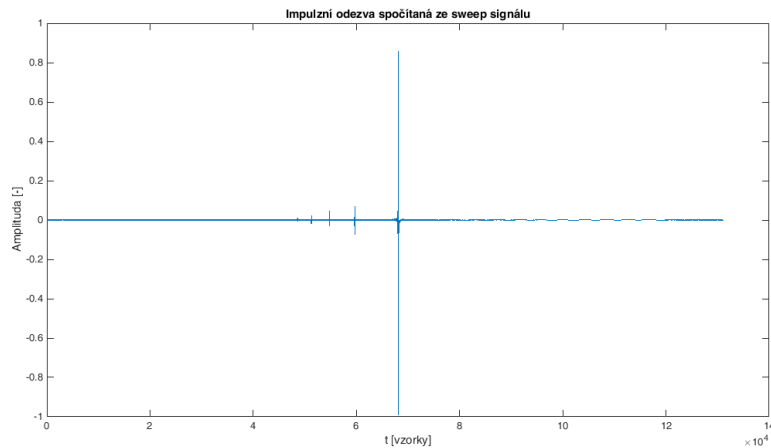
$$THD_n = \frac{\sum_{k=2}^n P_k}{P_1} \cdot 100[\%] = \frac{P_2 + P_3 + P_4 + \dots + P_n}{P_1} \cdot 100[\%] \quad (9.2)$$

Ze spektrogramu rekonstruovaného sweep signálu za použití Kaiserova okna délky 256 a 512 body DFT s překryvem 220 bodů (obrázek 9.6) jsou tyto vyšší harmonické frekvence jasně zřetelné. Zřetelné jsou i na grafu 9.10, zobrazujícím absolutní hodnotu impulzní odezvy v závislosti na čase, kde jsou zřetelné v podobě menších špiček před hlavní špičkou, reprezentující základní signál [19].

Pokud umocněním hodnoty špiček převedeme na výkon a dosadíme do výše uvedeného vzorce, vyjde harmonické zkreslení pro 4 harmonické rovné hodnotě 0.82%, což je zdánlivě výsledek velmi dobrý, nicméně vzhledem k parazitním rušivým signálům není až tak důležitý



Obr. 9.6: Spektrogram rekonstruovaného signálu



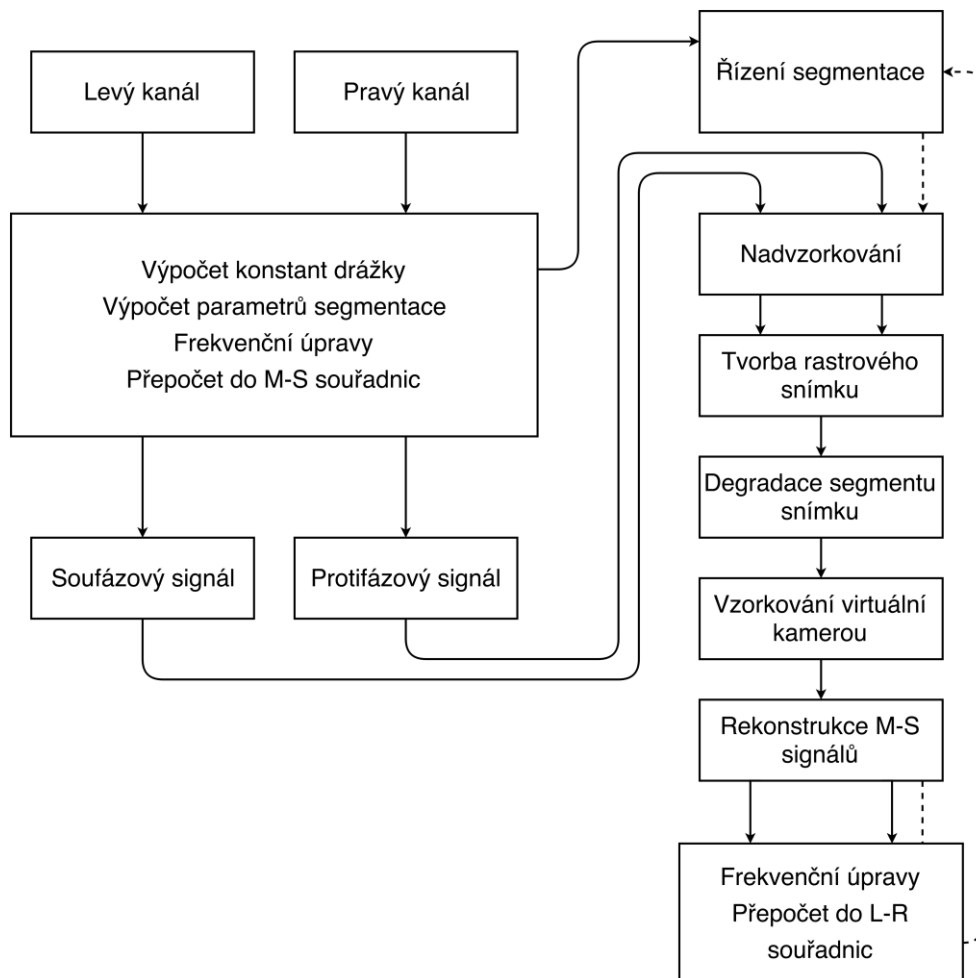
Obr. 9.7: Impulzní odezva s patrnými artefakty harmonického zkreslení



