

ČESKÝ METROLOGICKÝ INSTITUT

OBLASTNÍ INSPEKTORÁT BRNO

ZÁVĚREČNÁ ZPRÁVA Č. 6014-TR-Z0001-22

o plnění úkolu TR

AFM měření na velkých plochách

Odpovědný řešitel: Mgr. Petr Klapetek, Ph.D.

Vypracovali:

Mgr. Petr Klapetek, Ph.D. Mgr. Miroslav Valtr, Ph.D. doc. Mgr. Jan Martinek, Ph.D.

Ředitel VOJ:

Ing. Radovan Wiecek ředitel OI Brno

Datum:

30. listopadu 2022

Schválil:

RNDr. Pavel Klenovský odborný ředitel pro FM



Počet listů: 2/39

Resumé

Cílem úkolu bylo vyvinout nástroje pro měření rozměrů a úhlů na mikrostrukturách využívaných v optickém a elektronickém průmyslu na plochách až 5x5 cm s nejistotou menší než 10 nm pro měření rozměrů a menší než 50 mikroradiánů pro měření úhlů na povrchu vzorků. Pro tyto účely byly vyvinuty nástroje pro kompenzaci rotací při pohybu polohovacích systémů dvou mikroskopů, nástroje pro přenos návaznosti při měření úhlu a postupy pro vyhodnocování periody z mikroskopických dat. Součástí úkolu byly také přípravné kroky v další dlouhodobé aktivitě oddělení 6014 - měření teploty pomocí Johnsonova šumu, kde jsme se zaměřili na technickou přípravu pro porovnání s alternativní primární realizací na oddělení primární metrologie tepelně-technických veličin.



Obsah

Re	sumé								
O	bsah			3					
1	ÚVO	DD		4					
2	POI	PIS ŘEŠENÍ ÚKOLU							
	2.1	Měřen	í úhlů	6					
		2.1.1	Parazitní rotace při měření na metrologickém SPM	9					
		2.1.2	Nová detekce rotací pomocí interferometrů	10					
		2.1.3	Kompenzace rotací	12					
		2.1.4	Generátor malých rotací	16					
		2.1.5	Senzor autokolimátoru pro přenos návaznosti úhlu	16					
		2.1.6	Rozšíření kompenzace parazitních rotací pro měření na velkých plochách	19					
	2.2	Měřen	í periody	23					
	2.3	Termo	metrie na základě Johnsonova šumu (JNT)	29					
		2.3.1	Nová verze zesilovačů	29					
		2.3.2	Předběžné testy	30					
		2.3.3	Teplotní závislost zesilovačů a jejich kalibrace	31					
		2.3.4	Spolupráce s oddělením 1012	32					
3	DOSAŽENÉ VÝSLEDKY								
	3.1	Měřen	í úhlu	34					
	3.2	Měřen	í periody	34					
	3.3	Komp	lexní charakterizace mřížky	35					
4	ČEI	RPANÉ	NÁKLADY NA ÚKOL	36					
5	ZÁV	VĚRY		37					
Se	znam	bibliog	grafických citací	38					
PÌ	ŔÍLO	HY		39					



Počet listů: 4/39

1. ÚVOD

Smyslem měření na velké ploše (v kontrastu s měřením na typických komerčních zařízeních s omezeným rozsahem), při charakterizaci mikro- a nanostruktur, je především snížení nejistoty měření. S tímto přístupem se proto můžeme setkat všude, kde k tomu existují potřebná zařízení tj. například v německém metrologickém institutu PTB. Požadavky na nejistoty v řádu jednotek nanometrů jsou přitom ze strany zákazníků velmi časté. Typické vzorky určené k měření jsou v difrakční mřížky, využívané jako přenosové etalony délky při kalibraci dalších mikroskopů. S postupným pronikáním mikro- a nanotechnologií do různých oblastí průmyslu se objevují také nové požadavky na speciální měření period a úhlů různých velkoplošných struktur, které jsou takovým typickým vzorkům velmi vzdálené - například měření vad v rozmístění jednotlivých optických prvků na CCD detektoru. I metrologická zařízení schopna měřit s dostatečnou přesností (jednotky nm) jsou často schopna měřit jen na poměrně malé ploše (například 200x200 mikrometrů v případě našeho státního etalonu využívaného v nanometrologii), což může být pro takové úlohy nedostačující a vede to ke snaze o rozšíření měření na větších plochách. To s sebou přináší nové metodické otázky stran volby měřené oblasti, i nové výzvy v oblasti konstrukce mikroskopů Měření úhlů je ještě obtížnější – obecně se mikroskopy atomárních sil (AFM) využívají pro měření v rychlé ose skenování a data v pomalé ose skenování jsou považována za méně přesná. Měření úhlů pomocí AFM je proto obecně špatně popsané a málo rozvinuté téma metrologie a to navzdory tomu, že požadavky zákazníků na taková měření tvoří cca polovinu poptávek.

Cílem úkolu je rozvinout oblast měření periody a úhlu, a to jak rozšířením měřeného rozsahu, tak zpřesněním měření pomocí kompenzace úhlových chyb a vytvořením metrologické návaznosti pro malé úhly. Tyto dva úkoly jsou z velké míry provázané - kompenzace úhlových chyb se přímo projeví i ve snížení nejistoty měření periody díky snížení vlivu Abbého chyb, nicméně při popisu našich aktivit se pro jednoduchost zabýváme každou oblastí samostatně.

V oblasti měření úhlů jsme se zaměřili na dva aspekty. Prvním je minimalizace nejistot souvisejících s úhlovým natáčením skeneru mikroskopu při jeho pohybu. Natáčení skeneru formou Abbého chyb přispívá k nejistotě měření geometrických veličin, což s ohledem na to, že měření úhlů mezi strukturami na povrchu zkoumaného vzorku vyžaduje měření ve dvou směrech, má na výsledky měření úhlu podstatně větší vliv než na měření lineárních rozměrů. Pro aktivní kompenzaci úhlových chyb byly proto vytvořeny hardwarové i softwarové nástroje, které byly implementovány jak na státním etalonu délky a tvaru v oblasti nanometrologie, tak na platformě pro hybridní měření vyvinuté v roce 2021. Druhým aspektem je vytvoření řetězce návaznosti pro malé úhly, který by bylo možné využít jak směrem k porovnáním s jinými odděleními ČMI, tak pro určení úhlových nejistot dalších mikroskopů, které samy o sobě kompenzaci parazitních rotací nemají.

V oblasti měření periody se, kromě výše uvedených testů systému pro měření na velké ploše, jednalo především o výzkum v oblasti zpracování dat. Úsilím metrologických institutů, zejména PTB, vznikly v minulých letech některé normativní dokumenty, například ISO 11952, které se měřením a vyhodnocením periodických struktur zabývají. Problematika volby periody a jejího vyhodnocení z měřených dat je v nich nicméně podána poměrně stručně a zdaleka nevylučuje mylnou interpretaci, jak je možno vidět i při studiu výsledků různých porovnání, či při diskuzi s partnery z jiných metrologických institutů. Analýza, do které jsme se zapojili a na které se zcela zásadní měrou podílel Mgr. David Nečas Ph.D. (nyní CEITEC VUT) ukázala, že nejistotu měření periodických struktur je možné zásadním způsobem snížit volbou co možná nejdelších měřených profilů, dalece nad doporučení normy. Naopak, mnohé postupy, které měly být podle praxe, normy i literatury vhodné pro snížení nejistoty se ukázaly jako zcela kontraproduktivní. Výsledná publikace (nyní v recenzním řízení, připojena v příloze zprávy) tak po teoretické stránce podporuje koncepci měření na co možná největších plochách, kterou dlouhodobě sledujeme.

V rámci řešení úkolu bylo také pokračováno v některých dlouhodobých rozvojových aktivitách, které



Počet listů: 5/39

jsou ve zprávě krátce zmíněny. V oblasti měření teploty prostřednictvím šumu se jednalo například o technické přizpůsobení prototypu teploměru pro navázání spolupráce s Oddělením primární metrologie tepelnětechnických veličin a pro přípravu budoucího experimentu k porovnání prototypu Johnsonova šumového teploměru a sestavy primárního akustického plynového teploměru. Šumový teploměr je na AFM měření, které jsou primárním tématem úkolu navázány prostřednictvím měřených vzorků (etalony drsnosti) a sond (rastrovací termální mikroskopie). Z pohledu dlouhodobých aktivit je v příloze zprávy také zmíněna analýza vlivu drsnosti povrchů na optická měření, která souvisí s naší snahou o postihnutí vlivu drsnosti na měřená data v co možná nejširších souvislostech a také se snahou o větší využuití AFM jako metody pro nezávislé stanovení drsnosti při vyhodnocování dat z optických měření.



POPIS ŘEŠENÍ ÚKOLU 2.

Jak bylo zmíněno v úvodní části zprávy, popis řešení je rozdělen na část zabývající se měřením úhlů na mikro- a nanostrukturách a část zabývající se měřením periody. Je nicméně vhodné opět zdůraznit, že část zabývající se úhly je do velké míry propojena i se všemi dalšími měřeními pomocí AFM, a to prostřednictvím systému pro kompenzaci parazitních rotací, jehož využitím se měření pomocí mikroskopu zpřesní ve všech ohledech.

2.1. Měření úhlů

Požadavky na měření úhlů pomocí mikroskopie atomárních sil (AFM) mohou být v principu dvojího druhu, jak je naznačeno v obrázku 1. Může se jednat o měření sklonů vůči rovině vzorku, například při měření sklonu faset struktur pro zachytávání světla v solárních článcích, či analýze otisků vnikacího tělesa v instrumentované indentaci. Pro taková měření máme k dispozici již řadu nástrojů pro zpracování dat v programu Gwyddion, na jehož vývoji se naše oddělení podílí. Pokud je měření provedeno tak, že měřený sklon je možné zachytit v rámci profilů v rychlé ose měření mikroskopu, nejsou nejistoty měření úhlu zásadně jiné, než u jakýchkoliv jiných měření rozměrovýách parameterů pomocí AFM. Hlavním problémem jsou omezení rozsahu měřených úhlů daná sklonem stěn hrotu - tj. není možné měřit velmi příkré sklony. Druhým případem, ve kterém mluvíme také u úhlech, je měření úhlů v rovině povrchu vzorku, například při analýze neortogonality difrakční mřížky. Zde je situace podstatně složitější - požadavky přesnost měření ortogonality jsou vysoké a měření není možno provést jen v rychlé ose měření (je nutné stanovit směr struktur tvořících difrakční mřížku ve dvou osách), což přináší zásadní problémy z pohledu ortogonality měřicího systému i z pohledu mechanických či tepelných driftů. Právě fakt, že obě osy, rychlá, v jejímž směru se měří "řádky"ve výsledném obrázku a pomalá, v jejímž směru tyto řádky přibývají, nejsou ekvivalentní, je hlavní příčinou problémů. Problomatika správného měření úhlů v rovině vzorku, tj. tento druhý případ, byly proto hlavní náplní našich aktivit při řešení úkolu technického rozvoje.



Obrázek 1: Ilustrace základních typů měření úhlových veličin na AFM v nejjednodušším možném uspořádání: (vlevo) měření sklonů na anizotropně leptané mřížce s využitím jednoho profilu, (vpravo) měření úhlů v rovině vzorku na difrakční mřížce využívané pro kalibraci mikroskopů s využitím dvou profilů.

Pokud nahlédneme do odborné literatury, zjistíme, že problematika měření úhlů je podstatně méně



pokryta, než je tomu v případě jiných délkových měření v AFM, například periody mřížky. Měřením úhlů se zabývá například PTB v publikaci věnované měření difrakčních mřížek [1]. Pro tyto účely využívají vysoce přesný polohovací systém Nanopositioning and Nanomeasuring Machine (NMM) [2], který umožňuje měření na plochách až 25x25 mm. Ze sad velmi dlouhých profilů ve dvou osách stanovují orientaci obou os mřížky a jejich vzájený úhel stanovují s nejistotou cca 40 μ rad. Klíčovým prvkem pro dosažení tak malé nejistoty je přitom velmi malá úhlová chyba samotného mikroskopu (neortogonalita os), kterou autoři studie překvapivě získávají jen z dokumentace výrobce a jedná se o hodnotu 10 μ rad. Při využití jiného metrologického mikroskopu v korejském institutu KRISS byla pro srovnání dosažena o něco vyšší hodnota nejistoty, 100 μ rad [3], která se nicméně, ve srovnání se zcela nepředvídatelnými hodnotami, které by bylo možné získat na komerčních mikroskopech, dá považovat za velký úspěch.

Jak je zmíněno již v úvodu zprávy, naší snahou je dosáhnout nejistoty měření úhlu nižší, než je 50 μ rad. Měření dat pro vyhodnocení úhlů provádíme v praxi nejčastěji v souladu s postupem navženým v Ref. [1], který spočívá v natočení mřížky tak, aby byla zhruba orientovaná stejně jako jsou osy skeneru mikroskopu a následném měření dvou obdélníkových snímků topografie, jednoho pokrývajícího jednu řadu objektů tvořících periodickou strukturu mřížky a jednoho pokrývajícího jeden sloupec objektů, tak jak je to naznačeno na obrázku 2. Tato data jsou dále zpracována, v našem případě v programu Gwyddion.



Fast scan direction

Obrázek 2: Schema měření dat pro vyhodnocení úhlu dle Ref. [1]

Pro úvahy o vhodné volbě parametrů při měření je vhodné otestovat našich schopnost vyhodnotit malé rozdíly v úhlech z měřených dat. Pokud bychom uvažovali o typickém AFM snímku s velkým, avšak běžně používaným, rozlišením (1024×1024 pixelů), jako je to v obrázku 3, prostou geometrickou úvahou zjistíme, že z poměru počtu pixelů a velikosti pixelu by melo být největší možné úhlové rozlišení 1 mrad. To odpovídá rozdílu mezi profilem, pro ilustraci například celým v ose x, a profilem, který je na jedné straně v ose y posunut právě o jeden pixel. Vyhodnocení proto není vhodné provádět přímým odečítáním profilů, kde bychom na toto omezení narazili. Abychom dosáhli vyššího rozlišení, využíváme faktu, že mřížka je tvořena stejnými objekty (například čtvercovými otvory). Je proto možné detekovat polohy těchto objektů nějakou pokročilejší metodou se sub-pixelovým rozlišením, což je v našem případě kros-korelace. Vybereme jeden objekt, a hledáním pomocí korelace nalezneme polohy všech dalších obdobných objektů na povrchu. Výsledné polohy středů objektů v rámci jednoho řádku či sloupce mřížky pak můžeme například proložit přímkou a z ní odečíst požadovaný úhel.