## Kapitola 10

# Simulace snímání řádkovou kamerou

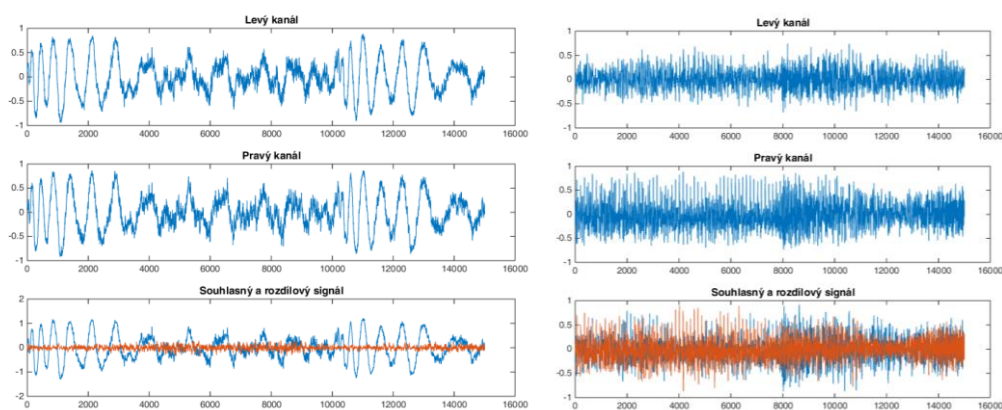
Poslední částí práce je simulace jak by probíhalo snímání obrazu za pomoci řádkové kamery s pevně danými parametry. Tato kapitola již bude stručnější a bude se jednat spíše o naznačení, jak by takové snímání mohlo probíhat v případě, že by byly navrženy dobré podmínky včetně osvětlení apod., jejichž reálný návrh už je mimo téma a rozsah práce. Dá se předpokládat, že praktické řešení by těmito předpokladům bylo i velice vzdálené. Vzorek drážky pro tuto část bude na rozdíl od části předchozí získán simulací a to tak, že ze zvukového souboru nejprve pomocí stanovených teoretických pravidel vytvoříme model drážky, jak by se asi mohla pod mikroskopem jevit, tento model bude následně upraven do podoby, v jaké by ho mohl vidět řádkový snímač (pohybové rozmazání, šum, apod...) a následně pomocí virtuální kamery převeden zpět do zvukové podoby. Blokové schéma postupu je naznačeno na obrázku 10.1.



Obr. 10.1: Schéma simulace snímání řádkovou kamerou

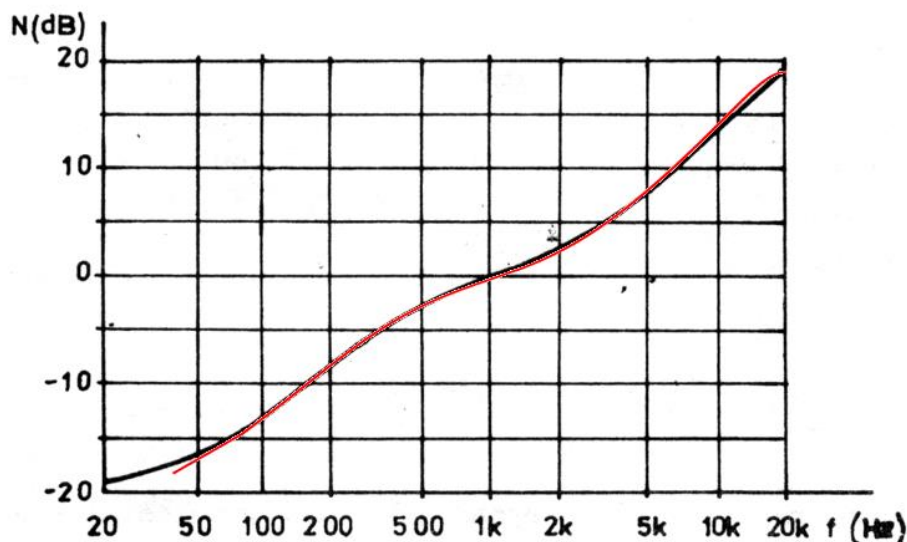
## 10.1 Frekvenční úprava audiostop

Jelikož se jedná o kompletní simulaci, je třeba nejprve vytvořit model drážky. K tomu je zapotřebí nejprve převést stereosignál na soufázovou a protifázovou složku (neboli monofonní a stereofonní složky signálu). K tomu lze snadno použít vzorce (2.3-2.6) z kapitoly 2.3.2. Již teď lze dobře z obrázku 10.2 vidět, že různé skladby mají různý poměr těchto složek, a je možné odhadnout, jak by se asi chovaly při řezání reálným záznamovým strojem. Za zmínku ještě stojí, že úseky obou skladeb mají časovou osu ve shodném měřítku, je tedy vidět, že úsek druhé skladby obsahuje výrazně více vyšších kmitočtů, než úsek první skladby.



Obr. 10.2: Porovnání množství soufázové (modře) a protifázové (oranžově) složky dvou různých skladeb

Druhým krokem je implementace RIAA záznamové charakteristiky. Pokud by se nepoužila, dalo by se říci, že horizontální a vertikální výchylka budou přímo úměrné příslušným signálům. Koeficienty filtru byly převzaty z [3] a platí pro použitou zdrojovou vzorkovací frekvenci 44,1 kHz. Na obrázku 10.3 je přes sebe vložena skutečná charakteristika a její použitá aproximace a je zřejmé, že přesnost je více než dostatečná (minimálně pro účely simulace, kde bude v obou směrech konverze použita tatáž křivka, resp. její inverze).



Obr. 10.3: Srovnání skutečné a aproximované záznamové charakteristiky- červená linie je graf aproximace z matlabu v odpovídajícím rozsahu os (podklad převzat z [2])

## 10.2 Simulace geometrických vlastností drážky

V další fázi už je jen potřeba filtrovaný signál převést na geometrický model drážky podle vzorců (2.3-2.6) v kapitole 2.3.2. Jak je uvedeno v kapitole 4.1, je třeba zvolit vhodnou konstantu  $K$ . Tu můžeme zvolit například takovým způsobem, aby nám rozsah virtuálního vertikálního řezu (tedy  $\pm 15 \mu\text{m}$ ) plně pokryl rozsah hodnot filtrovaného signálu  $S$ , který právě díky filtraci (v závislosti na množství protifází a jejich frekvenci) dosahuje o 2-3 řády nižších než signál jednotlivých kanálů před filtrací). Konstanta tedy může být rovna  $K = 15 / \max(|S(t)|) [\mu\text{m}/-]$ . Druhou konstantou nutnou pro simulaci obrazové informace o drážce je počet pixelů na  $1 \mu\text{m}$ . Pro konkrétní příklad řekněme, že budeme chtít simulovat záznam z kamery o rozlišení 2048 pixelů. Jelikož předem neznáme maximální horizontální výchylku, musíme počítat s tím, že drážka bude místy dosahovat šířky  $70 \mu\text{m}$  a možné výchylky o stejné hodnotě na obě strany. Konstantu tedy můžeme stanovit podle vzorce (10.1):