Přesnost postupu založeném na kros-korelaci jsme otestovali jednoduchým numerickým experimen-





Obrázek 3: Simulace schopnosti algoritmu kros-korelace detekovat rotaci vzorku pro různá rozlišení. Vlevo jsou vstupní data (vložený obrázek) a detekované posuny, vpravo závislost délky vektoru posunutí na vzdálenosti od středu obrázku, pro dvě různá rozlišení vstupních dat.

tem. Pomocí nástrojů programu Gwyddion jsme vygenerovali dvě sady topografií mřížek, vzájemně pootočené o 0.1 stupně. Vyhodnotili jsme polohy středů objektů a vynesli jsme do grafu jejich vzájemné posunutí, v závislosti na vzdálenosti od středu otáčení. Data jsme proložili přímkou a určili její směrnici a rozptyl této směrnice. Stejný experiment jsme zopakovali na dvou sadách dat, reprezentujících mřížku o fyzicky stejných rozměrech, nicméně naměřenou s rozlišením 1024×1024 pixelů a 5000×5000 pixelů. Výsledné grafy je možné vidět v obrázku 3. Detekované úhly byly $(0.1007 \pm 0.0004)^{(\circ)}$ pro data s menším rozlišením a $(0.1 \pm 5 \times 10^{-7})^{\circ}$ pro data s vyšším rozlišením. Nejistota v určení úhlů daná procesem zpracování dat je v tomto případě 14 µrad resp. 0.9 µrad, což je pro naše účely dostačující a to dokonce i pro měření v nižším rozlišení (pokud by ostatní nejistoty byly velmi nízké).

Generováná data byla nicméně ideální, taková, jaká by byla naměřena polohovacím systémem, který je dokonale ortogonální a prostý jakýchkoliv driftů. Metrologické mikroskopy využívají odměřování ve všech osách pomocí interferometrických systémů a takovému stavu se snaží co možná nejvíc přibližovat.

Klíčovým přístupem pro získání co možná nejmenší nejistoty při měření úhlu je využití skenovacího systému, který je dostatečně stabilní a prostý parazitních posunů či rotací — to byl také případ měření v PTB. Mohlo by se zdát, že je také nutné, aby i osy mikroskopu byly vůči sobě co možná nejvíc kolmé. Nedokonalost v tomto směru je nicméně možné detekovat velmi jednoduchým postupem - druhým měřením stejné mřížky pootočené o 90 stupňů. To umožňuje odseparovat chybu mřížky od chyby skenovacího systému. I tento postup předpokládá, že skenovací systém se chová stále stejně a právě různé parazitní jevy tento princip narušují. Většina námi vyvinutých systémů využívá k odměřování polohy interferometry, tj. poloha by měla být známa přesně. To nevylučuje, že se může polohovací systém při skenování různě natáčet. Pokud by náš senzor byl umístěn přesně v pomyslném středu svazků interferometrů, měl by být vliv natáčení nulový. V praxi je nicméně velmi obtížné takového stavu s rozumnou přesností dosáhnout a jako výhodnější se jeví parazitní rotace detekovat a kompenzovat.

V rámci řešení úkolu jsme se proto zabývali kompenzací parazitních rotací Státním etalonu délky a tvaru v oblasti nanometrologie (dále pro stručnost nazývaném metrologický SPM) a na systému pro hybridní měření na velkých plochách, který jsme zkonstruovali v minulém roce.



2.1.1. Parazitní rotace při měření na metrologickém SPM

Parazitní rotace souvisí s nedokonalostí lineárních vedení, která by měla zajistit, že je pohyb skenovacího systému v každé ose zcela nezávislý a přímý. V případě metrologického SPM se jedná o polohovací systém Physik Instrumente, který umožňuje pohyb ve všech třech osách a který obsahuje i integrované kapacitní senzory vzdálenosti. Úhel natočemní detekovat ani kompenzovat nedokáže. Úhlové chyby při stolku byly charakterizovány již před mnoha lety kolegy z oddělení kvantové metrologie délky. Pro jednotlivé rotace, s označením os prezentovaném v grafu 4), byly naměřeny maximální hodnoty 12 μ rad pro rotaci označenou jako "pitch", 35 µrad pro rotaci "yaw"a 12 µrad pro rotaci "roll". Pro měření planárních úhlů je kritická především rotace "yaw", jejíž hodnota je poměrně vysoká, pokud ji srovnáme s nejistotou 50 μ rad, které chceme dosáhnout. Proto se již krátce po konstrukci mikroskopu společným úsilím ČMI a ÚPT AVČR začalo uvažovat o instalaci kompenzačního systému.



Obrázek 4: Označení os rotací stolku Physik Instrumente dle měření na oddělení kvantové metrologie délky v roce 2007.

Pro kompenzaci rotací je nejprve nutné tyto rotace s dostatečnou přesností měřit. V případě metrologického SPM je tento úkol usnadněn tím, že v SPM je soustava interferometrů určená pro měření pohybu ve všech šesti stupních volnosti, což zahrnuje i rotace. Umístění interferometrů je schematicky naznačeno v obrázku 5. Metrologický SPM je systém skenující vzorkem, tj. veškerý pohyb mezi hrotem a povrchem v laterálním směru je zajišťován pomocí polohovacího systému Physik Instrumente, jehož parazitní rotace byly diskutovány v minulém odstavci. Při prvním pokusu o aktivní kompenzaci rotací byly mezi polohovací systéma a rovinu vzorku instalovány tři kombinované piezoelektrické prvky, každý umožňující pohyb ve dvou osách, laterálně i v ose z. Tyto prvky byly instalovány zhruba na kružnici naznačenou na obrázku 5, tak, aby laterální pohyb všech třech vyvolal natočení okolo osy odpovídající normále povrchu stolku, tj. rotaci "yaw"z předchozího textu. Nezávislý pohyb jednotlivých piezoelektrických prvků v ose z pak může kompenzovat zbylé dvě rotace. Vyhodnocení všech rotací a jejch kompenzace probíhaly poměrně složitým způsobem. Signály odměřované všemi interferometry detekovala elektronika ÚPT AVČR a délku z nich vyhodnocoval program v Labview běžicí na osobním počítači. Výsledky se předávaly přes průmyslovou sběrnici CAN do dalšího počítače. Na něm bežel program pro kompenzaci, který nastavoval hodnoty napětí pro jednotlivé kompenzační prvky jako analogový výstup na multifunkční kartě National Instruments. Ty byly dále zesilovány pomocí další elektroniky ÚPT AVČR. Celá tato sestava byla testována a schopna udržovat konstantní natočení celého polohovacího systému, jak bylo také prezentováno



v odborném článku [4], z něhož přejímáme jeden z výsledků v obrázku 6.

Experimenty provedené s první verzí aktivní kompenzace byly velmi slibné, ale jednalo se spíše o měření typu "proof of concept", než o sestavu, kterou by bylo možné rutinně využívat, jak se ukázalo posléze. Kombinace všech prvků detekce, zpracování, vyhodnocení a interpretace dat byla poměrně těžkopádná, což mimo jiné ovlivňovalo rychlost kompenzace, cože je ostatně vidět i na nízkém rozlišení snímku demonstrujícícho efektivitu kompenezace v obr. 6. V tomto roce jsme proto celý systém rekonstruovali a značně zjednodušili, jak je to popsáno v dalších odstavcích. Velkou výhodou získanou experimenty z předchozích let bylo, že jsme mohli využít již existující interferometry a kompenzační piezoelektrické prvky.



Obrázek 5: Umístění šesti interferometrů v metrologickém SPM

2.1.2. Nová detekce rotací pomocí interferometrů

Stejně jako v minulých letech je klíčovým prvkem pro detekci rotací v metrologickém SPM využití šesti interferometrů, z nichž jeden odměřuje posuny v ose x, dva posuny v ose y a tři odměřují posuny v ose z. Z jejich výstupů je možno zrekonstruovat pohyb stolku ve všech směrech i rotacích, tj. ve všech šesti stupních volnosti. Surová data z interferometrů nicméně mají k takovým výsledkům velmi daleko. Primárním signálem z interferometru s detekcí fáze jsou dva signály, které je možné si představit jako sinus a kosinus vektoru opisujícího pohyb po kružnici. Velikost této kružnice a její případná nepravidelnost je dána kontrastem interferenčních proužků, fáze odpovídá poloze, tj. její změna o 2π odpovídá posunu pohybujícího se zrcadlo vůči referenčnímu o určitý zlomek vlnové délky - v případě našich interferometrů je to její čtvrtina.

Nové uspořádání detekce posunutí je založeno na využití desky Red Pitaya a přídavné desky s AD a DA převodníky (viz obr. 7), kterou jsme popisovali již ve zprávě ÚTR v předchozích letech při popisu kontroléru pro rastrovací sondovou mikroskopii. Pro účely detekce a kompenzace natočení je na tomto systému podstatné, že obsahuje 16 simultánních analogových vstupů, které můžeme použít pro snímání signálů z interferometrů. Kromě toho obsahuje také 16 analogových výstupů, které můžeme využít na





Obrázek 6: Výsledky pokusů o kompenzci rotací publikované v Ref. [4]: a) úhlová chyba bez kompenzace a s kompenzací, b) reziduální úhlová chyba po kompenzaci rotací.

řízení kompenzací, jak bude zmíněno v dalším textu, v kombinaci s vysokonapěťovým zesilovačem zobrazeným na obr. 7. Analogové vstupy jsou realizovány prostřednictvím dvou osmikanálových 18bitových AD převodníků Od Texas Instruments (ADS8598H), které sdílí jeden signál pro konverzi hodnot. I přesto, že je měření realizováno dvěma převodníky, je díky tomu možné načíst simultánně 16 napětí. Analogové výstupy jsou implementovány pomocí 16kanálového 16bitového DA převodníku DAC81416 rovněž od Texas Instruments. Využili jsme celkem 4 kanály, kdy jeden slouží pro řízení napětí na všech střižných piezech a zajišťuje tak rotaci okolo osy z. Zbylé tři signály nastavují napětí na jednotlivých z piezech a umožňují tedy rotace okolo os x a y. Pro zesílení řídicích signálů jsme využili kompaktní zesilovače firmy Piezodrive. Moduly BD250 mají integrovaný zdroj vysokého napětí a vstupní rozsah 0 V až 3 V zesílí na ± 250 V. Napětí 1,5 V na vstupu udělá nulové napětí na výstupu.

Pro vyhodnocení vzdáleností měřených jednotlivými interferometry využíváme vždy dvojici signálů z příslušné detekční jednotky. Před detekcí signály normalizujeme jejich přepočtem na jednotkovou kružnici s korekcí případné elipticity, která může být dána špatným seřízením interferometru (tzv. Heydemannova korekce, popisovaná ve zprávách ÚTR v minulých letech). Samotná detekce je velmi jednoduchá, fázi detekujeme pomocí funkce arkustangens a průběžně počítáme průchody nulovou fází abychom mohli kumulativně sledovat počet proběhlých interferenčních proužků.

Parazitní rotace detekujeme jako rozdíl mezi vzdálenostmi čtenými interferometry umístěnými ve stejném směru, viz obr. 5, tj. například pro nás kritickou rotaci "yaw"z předchozího textu využijeme rozdíl mezi vzdálenostmi měřenými dvěma interferometry v ose y. Z pohledu rozlišení úhlu je tedy zásadní, aby naše odměřování bylo zatíženo co možná nejmenším šumem, což zahrnuje faktory jako seřízení interferometrů, zesílení a detekci signálů z detekčních jednotek i mechanický šum celého systému.

Na obr. 8 a 9 jsou uvedeny výsledky analýzy šumu v systému pro různé stavy elektroniky. Na obr. 8 je porovnání stavu, kdy je řídicí elektronika vypnutá, se stavem, kdy je kompletně zapnutá. Vynesena jsou měření ze všech interferometrů, které odměřují pohyb stolku, doplněná o měření interferometru, který odměřuje pohyb sondy (Head). Škála na ose *y* byla záměrně nastavena u obou stavů elektroniky stejně, aby byl rozdíl zřejmý již na první pohled. Obecně se dá říct, že nejmenší šum byl pozorován na Head interferometru, zatímco největší byl na *y* interferometrech. Ve vypnutém stavu se šum pohyboval od 0,6 nm u Head interferometru do 1,2 nm u *y*2 interferometru, zatímco v zapnutém stavu to bylo v rozmezí 1,5 nm až 3 nm. Zajímalo nás dále, jaké jsou příspěvky od jednotlivých částí systému. Srovnání je uvedeno na



Zpráva úkolu TR č.	Počet příloh: 4
6014-TR-Z0001-22	Počet listů: 12/39

obr. 9. Pro přehlednost jsou uvedeny pouze výsledky měření na *x* interferometru. Je vidět, že pokud se zapnulo napájení zesilovačů pro kompenzační pieza a drželo se řídicí napětí na konstatních hodnotách, šum v ose *x* to nezvýšilo. Pokud ovšem byla zapnutá jen stabilizace vlnové délky, šum se zvýšil několikanásobně. Podstatné zvýšení šumu také přineslo zapnutí napájení piez, které pohybují stolkem, a případ, kdy kromě piez byla zapnutá i zpětná vazba, která na základě hodnot z integrovaných kapacitních senzorů držela stolek na dané pozici. Z porovnání jde tedy vidět, že systém kompenzace rotace zdaleka není podstatným zdrojem šumu celé sestavy.



Obrázek 7: Fotografie desky s AD a DA převodníky s modulem Red Pitaya (vlevo) a zesilovačů Piezodrive BD250 (vpravo).



Obrázek 8: Porovnání šumu interferometrů v případě, že je veškerá řídicí elektronika vypnutá (vlevo) a v případě, že je kompletně zapnutá (vpravo).

2.1.3. Kompenzace rotací

Jak bylo zmíněné výše, pro kompenzaci rotací využíváme piezoelektrické převodníky umístěné na polohách shodných s polohami interferometrů v ose z na obrázku 5, přičemž pro rotaci typu "yaw"je využito





Obrázek 9: Porovnání jednotlivých příspěvků do šumu na x interferometru. Off - řídicí elektronika je vypnutá, rotation - zapnutá pouze elektronika pro rotace stolku, stabilization - zapnutý pouze systém stabilizace vln. délky, table - zapnuté polohování stolku, table+servo - stolek aktivně drží pozici.

střižné funkce převodníků - jejich pohybu v laterálním směru. Pro kompenzaci využíváme program implementovaný na zařízení Red Pitaya, který pomocí jednoduché zpětné vazby udržuje detekované rotace konstantní tím, že na jednotlivé piezoelektrické prvky posílá napětí. Efektivita kompenzací je znázorněna na obr. 10 a 11, kdy jsme se stolkem popojeli z výchozí nulové polohy o 10 μ m v osách x a y a o 1 μ m v ose z a o stejné vzdálenosti zase zpět. Rychlosti stolku byly 10 μ m/s v laterálním směru a 1 μ m/s ve směru z. V grafech je pro srovnání uveden i případ, že rotace stolku kompenzovány nebyly. Z porovnání těchto závislostí jde vidět, že pokud je kompenzace vypnutá, interferometry v ose y se vzdálí cca o 300 nm od referenční hodnoty nameřené kapacitním senzorem. V případě našeho systému se jedná o rotaci cca 4 μ rad, která byla bezezbytku vykompenzována.