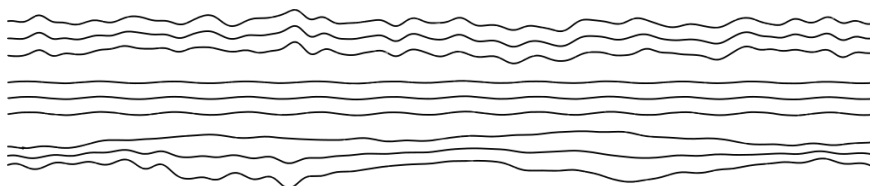
$$f s_{\mu\text{m}} = \text{floor} \left( \frac{N}{3 \cdot w_{\text{max}}} \right) = \text{floor} \left( \frac{2048}{210} \right) = 9 [\text{pix}/(\mu\text{m})]. \quad (10.1)$$

Pro další krok simulace potřebujeme znát rychlost otáček desky a poloměr drážky, na které se vyskytujeme, abychom spočetli drážkovou rychlost podle dříve zmíněného vzorce (2.1). Z drážkové rychlosti a délky zpracovávaného audiosignálu spočítáme délku simulované drážky v  $\mu\text{m}$  a po přenásobení konstantou  $K$  i v pixelech. Na základě těchto hodnot je pak nutné nainterpolovat audiosignály tak, aby jejich délka ve vzorcích odpovídala délce drážky v pixelech a bylo možné obraz drážky vůbec vytvořit. Interpolace lze snadno udělat pomocí přímého převzorkování v poměru  $\frac{v d_{\mu\text{ms}} \cdot f s_{\mu\text{m}}}{f_s}$ , kde  $v d_{\mu\text{ms}}$  je drážková rychlost v

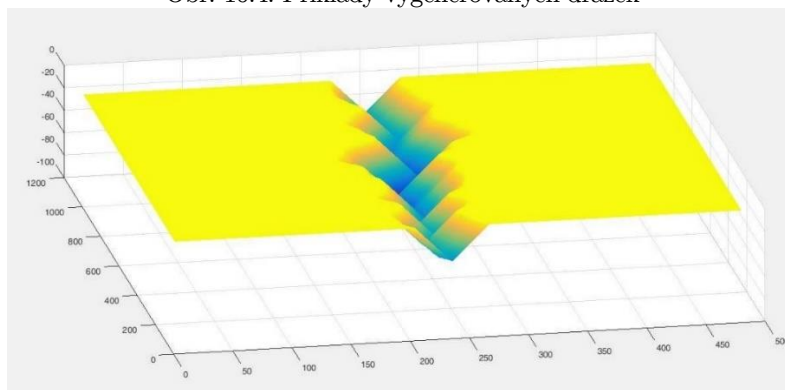
$\mu\text{m/s}$ ,  $f_{s_{\mu\text{m}}}$  zmíněná převodní konstanta snímacího systému a  $f_s$  vzorkovací frekvence zdrojového audia. Nakonec je potřeba (obdobně jako u rekonstrukce z rastrového obrazu) vzít v úvahu, že signál jako takový reprezentuje pouze okamžitou rychlost rycí jehly. Pro zjištění skutečné výchylky je tak potřeba provést opak numerické diference. V Matlabu je onou reverzní funkcí funkce *cumsum()*, neboli kumulovaný součet.

### 10.2.1 Obraz nasimulované drážky

Výsledný obraz drážky lze zobrazit různými způsoby. Buď ve 2D rovině, která nám umožňují veškerá následující zpracování (tedy zkreslení, rozmazání a následnou zpětnou rekonstrukci) nebo jako 3D model, který je sice efektní, pro jakékoliv další zpracování bohužel naprosto nevhodný. Na následujících obrázcích 10.4 a 10.5 jsou příklady drážek 3 různých skladeb a ukázka 3D modelu drážky. Drážka číslo 3 byla naschvál zvolena jako příklad drážky s takovým množstvím protifázového signálu, kterou by v reálném prostředí nebylo možné realizovat a bezpečně přehrát



Obr. 10.4: Příklady vygenerovaných drážek

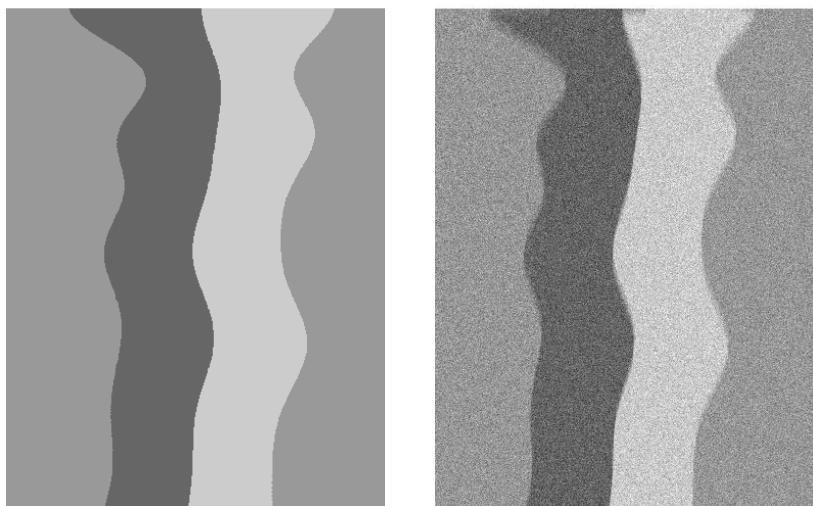


Obr. 10.5: 3D model drážky

### 10.2.2 Simulace degradace obrazu

V reálných aplikacích (tedy skutečném sejmutí kamerou) ovšem nelze čekat, že výstupem budou ostře definované 3 hrany. Je mnohem pravděpodobnější, že v závislosti na úhlu osvětlení získáme obraz obsahující alespoň 3 různé odstíny šedé barvy doplněné o rozmazání, šum apod. Pro účel simulace jsem původně využil funkci *Degrade()*, která kombinuje v rámci jedné funkce schopnost

kombinovat rozmazání v určeném směru a množství, přidání šumu různého typu a celkové ztmavení či zesvětlení obrazu. Nakonec jsem z důvodu pomalosti výpočtu zvolil rozmazání pomocí konvolučního jádra (z funkce *fspecial()*) v kombinaci s Gaussovským šumem. Zvolená filtrace odpovídá metodě pohybového šumu simulující pohyb kamery v rozsahu 15 pixelů o úhel 84°. Hodnoty byly zvoleny spíše experimentálně, než se snahou nasimulovat konkrétní rychlost pohybu (který by při správně seřízené kameře měl úhel rovný 90°. Na obrázku 10.6 je ukázáno, jak vypadá tato „realističtější“ podoba drážky a následně její degradovaná varianta.