V případě kompenzace rotace okolo osy z byla situace jednoduchá v tom smyslu, že na modulu Red Pitaya byla implementována jednoduchá zpětná vazba pouze s P členem, kdy se napětí na střižných piezech nastavovalo tak, aby se vynuloval rozdíl interferometrů y1 a y2. V případě zbylých dvou os je situace o něco komplikovanější. Možností jak nastavovat rotace pomocí tří piez je hned několik:

- Nulovat rozdíl jednoho z interferometrů oproti průměru spočítaném ze všech tří interferometrů, tj. např. nastavit napětí na piezu z1 tak, aby byl nulový rozdíl z1 (z1 + z2 + z3)/3.
- Nulovat rozdíly sousední dvojice interferometrů, tj. např. napětí na piezu z1 nastavit tak, aby byl nulový rozdíl z1 z2.
- Nulovat v jednom kroku rozdíl interferometrů z^2 a z^3 pomocí napětí na piezech z^2 a z^3 a ve druhém kroku nastavit napětí na piezu z^1 tak, aby byl nulový rozdíl $z^1 (z^2 + z^3)/2$.



Počet listů: 14/39

V aktuální verzi řídicího software je zpětná vazba implementována podle třetí varianty. Důvodem byla názorná geometrická představa, kdy v prvním kroku se stolek otočí okolo osy y a ve druhém pak okolo osy x. Toto ale bude předmětem zkoumání, protože už teď je z obr. 11 (vlevo) patrné, že piezu z1 trvalo výrazně delší dobu otočení stolku podle osy x v porovnání s natáčením podle osy y, kdy se na rotaci podílela pieza dvě. Nicméně kompenzace rotací funguje, jak má. Obrázek 11 vpravo jasně demonstruje, že jsme byli schopní vykompenzovat rozdíl např. mezi piezy z1 a z2, který činil cca 150 nm, což opět odpovídá řádově jednotkám µrad.

Dále nás zajímalo, zda můžeme vykompenzovat natáčení okolo osy z při pohybu stolku na celém svém rozsahu. Nastavili jsme proto výchozí polohu stolku ve všech osách do středu příslušného rozsahu a pak jsme jeli o polovinu rozsahu směrem na začátek, dále přes celý rozsah nakonec, abychom se na závěr opět vrátili do středové polohy. Průběh řídicího napětí na střižných piezech a rozdíl interferometrů y1 a y_2 je uveden na obrázku 10. V grafu je pro porovnání uveden i případ, kdy byla kompenzace vypnutá. Jde vidět, že v tomto případě se interferometry rozešly až o 1,7 μm, to odpovídá rotaci až o 11 μrad. Při zapnuté kompenzaci byl rozdíl mezi interferometry menší než 200 nm, přičemž tato hodnota by jistě šla zmenšit buď zpomalením pohybu stolku nebo zvýšením proporčního členu ve zpětné vazbě. Při hodně vysoké hodnotě ale hrozí, že se zvýší šum. Z grafu je také patrné, že se řídicí napětí během kompenzace pohybovalo v rozmezí 0,5 V až 2,7 V. Vzhledem k tomu, že plný rozsah je nula až tři volty, je zde potenciál pro kompenzaci ještě větších natočení.



Obrázek 10: Kompenzace rotace stolku při současném pohybu ve všech osách. Hodnoty y interferometrů, kdy byla kompenzace vypnutá, byly pro lepší názornost posunuty o 400 nm. Zobrazeno je i měření z kapacitního senzoru samotného stolku (CY).





Obrázek 11: Kompenzace rotace stolku při současném pohybu ve všech osách. Hodnoty z interferometrů byly při vypnuté kompenzaci posunuty o 400 nm. CZ je měření z kapacitního senzoru.



Obrázek 12: Kompenzace rotace stolku při jeho pohybu včetně řídicích napětí. Je uveden i případ, kdy byla kompenzace vypnutá.



2.1.4. Generátor malých rotací

Zařízení pro kompenzaci rotací můžeme použít i jako generátor malých úhlů. Pokud stolek stojí a my pouze nastavujeme různá napětí na piezo ovládající otáčení ve směru "yaw", můžeme pomocí interferometrů detektovat natočení s metrologickou návazností na stabilizovaný Nd:YAG laser, který je součástí státního etalonu. Na obrázku 13 je výsledek takového experimentu. Napětí na střižných piezech bylo nastavováno v rozmezí ± 220 V, čemuž odpovídal rozdíl interferometrů v rozmezí -3,3 µm až +3,2 µm. Převedeno na úhel to dělá celkem 43 μ rad. Celkový rozsah natáčení by tedy pro rozsah napětí ± 250 V měl činit přibližně 49 μ rad. Co se šumu týče, ukázali jsme, že šum interferometru v ose y byl na úrovni 3 nm. Pokud tento údaj opět převedeme na úhel, vyjde nám šum přibližně 20 nrad.



Obrázek 13: Rotace stolku okolo osy z, uveden je rozdíl interferometrů y1 a y2. Rozsah napětí na střižných piezech byl ±220 V.

2.1.5. Senzor autokolimátoru pro přenos návaznosti úhlu

Situace, kdy máme v mikroskopu odměřovací systém detekující všechny stupně volnosti pohybu je poměrně vzácná. V praxi by nás zajímaly také parazitní rotace u jiných polohovacích systémů, včetně komerčních mikroskopů. Kromě toho plánujeme ještě v tomto roce porovnat naše schopnosti měření malých úhlů s přístroji na Oddělení primární metrologie rovinného úhlu a délky ČMI v Liberci a šestiosý odměřovací systém na státním etalonu se pro takové účely nehodí, byť může pracovat i jako generátor malých úhlů. Proto jsme se rozhodli vyrobit malý etalon pro přenos návaznosti v oblasti velmi malých úhlů.

Koncepce etalonu je založena na využití autokolimátoru, což je zařízení, které sleduje posun svazku světla na pozičně citlivém detektoru (PSD), poté co se světlo odrazilo od zrcátka, jehož pootočení chceme detekovat. Předpokládáme že zrcátko bude součástí jiného etalonu generujícího úhel (buď na našem oddělení nebo na OI Liberec), při konstrukci jsme se proto zaměřili jen na optickou a elektronickou část





Obrázek 14: Fotografie DPS pro zpracování signálů z pozičně citlivých detektorů. Vyobrazeny jsou i destičky s předzesilovači.

detektoru. Optická sestava je poměrně jednoduchá, skládající se z laserového modulu, kolimační optiky a děliče svazku. Pro detekci využíváme pozičně citlivý detektor, na nějž dopadá odražený svazek. S ohledem na to, že pozičně citlivý detektor a jeho předzesilovač jsou klíčové pro dosažení potřebné citlivosti a stability, vyvinuli jsme novou elektroniku znázorněnou na obr. 14. Schéma zapojení jak předzesilovačů, tak samotné desky pro analogové zpracování signálů z PSD je v příloze zprávy. Umístění předzesilovačů do těsné blízkosti senzoru je důležitým faktorem pro minimalizaci elektromagnetického rušení. Na obrázku jsou destičky s předzesilovači osazeny kvadrantními detektory Hamamatsu S5980 a S4349, které chceme využít především pro optickou zpětnou vazbu v AFM. Destička AFMTIA02B však přímo umožňuje osadit detektor Hamammatsu S5990, který byl použit v dále uvedených měřeních šumu (obr. 15) a citlivosti (obr. 16). Máme ještě destičku pro detektor S5991 s aktivní plochou o rozměru $9 \times 9 \text{ mm}^2$, což by mělo vést na větší rozsah autokolimátoru. Námi testovaný menší z detektorů má aktivní plochu jen $4 \times 4 \text{ mm}^2$, má ale zase menší šum.

V testovací aparatuře jsme použili zrcátko, které umožňovalo přesné natáčení pomocí piez. Při odhadu rozsahu našeho autokolimátoru jsme použili katalogové údaje výrobce (Thorlabs), který udává rozsah ± 275 µrad pro rozsah 0 V až 150 V na piezech. První experiment spočíval v měření šumu. Nejdříve jsme měřili maximální samplovací rychlostí, co použitá měřicí karta umožňovala - 125 kS/s. Na obr. 15 vlevo jde vidět, že autokolimátor v prvním okamžiku zaznamenal otřes zrcátka při inicializaci analogových výstupů měřicí karty, které byly přivedeny na zesilovač, přes který bylo nastavováno napětí na piezech zrcátka. Dále je vidět, že šum se pohyboval na úrovni 100 mV. Toto jsou data naměřená přímo na výstupech z desky pro analogové dělení signálu. Vzhledem k tomu, že pohyby odměřovaných zrcadel budou v naprosté většině případů pomalé, plánujeme tato napětí filtrovat a řádově tak snížit šumy. Zkusili jsme rovněž dlohodobé měření, kdy jsme během 3hodinového pozorování nezaznamenali větší drift signálů autokolimátoru.

Další experiment spočíval v měření citlivosti, kde jsme zrcátko nastavili do pevné polohy v jedné ose a proskenovali většinu rozsahu v druhé ose. Toto se zopakovalo s prohozenými osami. Pro vzdálenost zrcátka od detektoru 420 mm jsme zjistili, že citlivost dané osy nezávisí na poloze v druhé ose a činí 55 μ rad/V v ose x a 49 μ rad/V v ose y. Rozsahy jsme odhadli na více než 250 μ rad v ose x a více než 220 µrad v ose y. Šumy okolo hodnoty 100 mV by tedy odpovídaly cca 5 µrad, tedy přibližně jedné úhlové





vteřině, ale jak bylo zmíněno dříve, tato hodnota by šla ještě řádově snížit filtrováním signálů.





Obrázek 16: Měření citlivosti v ose X (vlevo) a v ose Y (vpravo) autokolimátoru. Vzdálenost zrcadla od detektoru autokolimátoru byla 420 mm.



2.1.6. Rozšíření kompenzace parazitních rotací pro měření na velkých plochách

Koncepci otestovanou na metrologickém SPM jsme rozšířili i pro druhé zařízení, které v tomto kontextu plánujeme využívat, tj. na systém pro hybridní měření v nanometrologii vyvinutý v roce 2021 v rámci Úkolu technického rozvoje. Jedná se o polohovací systém s velkým rozsahem $(5 \times 5 \text{ cm}^2)$, využívající vzduchová vedení a generování síly pomocí cívek. I tento systém je řízen pomocí interferometrů, základní metrologická návaznost je tedy zajištěna, ale i u tohoto systému se setkáváme s parazitními rotacemi. Ty jsou pozorovatelné i v praxi a projevují se například změnami v kontrastu interferenčních proužků, když je zařízení v pohybu.



Obrázek 17: Fotografie stolku s orientací os.

Narozdíl od metrologického SPM se v systému pro hybridní měření neodměřují všechny stupně volnosti pohybu, ale v každé ose je instalován jen jeden interferometr. Pro detekci úhlů jsme proto využili senzory autokolimátorů popsané v předchozí kapitole. Pro kompenzaci samotnou jsme použili velmi podobné a obdobně uspořádané piezoelektrické prvky jako v metrologickém SPM, viz obr. 17, také elektronika a software pro kompenzaci byly stejné.

První, co nás zajímalo, byl opět šum. Na obr. 18 jsou měření z interferometrů. Na grafu vlevo jsou uvedeny případy kdy byl stolek aktivní a držel polohu, dále když se vypnula zpětná vazba, případ, kdy se zastaví přívod vzduchu do ložisek, nebo když se ještě navíc vypne napájení střižných piez, které zajišťují rotaci horní desky se zrcadly. Pro přehlednost jsou v pravém grafu jenom případy, kdy stál stolek na místě, tj. není zobrazen případ bez zpětné vazby, kdy stolek pomalu odjížděl z původní polohy. Jde vidět, že zdaleka největší příspěvek do šumu měla zpětná vazba, která držela stolek na místě. Parametry zpětné vazby bude tedy potřeba ještě optimalizovat.

Dále jsme se zajímali o natáčení stolku. Na obr. 19 je uvedeno měření na autokolimátoru, který byl umístěn v ose x stolku. Uvedeny jsou opět případy analogicky k měření poloh pomocí interferometrů. Vzdálenost zrcadla od detektoru byla oproti testovací sestavě jenom čtvrtinová a navíc muselo být přenastaveno i zesílení desky pro analogové dělení signálů autokolimátoru. Nemůžeme tedy použít citlivostní faktory odhadnuté dříve. Místo toho ale můžeme rozsah natáčení a příslušné úhly odhadnout z geometrie a rozsahů piez. Vyjde nám v tomto případě citlivost 390 µrad/V. S tímto citlivostním faktorem nám vyjde, že v případě, kdy stolek drží aktivně polohu, se jedná o úhly 40-80 µrad.





Obrázek 18: Šum v poloze odměřované 2D interferometrem v různých stavech stolku. Vpravo jsou vyobrazena pouze měření, kdy stolek držel polohu.



Obrázek 19: Šum odměřovaný autokolimátorem umístěným v ose X stolku, resp. X interferometru. Osa X na autokolimátoru tedy odpovídá rotaci okolo osy Z stolku a osa Y na autokolimátoru odměřuje rotaci okolo osy Y stolku.



Na obr. 20 a 21 jsou měření na autokolimátoru, kdy stolek postupně jede v jedné či druhé ose o různé vzdálenosti. Můžeme vidět, že pokud stolek popojede maximálně o 1 mm, natočení jsou kolem 40 µrad. Zde neuvažujeme otřesy způsobené zakolísáním zpětné vazby. Pokud ovšem jedeme na větší vzdálenost, úhly jsou podstatně větší. Pokud například jedeme podél osy x o 4 cm, napětí na vodorovné ose autokolimátoru postupně narůstá z cca 0,5 V na 3,5 V. Tomu by odpovídala rotace o úhel 1170 µrad.

V dalším jsme se pokusili odhadnout, jakých rotací jsme schopni dosáhnout přivedením napětí na jednotlivá pieza. Na obrázku 22 vlevo je znázorněno měření na autokolimátoru, když se měnilo napětí na střižných piezech. Můžeme odhadnout, že změna napětí o cca 200 mV odpovídá rotaci o 78 µrad. Rozsah střižných piez je ± 200 V, takže teoreticky jsme schopni kompenzovat rotace do 156 µrad, což ovšem zdaleka nepokryje rozsah rotací při pohybu na plném rozsahu stolku. Na obrázku vpravo je případ, kdy se přivádělo napětí na piezo z1. Zde jde vidět, že k žádnému natáčení nedochází, což vyplývá z geometrie systému a je tedy v souladu s očekáváním. Analogicky odhadujeme rozsah natáčení při změně napětí na piezech z2 a z3. Rozsah možných natočení odpovídá v obou případech 156 µrad.