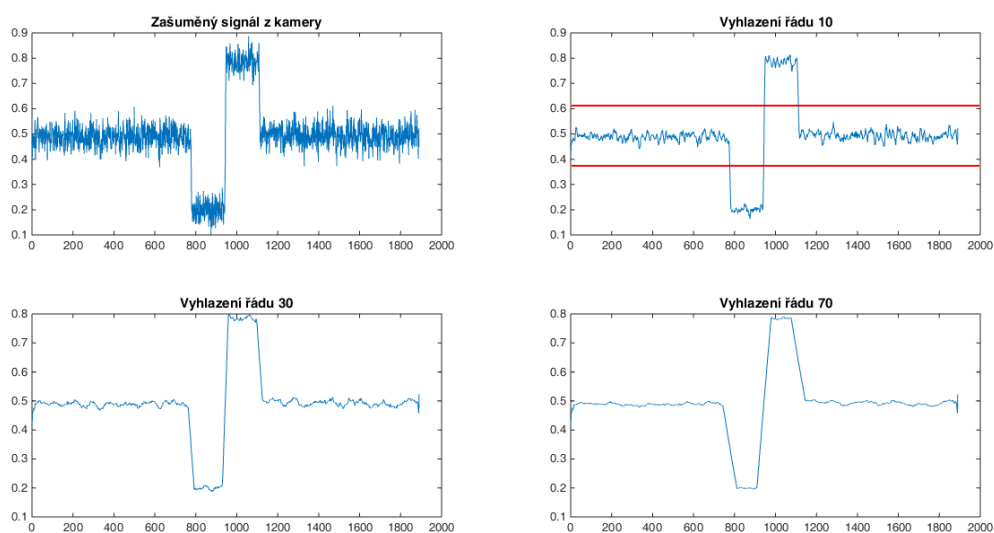
Obr. 10.6: Simulace obrazu drážky a její degradovaná podoba

### 10.2.3 Simulace čtení řádkovou kamerou a detekce hran

Jelikož snímání v ideálním případě probíhá pomocí řádkové kamery, nelze proto při simulaci zpracovávat celý obrazový signál, nýbrž pouze jednotlivé řádky v intervalech odpovídajících řádkové frekvenci kamery. Abych měl v projektu jistotu, že nebudu snímání simulovat na stejných vzorcích původního signálu, vzorkovaného na 44,1 kHz, zvolil jsem řádkovou frekvenci 48 kHz. Signál z řádkové kamery je jednorozměrný, je proto poměrně jednoduché jeho zpracování. Obdobně jako v opačném procesu simulace pro určení horizontální a vertikální výchylky stačí určit polohu hran drážky. V tomto případě dokonce ani nepotřebujeme znát polohu dna drážky, protože lze dopočítat jako aritmetický průměr poloh okrajů drážky. Na rozdíl od detekce hran v kapitole, kde došlo ke skutečné realizaci na skutečném vzorku, při použití řádkové kamery není třeba řešit žádné spojování snímků, neboť se jednotlivé řádky nijak nepřekrývají. V případě, že by bylo použito osvětlení, které by dosáhlo podobného výsledku jako na snímku 10.5, lze jako detektor hran použít prosté prahování.

V závislosti na množství šumu může být před detekcí vhodné signál vyhladit. Existuje více metod jak vyhlazení, tak detekce hran, jejichž vhodnost

či nevhodnost závisí na více parametrech, jakými je například množství šumu, reálná podoba nasnímaného signálu apod. Pro účel simulace a ověření algoritmu jsem zvolil nejjednodušší metody a to vyhlazení pomocí klouzavého průměru a následné prahování. Na obrázku 10.7 je znázorněn výstupní signál kamery odpovídající 2D obrazci z obrázku 10.6, různé úrovně vyhlazení a nastavené prahové úrovně pro stanovení pozice hran.



Obr. 10.7: Signál z řádkové kamery, různé úrovně vyhlazení a nastavení prahových úrovní pro detekci hran – příliš velké vyhlazení způsobí rozmazání hran a tím méně přesnou detekci

## 10.3 Rekonstrukce zvuku ze snímku drážky

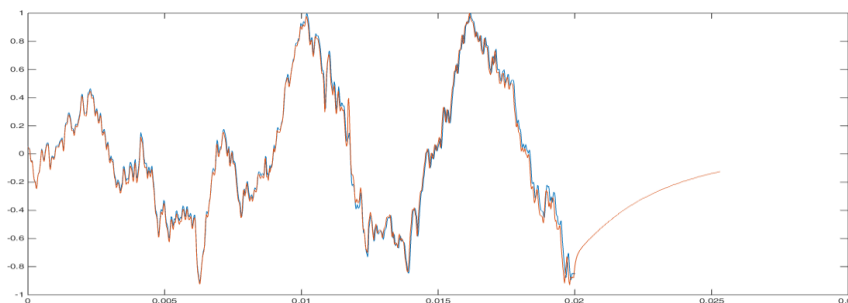
Ve chvíli, kdy jsou správně detekovány hrany, je potřeba navzorkovaná data z drážky převést zpět na zvuk. Postupy jsou zde v zásadě dva: Buď lze exaktně použít přesně tytéž vzorce jako při simulaci podoby drážky a z nich vyvodit opačný postup, nebo lze pracovat v relativním měřítku, signály zrekonstruovat takto a teprve poté celý výsledný signál normalizovat.

Mějme tedy detekovány polohy hran  $H1$ ,  $H2$  a  $H3$  pro každý jednotlivý „snímek“ řádkové kamery. Okamžitou hodnotu horizontální ( $M$ ) a vertikální výchylky ( $S$ ) – oboje v pixelech - tedy zpětně spočítáme jako  $H2 - wid/2$ , resp  $(H3 - H1)/2$ . Odpovídající signály pro levý a pravý kanál pak vypočítáme pomocí vzorců (4.3 a 4.4), s přihlédnutím k nutnosti signál dále normalizovat možno i bez konstanty, tedy pouze jako prostý součet a rozdíl mono a stereo složky rekonstruovaného signálu. Na ten je ještě nutné aplikovat inverzní RIAA filtr, pomocí funkce *filter()*, ve které se pouze vymění vektory koeficientů  $a$  a  $b$ , a opět číselnou diferencí.

Po těchto frekvenčních úpravách už jen dojde k normalizaci signálu pro úroveň odpovídající PCM kódování.

## 10.4 Vyhodnocení rekonstruovaného zvuku

Samozřejmě nejdůležitějším kritériem pro hodnocení algoritmu je jeho funkčnost. Na obrázku 10.8 je proto srovnání části signálu originálního a rekonstruovaného. Jak uvádím v předchozích kapitolách, oba signály mají rozdílnou vzorkovací frekvenci, aby bylo zřejmé, že se nejedná o odpovídající vzorky (ač vzhledem k povaze simulace zde určitá závislost samozřejmě je).



Obr. 10.8: Srovnání originálního a rekonstruovaného signálu v jednom kanálu

## 10.5 Vyhodnocení použité metody

Metoda simulace ukázala, že pokud by byla drážka řádkovou kamerou dobře sejmuta a nasvícena tak, aby se praktická podoba řádku alespoň přibližovala podobě uvedené na obrázcích 10.5, bylo by zpracování těchto dat poměrně jednoduchým úkolem. Ve skutečnosti nicméně celá simulace zpracování 5 vteřin dlouhého úseku hudby trvala téměř hodinu. To je v praxi samozřejmě naprosto nepoužitelný čas. Je ovšem nutné poznamenat, že nejvíce času zabraly operace související s 2D signálem, tedy samotná tvorba drážky a simulace její degradace.

Právě kvůli časové náročnosti, která – jak se ukázalo – při změně délky signálu rostla prudčeji než lineárně, byla implementována segmentace signálu na úseky dlouhé cca 30.000 pixelů. Při onom zpracování signálu dlouhého 5 vteřin tedy musel cca 780krát vzniknout, zkreslit se a vyhodnotit snímek s rozlišením přes 50 MPx. Při použití řádkové kamery a skutečné desky tyto operace samozřejmě odpadnou a zůstane pouze nutnost zpracovat velké množství jednorozměrných signálů, což již zdaleka nejsou tak náročné operace.

Bohužel vzhledem ke skutečné podobě obrazu drážky, jak ji známe z kapitol věnovaných rastrové kameře, lze předpokládat, že simulovaná podoba drážky je velmi optimistická, až téměř nereálná. Lze proto předpokládat, že ani samotné vyhodnocení v praxi nebude tak jednoduché a nebude možná každý navzorkovaný řádek zpracovávat zcela zvlášť, ale minimálně aplikovat nějaké modely, které budou brát v úvahu i podobu předchozích řádků.

## Kapitola 11

# Návrh na vylepšení metody a směr výzkumu

V kapitolách 6-10 je navržena a zkonstruována konkrétní metoda, umožňující bezkontaktní převod kontur drážky na zvukovou informaci. Výsledná kvalita zvuku a rychlost zpracování však není dostatečná na to, aby bylo možné systém používat coby optický gramofon a zároveň schopnost detekce chyb není na takové úrovni, aby byl systém použitelný coby systém pro optickou kontrolu. Existuje tak spousta možností, co lze do budoucna dále vylepšovat a jakým směrem by se mohl ubírat další výzkum v dané oblasti. Tyto možnosti nemohou být všechny vyřešeny v rámci této práce ať už pro jejich komplikovanost, časovou a výpočetní náročnost, ale i proto, že dedikovaný systém může být velmi drahý. Tato kapitola tedy bude zaměřena na tyto návrhy.

## 11.1 Specializované pracoviště optické kontroly

Prvním krokem by pak samozřejmě byla stavba pracoviště určeného právě pro specializovaný účel optické kontroly. Stavbou by se mohla podobat řezačce (tedy pracovišti, které jsem v práci použil a modifikoval pro snímání i já) s obdobnou variací krokového motoru s páskovým převodem. Jednou z konstrukčních nutností je použití vakuového přísávacího systému, neboť při velkém zvětšení mikroskop nabízí velmi omezenou hloubku ostrosti.

Zatímco pohonný systém by mohl zůstat v podstatě totožný, bylo by nutné vyřešit pohyb v kolmém směru dle kapitoly 6.4. Vzhledem k citlivosti systému na tento (při demonstraci metody spíše parazitní) pohyb by bylo nejvhodnější použít řešení s poháněnou závitovou tyčí. Toto řešení je jednak nejpresnější a zároveň tato tyč dodá systému další dodatečnou konstrukční tuhost. Při použití rastrové kamery a skládání snímků by pak stačilo, aby se po každé nasnímané rotaci systém posunul o danou vzdálenost. Pro snímání řádkovou kamerou by bylo možné pohyb optické soustavy řídit modulačním signálem shodným s tím, který byl použit pro řízení řezací hlavy.

Konstrukce umístěná na této rotační tyči by mohla kromě optického systému a snímací kamery obsahovat současně osvětlovací systém v podobě různých „pluginů“, neboť na určité účely může být vhodný jiný typ osvětlení z hlediska úhlu, barvy, apod. Například na detekci prořezů bude určitě vhodné takové světlo, které co nejvíce zdůrazní hrany drážek pro následné snímání (obdobně tomu, s jakým pracovala praktická část této práce), zatímco na detekci vytrhaného dna drážky bude lepší nasvítit negativní matici více ze strany pro co nejkontrastnější oddělení špiček drážky.



## 11.2 Synchronizované paralelní zpracování dat

Datová náročnost nekomprimovaného obrazu je obrovská a často je zbytečné všechny tyto informace ukládat. Rychlost celého systému nelze ani při použití výkonného počítače zrychlovat do nekonečna. Jak je uvedeno v kapitole 7, při použití snímkovací frekvence 25 snímků za vteřinu (tedy snímkování s periodou 40ms) byly vždy 3-4 snímky rozmazané vlivem pohybu a setrvačnosti rotačního talíře. Znamená to tedy, že při totožných mechanických vlastnostech bude talíř v pohybu vždy cca 160 ms. Pokud vezmeme v úvahu, že cca 11 vteřin záznamu vyžadovalo 3746 snímků a běžná deska může mít délku až 22 minut, dostaneme se k tomu, že celkový počet vyfocených snímků může být roven téměř půl milionu obrázků. Takový počet samozřejmě bude mít obrovské paměťové nároky a navíc se bude snímat téměř 20 hodin (při zachování krokovacího minima 160 ms). I proto by bylo ideální, aby byla data průběžně zpracovávána již při natáčení a nebyla zbytečně skladována. Pokud by byl postup stejný, jako mnou použitý, nasnímané video (komprimované) by mělo až 720 GB a vyselektované obrázky převedené do odstínů šedi (cca 2 MB každý) pak až k 1 TB. Byl by tedy velký problém už jenom skladovat data na pevném disku, natož s nimi pak pracovat v operační paměti.