Obrázek 20: Odměřování rotací stolku při jeho pohybu pomocí autokolimátoru. Vlevo pohyb stolku o 100 μm, resp. o 1 mm v ose X, vpravo pohyb v ose Y.



Obrázek 21: Odměřování rotací stolku autokolimátorem. Vlevo pohyb stolku o 1 cm, resp. o 4 cm v ose X, vpravo pohyb v ose Y.





Obrázek 22: Autokolimátorem změřené natáčení stolku pomocí střižného (vlevo) a Z1 pieza (vpravo).



Obrázek 23: Vlevo je autokolimátorem naměřené natáčení stolku při změně napětí na piezu Z2, vpravo jsou data z autokolimátoru při změně napětí na piezu Z3.



2.2. Měření periody

Měření periody je jednou z typických úloh v oblasti metrologie délky a to i mimo oblasti mikro- a nanotechnologií. Periodické struktury jsou využívány v řadě oblastí, ať už se jedná o konečné produkty, nebo o součásti metrologických zařízení. Typickými příklady jsou difrakční mřížky, senzorové matice různých detektorů (CCD, CMOS) či fotonické krystaly. Periodické struktury se také využívají jako pomocné prvky ve velmi pokročilých výrobních technologiích, například pro sesazování masek v polovodičovém průmyslu. S ohledem na to, že řada periodických struktur má rozměry ideální pro charakterizaci pomocí mikroskopie atomárních sil se jedná o jedno z nejčastějších měření i z pohledu požadavků zákazníků. Ilustraci různých periodických struktur se kterými se v tomto oboru potkáváme můžeme vidět na obrázku 24. V oblasti metrologie je zdaleka nejčastějeím typem vzorku difrakční mřížka, která se kromě své primární funkce používá také jako etalon pro přenos návaznosti mezi různými typy mikroskopů, včetně optických či elektronových.

Periodu nejjednodušeji získáme naměřením topografie povrchu zkoumaného vzorku pomocí AFM, následným výběrem směru ve kterém má být perioda vyhodnocena a manuálním vyhodnocením poloh jednotlivých útvarů tvořících periodickou strukturu na profilu získaném v tomto směru. Již takový postup s sebou přináší různé otázky, například: je lepší naměřit více period s nižším prostorovým rozlišením, nebo ve stejném čase méně period s vyšším prostorovým rozlišením? Je lepší vyhodnotit data z hran struktur, nebo spočítat nějakou statistickou funkci reprezentující vlastnosti vzorku v laterálním směru, například autokorelační funkci a výsledek získat z ní? Na tyto otázky by mohl odpovědět příslušný normativní dokument, například norma ISO 11952. Ta však bohužel o této problematice mluví jen velmi kuse - je v ní uvedeno několik metod vyhodnocení a doporučení, aby se vyhodnocovalo více než 5 period. Odpověď proto musíme nalézt sami. To se podařilo především díky spolupráci s Mgr. Davidem Nečasem, který se otázkám statistických vlastností struktur na površích pevných látek věnuje dlouhodobě. V následujícím textu prezentujeme hlavní přístupy, které byly ve společné studii použity a hlavní výsledky, které se nějakým způsobem vztahují k měření na velkých plochách, které je náplní tohoto projektu. Jako příloha této zprávy je uveden rukopis článku, toho času v recenzním řízení, který jde daleko za rámec problematiky tohoto úkolu technického rozvoje a podává podrobnou analýzu chování různých metod vyhodnocení periody, včetně řady praktických doporučení jak stran jejich použití tak stran jejich implementace.

Prezentovaná analýza chování různých metod vyhodnocení periody je založena na masivních numerických simulacích na syntetických datech povrchů. To má dobrý důvod: pokud data generujeme, známe všechny jejich parametry, tj. víme, jaký má být výsledek vyhodnocení periody. Můžeme také konzistentně přidávat různé chybové vlivy abychom zjistili, který z nich má na efektivitu zkoumaných algoritmů největší vliv. Vytvářením metod pro generování simulovaných dat v Gwydddionu se věnujeme dlouhodobě (viz např. Ref. [5]). Generované povrchy přitom mohou být velmi blízké tomu, co mikroskop změří - i všechny struktury prezentované v obrázku 24 byly získány jen tímto způsobem a nereprezentují žádné reálné měření.

Metod pro vyhodnocení periody je celá řada a nacházíme je nejen v normě ISO 11952, ale i v řadě odborných publikací. Můžeme je rozdělit do různých kategorií. **Měření v přímém prostoru** je patrně nějintuitivnější postup a spadá do něj i výše uvedený nejjednodušší přístup - odečíst dvě polohy hran objektů tvořících periodickou strukturu. Zkoumané metody byly sofistikovanější, konkrétně tyto:

- **Těžiště** (Gravity centre, GC) měřený profil je pomocí prahové hodnoty rozdělen na horní a dolní oblasti. Pro každou oblast je vyhdodnocena poloha těžistě, viz obr. 25c a polohy všech těžišť jsou pak fitovány lineární funkcí.
- **Průchod nulou** (Zero crossing, ZC) hledáme polohy křížení profilu s prahovou hodnotou (která opět profil rozděluje zhruba v polovině na horní a dolní oblast), viz obr. 25d a výsledky fitujeme lineární funkcí.



Obrázek 24: Syntetická data reprezentující různé typické periodické struktury měřené pomocí AFM: (a) 1D difrakční mřížka, (b) 2D difrakční mřížka, (c) pole křemíkových nanodrátků, (d) nedokonalá 1D mřížka s velmi vysokým rozlišením, (e) atomární mřížka, (f) samouspořádané polystyrenové částice využívané při měření čítačů částic.

Další skupina metod je založena **využití Fourierovy tranformace**, kterou v tomto kontextu známe například z minulých zpráv ÚTR, kde byla využívána pro vyhodnocení drsnosti prostřednictvím spektrální hustoty prostorových frekvencí. Z naměřených dat topografie je spočtena 1D nebo 2D Fourierova transformace a poloha maxima odpovídajícího prostorových frekvencím hledané periodické struktury je vyhodnocena různými způsoby:

- Základní FFT (Simple FFT) nejjednodušší přístup, kdy je přímo odečtena poloha ve kterém je spektrální hustota prostorových frekvencí spočtená pomocí FFT největší, viz obr. 25a.
- **Dai05 FT** zpřesněná metoda [6,7], umožňující výpočet prvků ve spektru i pro necelá čísla indexů, tj. hodnot mezi indexy, které jsou prostým výsledkem FFT.
- **Zoom FFT** obdobné zpřesnění, počítající koeficienty Fourierovy transformace pro neceločíselné indexy s využitím Bluesteinova algoritmu [8] umožňujícího zásadní zrychlení výpočtu.

Obdobnou statistickou veličinou jako je spektrální hustota prostorových frekvencí je **autokorelační funkce** - i tu jsme v minulosti využívali pro analýzu statistických vlastností drsných povrchů. I z jejích maxim je možné určit periodu různými způsoby:

- **Základní ACF** (Simple ACF) nejjednodušší přístup, kdy hledáme prostorovu frekvenci odpovídající prvnímu maximum (nikoliv v nule, kde je maximum vždy), viz obr. 25b.
- Více maxim z ACF (Multi-peak ACF) pro zpřesnění výsledku jsou využita i další maxima, která se v autokorelační funkci periodicky opakují. Jejich polohy jsou fitovány lineární funkcí.