## 11.3 Detekce vad

Jelikož má být systém použitelný zejména pro detekci vad ať už pozitivního originálu, či negativní matrice, bylo by dobré, aby kontrolní program obsahoval modelovou databázi vad, která by se s každou další nalezenou vadou učila a tak zvyšovala svoji účinnost. Kromě sledování hran pro účel přehraní zvuku by tak program sledoval veškeré abnormality a následně je vyhodnocoval, jejich parametry ukládal do zmíněné databáze a jejich snímky zobrazoval obsluze pro další vyhodnocení. Příkladem takto zobrazené vady je mnou implementovaná detekce dotyků či prořezů drážky.

Dále by bylo možné implementovat vícestupňové ostření, které by umožnilo sledovat hrany i dno drážky (případně vrchol v případě matic) a tak na základě rozdílu mezi skutečnou polohou dna a její teoretickou polohou vypočítanou jako aritmetický průměr polohy obou okrajů. Tento rozdíl by pak v praxi poukazoval na přítomnost vad galvanoplastiky, tedy zmíněná vytrhaná dna či zborčené hrany. Sledování ostrých kontinuálních hran pak zase lze využít pro detekci škrábanců, vzniklých například neopatrnou manipulací s nosičem.

## Kapitola 12

### Závěr

Ačkoliv se může možnost převodu optické informace o gramofonové drážce na zvuk zdát jako jednoduchý úkol, ukázalo se, že ač se díky samotné podstatě analogového mechanického záznamu v principu jedná o poměrně snadnou a dobře pochopitelnou záležitost, je zde hned několik omezení.

Tato omezení nejsou ani tak na straně techniky a HW, kde většina omezení jde vyřešit použitím rychlé řádkové kamery, jako spíše na straně SW zpracování signálu. To je úloha náročná a díky mikroskopickým rozměrům drážky a neuvěřitelně jemných detailů se pak při digitalizaci obrazu jedná o obrovské množství dat.

Myslím, že je ale možné prohlásit, že princip optického čtení byl naznačen, bylo ukázáno, že omezené rozlišení, které zdaleka neodpovídá 16 bitům v PCM kódu, pro demonstraci metody a hledání problémů na desce stačí.

Teoretická část práce byla zaměřena na shrnutí principů a možností samotného mechanického záznamu zvuku a zejména pak na principy, pomocí kterých je možné z obrazu drážky zpětně rekonstruovat zvukovou informaci. Dále byl prezentován přehled literatury a byla shrnuta většina dosavadních pokusů o optické čtení z hlediska použité technologie, techniky, pracovního postupu a u prací, u kterých byly výsledky dohledatelné, i zhodnocení zvukové kvality výstupu. Byly probány teoretické omezení dvou typů kamer a vypočteny některé maximální parametry, jichž je možné v ideálních případech dosáhnout. Závěrem teoretické části bylo shrnutí důležitých faktů a principů v oblastech, jejich znalost umožnila či usnadnila převod celého navrženého systému do funkčního celku vhodného pro demonstraci metod.

V rámci praktické části práce byl navržen, sestaven a otestován systém pro optické čtení měděných originálních nosičů mechanického záznamu. Na základě teoretických i praktických znalostí nabytých v tomto oboru a studia příslušné literatury, byla pro demonstraci možností optického čtení využita standardní rastrová kamera a snímací systém založený na krokové rotaci zvukového nosiče a postupném snímání pomocí staticky připevněné kamery. Všechna získaná data (kterých bylo skutečně velké množství) byla v programu Matlab následně zredukována odebráním irelevantních a redundantních snímků.

Již užitečná data byla následně pomocí navrženého algoritmu pro hledání podobnosti snímků posunutých v jedné ose spojena do několika velmi dlouhých snímků, ve kterých byla následně pomocí detektoru hran najita poloha veškerých významných a zřetelných hran. Zároveň došlo k převodu snímků v odstínech šedi na snímky pouze čistě černobílé, což ušetřilo velké množství paměťových prostředků.

Hrany byly následně sledovány pomocí opět mnou navrženého algoritmu se schopností nalézt i hrany přerušené a poškozené. Zároveň se díky povaze hran, odlesků apod. občas stalo, že byla hrana vyhodnocena chybně. Proto byl navržen opravný algoritmus, který tyto problémy řešil i za cenu ztráty informace v daném úseku. Na závěr byly veškeré informace o drážkách převedeny do zvukové podoby na základě teoretických znalostí o tom, jakým způsobem je drážka řezána a následovně čtena a bylo provedeno vyhodnocení zvukové kvality a parazitních signálů.

Velkou nevýhodou programu je ovšem jeho velká časová a paměťová náročnost. Natočená videa, která obsahují informaci o pouze cca 11 vteřinách záznamu, mají velikost okolo 9 GB. Po detekování hran už se nicméně jedná o pouhých 300 MB. Nicméně zejména na spojování snímků dohromady by bylo potřeba do praktického využití navrhnout lepší a časově úspornější algoritmy.

### **Osobní zhodnocení**

Jelikož povaha experimentu a nutné technické vybavení nenechává příliš prostoru pro využití systému veřejností (a to ani odbornou), podepsalo se to výrazně na podobě matlabového kódu programu pro zpracování dat. Zjednodušeně řečeno, nejedná se o typ programu, který by stačilo jednou zapnout a následně byla obsluha provedena od začátku až po závěrečný zvukový výstup. Ať už samotné nasegmentované zpracování dat, tak zejména různé opravné algoritmy vyžadují poměrně velkou míru spolupráce uživatele, který by v ideálním případě měl vědět, co dělá a čeho chce dosáhnout.

Do budoucna by samozřejmě bylo nejlepší sestrojít program s implementovaným grafickým rozhraním, které by uživatele programem skutečně provedlo, informovalo ho o průběhu a vyžadované informace by se zadávaly do připravených textových polí.

Skutečnost, že byla práce zadána ve spolupráci s firmou z praxe a skutečně během ní bylo dosaženo „hmataelného“ výsledku, hodnotím jako jednoznačný přínos. Ačkoliv ve fázi, v jaké se práce nachází, zatím v podstatě nemůže být přímým přínosem v rámci výrobního procesu, věřím, že přináší stabilní základy a přehled poznatků, na kterých bude moci být postaven další výzkum.



## Kapitola 13

### Použitá literatura:

- [1] Information and Measurement: Digital Vinyl?! [online]. University of St. Andrews, 100-109 [cit. 2016-02-13]. Dostupné z: [https://www.st-andrews.ac.uk/~www\\_pa/Scots\\_Guide/iandm/pdf/chapter12.pdf](https://www.st-andrews.ac.uk/~www_pa/Scots_Guide/iandm/pdf/chapter12.pdf)
- [2] GALO, Gary A. The Columbia Lp Equalization Curve. *ARSC Journal*. 2009.
- [3] UMMINGER, Frederick. Digital RIAA equalization filter coefficients [online]. [cit. 2016-02-13]. Dostupné z: <http://www.musicdsp.org/showone.php?id=122>
- [4] Line Scan Cameras: Function and Applications [online]. Schäfter & Kirchhoff, GmbH [cit. 2016-02-13]. Dostupné z: <http://www.sukhamburg.com/download/cam-introduction.pdf>
- [5] DIRJISH, Mat. What's The Difference Between Moving Magnet, Coil, And Iron Cartridges For Turntables? *Electronic Design* [online]. 2012 [cit. 2016-12-27]. Dostupné z: <http://electronicdesign.com/components/what-s-difference-between-moving-magnet-coil-and-iron-cartridges-turntables>
- [6] TIAN, Baozhong a Barron JOHN L. Using Computer Vision Technology to Play Gramophone Records. In: *JAES Volume 59 Issue 7/8*. 2011, s. 514-538.
- [7] HENSMAN, Arnold a Kevin CASEY. Optical Reading and Playing of Sound Signals from Vinyl Records. *The 7th IT&T Conference (ITT07)*. Dublin, Irsko: Institute of Technology Blanchardstown, 2007.
- [8] LABORELLI, Louis, Jean-Hugues CHENOT a Alain PERRIER. Non contact phonographic discs digitisation using structured colour illumination. *AES Convention Paper: Presented at the 122nd Convention*. Vídeň, Rakousko, 2007.
- [9] SPRINGER, Olaf. *Digital Needle: A Virtual Gramophone*[online]. 2002 [cit. 2016-12-21]. Dostupné z: <http://www.cs.huji.ac.il/~springer/DigitalNeedle/index.html>
- [10] STOTZER, Sylvain. *Phonographic Record Sound Extraction By Image Processing*. Fribourg, Švýcarsko, 2006. Faculté des sciences - Université de Fribourg. Vedoucí práce Prof. Rolf Ingold.
- [11] IWAI, T., T. ASAKURA, T. IFUKUBE a T. KAWASHIMA. Reproduction of sound from old wax phonograph cylinders using the laser-beam reflection method. *Appl. Opt.* 1986, 27(13).
- [12] FADEYEV, Vitaliy a Carl HABER. Reconstruction of Mechanically Recorded Sound by Image Processing. *LBNL Report 51983*. Berkeley, Kalifornie, 2003.

- [13] RYDLO, Pavel. *Krokové motory a jejich řízení* [online]. [cit. 2016-12-21]. Dostupné z: <http://forum.mcontrollers.com/download.php?id=1054&sid=acf2d6b4878bd6900d9981e443a17d07>
- [14] LEARN ENGINEERING. *How Does a Stepper Motor work?: (video)* [online]. [cit. 2016-12-21]. Dostupné z: <https://www.youtube.com/watch?v=eyqwLiowZiU>
- [15] HLAVÁČ, Václav. *Hledání hran: přednáška*. ČVUT, katedra kybernetiky. Praha: Centrum strojového vnímání, ČIIRK.
- [16] FATAFTA, Ishraq. *Edge Detection* [online]. 2012 [cit. 2016-12-21]. Dostupné z: <http://www.slideshare.net/ishraqabd/edge-detection>
- [17] MION, Endi. *Edge Detection* [online]. 2016 [cit. 2016-12-21]. Dostupné z: <http://www.slideshare.net/EndiMion3/edge-detection-62287272>
- [18] Total harmonic distortion. Wikipedia [online]. [cit. 2017-01-06]. Dostupné z: [https://en.wikipedia.org/wiki/Total\\_harmonic\\_distortion](https://en.wikipedia.org/wiki/Total_harmonic_distortion)
- [19] FARINA, Angelo. Simultaneous measurement of impulse response and distortion with a swept-sine technique. Università di Parma.
- [20] FENYK, Miloš. Modul zpracování obrazu s DSP-Blackfin a jeho využití v bezdotykovém měření. 2009. Diplomová práce. ČVUT, FEL. Vedoucí práce Ing. Jan Fischer, CSc.

## Seznam tabulek

Tabulka 1: Předpokládané parametry rekonstruovaného zvuku .....	38
Tabulka 2: Logické schéma fází krokového motoru:.....	45

## Seznam obrázků

Obr. 2.1: Schéma procesu výroby gramofonové desky v GZ Media .....	16
Obr. 2.2: DMM řezací aparatura .....	17
Obr. 2.3: kombinovaná záznamová charakteristika (převzato z [2]) .....	18
Obr. 2.4: RIAA charakteristika (převzato z [2]) .....	19
Obr. 2.5: Detail řezací hlavy (převzato z internetu).....	19
Obr. 2.6: Schéma směru výchylky v závislosti na okamžitém signálu, způsob skládání monofonní a stereofonní složky signálu .....	20
Obr. 2.7: Průřez drážky bez vertikální výchylky a v obou „extrémech“ .....	20
Obr. 2.8: Jedna z podob loga vydavatelství His Master's Voice (převzato z internetu).....	22
Obr. 2.9: Stavba přenosky MM (vlevo) a MC (vpravo) systému (převzato z [5])....	22
Obr. 3.1: Příklad poškozené archivní desky (převzato z [10]) .....	23
Obr. 3.2: Laserové čtení Edisonovského válečku (převzato z [11]) .....	24
Obr. 3.3: Diagram použití fotodiody jako snímacího prvku (převzato z [11]) .....	24
Obr. 3.4: Část desky v podobě, v jaké s ní pracoval Springer (převzato z [9]).....	26
Obr. 3.5: Překrývající se data před a po zprůměrování (převzato z [12]) .....	28
Obr. 3.6: Jednotlivé kroky VisualAudio Conceptu (převzato z [10]) .....	29
Obr. 3.7: Mechanismus pohonu a čtení VisualAudio Projectu (převzato z [10]) .....	30
Obr. 3.8: Srovnání zdrojového snímku a jejího gradientu v obou směrech (převzato z [10]).....	32
Obr. 4.1: Drážka v případech, že je signál (zleva): Pouze v pravém kanále, pouze v levém kanále, v obou stejný (mono), v každém kanále jiný (stereo), v kanálech vzájemně v protifázi (převzato z internetu) .....	34
Obr. 4.2: Naznačení maximální možné horizontální výchylky ve FullHD obraze....	35
Obr. 5.1: Různé typy hran – idealizované: a) skoková hrana, b), c) konvexní a konkávní svahová h., d) střeňová h., e) údolní h., f) schodišťová h., g.) špičková h. (převzato z [15]) .....	40
Obr. 5.2: Princip postupného buzení fází a krokové rotace (převzato z [13]).....	43
Obr. 5.3: Stavba typického krokového motoru s rozlišením 200 kroků (převzato z [14]).....	44
Obr. 5.4: Schéma zapojení krokového motoru do Arduina přes H-můstek .....	46
Obr. 6.1: Určení reálného rozlišení snímku v $\mu\text{m}$ .....	48
Obr. 6.2: Nákres konstrukce externího pohonu.....	51
Obr. 6.3: Princip páskového posuvu.....	52
Obr. 6.4: Princip posuvu pomocí závitové tyče .....	52
Obr. 6.5: Vliv osvětlení na obraz drážky (převzato z [10]).....	54
Obr. 7.1: Detail připevněného krokového motoru na řezací aparatuře.....	56

Obr. 7.2: Postupné střídání ostrých a pohybem rozmazaných snímků .....	57
Obr. 7.3: Porovnání jednotlivých kanálů barevného snímku .....	57
Obr. 8.1: Algoritmus spojení snímků a detekce hran .....	58
Obr. 8.2: Spojený obrázek s označením místa spojů („švy“).....	60
Obr. 8.3: Ukázka jiných detektorů hran. (zleva: Cannyho metoda v defaultním nastavení, Cannyho metoda se zvýšeným prahováním, metoda Prewittové v horizontálním směru) .....	61
Obr. 8.4: Srovnání bitmapy v ostínech šedi a výsledné detekované hrany (v negativu) .....	62
Obr. 8.5: Princip vyhledávání drážky v tolerančních pásech .....	63
Obr. 8.6: Grafické znázornění detekovaného průběhu drážky. Kromě rozdílného měřítka v horizontální a vertikální ose je analogické k bitmapovému obrazu drážky .....	63
Obr. 8.7: Ukázka chybně detekovaných – spojených – drážek .....	64
Obr. 8.8: Ukázka detekovaných prořezů .....	65
Obr. 9.1: Grafické znázornění složené informace o všech 8 drážkách .....	66
Obr. 9.2: Kompletně spojený vektor reprezentující 7 drážek z 8 sérií snímků .....	67
Obr. 9.3: Frekvenční charakteristika spočítaná ze sweep signálu .....	70
Obr. 9.4: Frekvenční charakteristika spočítaná z tiché drážky .....	71
Obr. 9.5: Rekonstruovaný signál v absolutních hodnotách a v dB stupnici .....	71
Obr. 9.6: Spektrogram rekonstruovaného signálu.....	72
Obr. 9.7: Impulzní odezva s patrnými artefakty harmonického zkreslení .....	72
Obr. 10.1: Schéma simulace snímání řádkovou kamerou .....	73
Obr. 10.2: Porovnání množství soufázové (modře) a protifázové (oranžově) složky dvou různých skladeb.....	74
Obr. 10.3: Srovnání skutečné a aproximované záznamové charakteristiky- - červená linie je graf aproximace z matlabu v odpovídajícím rozsahu os.....	75
Obr. 10.4: Příklady vygenerovaných drážek .....	76
Obr. 10.5: 3D model drážky.....	76
Obr. 10.6: Simulace obrazu drážky a její degradovaná podoba .....	77
Obr. 10.7: Signál z řádkové kamery, různé úrovně vyhlazení a nastavení prahových úrovní pro detekci hran – příliš velké vyhlazení způsobí rozmazání hran a tím méně přesnou detekci.....	78
Obr. 10.8: Srovnání originálního a rekonstruovaného signálu v jednom kanálu.....	79



## Seznam zkratek a symbolů

- GZ – gramofonové závody
- FTP – file transfer protocol (protokol na přenos souborů)
- VTM – visual track marker (vizuální oddělovač stop)
- DMM – direct metal mastering (technologie řezání nosiče do plechu)
- RIAA - Recording Industry Association of America
- MM – moving magnet (systém přenosky s pohyblivým magnetem)
- MC – moving coil (systém přenosky s pohyblivou cívkou)
- MIC – moving iron cardridge (obdoba MM s feromagnetikem)
- THD – total harmonic distortion (harmonické zkreslení)
- CCD – charge-coupled device (typ snímače)
- CMOS – complementary metal-oxide-semiconductor (typ snímače)
- SNR – signal to noise ratio (odstup signálu od šumu)
- $\delta[]$  – Diracův (či jednotkový) impulz
- LoG – Laplacian of Gaussian (Laplacián Gaussiánu – matematický operátor)
- TTL - transistor-transistor-logic (standard používaný pro logické obvody)
- LED – light emitting diode (svítivá dioda)
- MSE – mean squared error (střední kvadratická chyba)
- DFT – discrete Fourier transform (diskrétní Fourierova transformace)
- PCM – pulse-code modulation (pulzně kódová modulace – modulace audia)



## Příloha A

# Obsah příloženého DVD

MATLAB .....	složka se skripty a ‚mat‘ soubory z mezikroků
ZDROJOVE_VIDEO .....	jedno ze 2 zdrojových ‚avi‘ videí <sup>15</sup>
ZDROJOVE_AUDIO .....	audiosoubor zpracovatelný v rámci simulace
ZPRACOVANE_OBRAZKY .....	spojené snímky s detekovanými hranami <sup>16</sup>
Mohl_DP_text_final.pdf .....	text práce
Mohl_DP_export_final.wav .....	výsledné rekonstruované audio
readme.txt .....	doplňkový soubor specifikující jednotlivé přílohy

---

<sup>15</sup> Kvůli datové velikosti nejsou zahrnuta v příloze v KOSu.

<sup>16</sup> Kvůli datové velikosti bude v příloze v KOSu zahrnut pouze jeden z 8 těchto snímků.