Obrázek 25: Ilustrace principů různých metod vyhodnocení periody z měřené topografie.

Konečně můžeme periodický povrch **fitovat** nějakou vhodnou periodickou funkcí, jako je to naznačeno v obr. 25e. Pro tytou účely využíváme metodu nejmenších čtverců a perioda je jedním z parametrů fitu. Podrobnější popis všech algoritmů včetně jeho matematických základů a řady dalších zdrojů v literatuře je uveden v rukopisu v příloze.

Cílem analýzy chování výše uvedených metod bylo zjistit jaká je jejich přesnost a robustnost, tj. jak se chovají pro různé typy vstupních dat. To zahrnovalo jak data o s různým počtem "měřených" period, tak data s různými chybovými vlivy (šum, vlnitost povrchu, konvoluce hrotu, apod.). Ukázalo se, že hlavní trendy v chování jednotlivých metod jsou si velmi podobné i pro různé typy chybových vlivů, proto zde vybíráme jen dva příklady uvedené v obrázku 26 pro data zahrnující vliv vlnitosti a šumu. Pozitivním zjištěním bylo, že většina pokročilejších metod se chová velmi dobře, zejména v kontrastu s jejich "základními" variantami. Nicméně, ukázalo se také, že pro plné využití potenciálu metod je nutné měřit co možná nejdelší profily, zahrnující desítky až stovky period, což je podstatně více, než doporučují normativní dokumenty.



Obrázek 26: Příklad určené přesnosti různých metod vyhodnocení periody pro dva modelové defekty v datech.

S ohledem na to, že se (nejen) v této zprávě snažíme využívat systémy s velkým rozsahem, zajímalo



Počet listů: 26/39

nás jak se metody vyhodnocení chovají i pro různý poměr počtu bodů vůči periodě. Jak u námi konstruovaných skenovacích systémů tak u komerčních zařízení pro nanometrologii (jako je například zařízení NMM1 zmíněné v úvodu zprávy) je totiž možné získávat v rámci každého profilu desítky až stovky tisíc datových bodů. Také je, už z podstaty skenovacích systémů s velkým rozsahem, možné měřit velké množství period. Porovnání různých kombinací počtu bodů a počtu period (délek profilu) je uvedeno v obrázku 27. Z obrázku je patrné, že až do chvíle, kdy je počet bodů na jednu periodu velmi malý (cca 20) je vždy lepší měřit delší profil, než jen přidávat body.



Obrázek 27: Výsledky studie přesnosti různých metod vyhodnocení periody pro různé poměry počtu měřených bodů vůči periodě.

Z grafů je patrné, že "dobré" metody vyhodnocení periody mají velmi dobré škálování přesnosti s ohledem na počet měřených period (tj. zároveň délku profilu). Abychom se přesvědčili, že stejný efekt budeme pozorovat také na měřených datech, případně abychom odhalili, čím mohou být měřená data odlišná od simulací, využili jsme nepřesnější dostupný skenovací systém, zařízení NMM na Ceitec VUT. Pro měření jsme využili bezkontaktní Akiyama sondu a elektroniku pro její řízení vyvíjenou v rámci jiného grantového projektu. Měření proběhlo na testovacím vzorku od firmy Bruker, což je 2D difrakční



Počet listů: 27/39

mřížka. Profily tvořící AFM snímek byly snímány po celé šířce mřížky, abychom postihli co možná největší množství period. Výsledná data jsme zpracovali způsobem, který byl obdobný analýze přesnosti metod vyhodnocování. Rozdělili jsme je na různě dlouhé profily (vždy větší množství, náhodně vybírané) a vyhodnotili jsme rozptyl výsledků v rámci každé konkrétní délky profilu (v tomto případě nebylo možné srovnávat výsledek se "správnou"hodnotu, neboť ta není známa, ale je stále možné sledovat jak se s délku profilu snižuje rozptyl vyhodnocených hodnot). To by mělo poskytovat v principu stejné škálování přesnosti, jako analýza na simulovaných datech. Ve výsledném grafu, viz obr. 28 se ukázaly překvapivě odchylky od očekávaného trendu u velmi dlouhých profilů. Další analýzou dat jsme zjistili, že tento jev je patrně dán drobnými periodickými nepřesnostmi mřížky vzniklými při její litografické výrobě (tzv. stitiching error), které mají za důsledek periodické odchylky roztečí struktur tvořících mřížku. Tento jev bylo možné v měřených datech pozorovat, viz obr. 29b, byť je tato periodická chyba velmi malá, jen jednotky nanometrů. Simulací celého jevu jsme také byli schopni potvrdit, že tento jev může způsobit námi detekovanou odchylku ve škálování přesnosti (viz obr. 29c), kterou by bylo nicméně možné potlačit volbou ještě delších profilů (viz obr. 29d), nejedná se tedy vyloženě o porušení odhadované závislosti přesnosti na délce profilu.



Obrázek 28: Škálování rozptylu vyhodnocených period na reálné mřížce pro různě dlouhé profily, tj. různý počet period v profilu.

Výše uvedený experiment nám dále potvrdil význam měření na velkých plochách. Pokud bychom prováděli měření na náhodných místech mřížky v rozsahu typickém pro komerční mikroskopy (např. 100×100 mikrometrů), patrně bychom získávali různé hodnoty periody a nedokázali odlišit, jestli se jedná o statistickou nehomogenitu mřížky, nebo nejistoty související s chováním mikroskopu. Měření velmi dlouhých profilů nám poskytuje možnost provést analýzu homogenity mřížky jako v obr. (viz obr. 29b, což je v praxi velmi cenné.

S podrobnou analýzou různých metod vyhodnocení periody se může čtenář seznámit v přiloženém rukopisu, zde již pouze shrnujeme hlavní závěry, neboť ty jsou klíčové pro naše úvahy o nalezení nejvhodnější metodiky měření a vyhodnocení periody:

 Přesnost pokročilejších metod vyhodnocení se zvyšuje super-lineárně s počtem měřených period, typicky s koeficientem $P^{3/2}$ pro 1D měření a P^2 pro 2D měření.



Obrázek 29: Výsledky měření periody mřížky od firmy Bruker: nehomogenita periody napříč mřížkou, její vliv na měřená data a srovnání s numerickým modelem: (a) jeden z měřených profilů, (b) lokální perioda vyhodnocená z více profilů, osa x odpovídá celé šířce mřížky, každý bod reprezentuje 30 period, (c) simulované chování vyhodnocvacích metod v případě přítomnosti či absence periodických chyb, (d) totéž na větším simulovaném rozsahu měřené plochy (nad rozsah reálné velikosti této konkrétní mřížky).

- Všechny metody mají sub-pixelovou přesnost, z pohledu přesnosti není tedy nutné aby každý z motivů tvořících periodickou strukturu byl měřen velkým počtem pixelů.
- Přestože norma ISO doporučuje měření více než 5-7 pixelů na periodu, ideálně by to mělo být alespoň 20.
- Již naměřená data nemá smysl resamplovat na vyšší rozlišení.
- Metody založená na vyhodnocení tězišť a průsečíků s prahovou hodnotu mohou získat také odhad nejistoty typu A.
- Pokud dodržíme výše uvedená doporučení, dá se považovat nejistota související se zpracováním dat za zanedbatelnou s ohledem na jiné nejistoty vyskytující se v AFM.



2.3. Termometrie na základě Johnsonova šumu (JNT)

Část prací navazuje na úkol řešený v minulém ÚTR, který spočíval v realizaci primárního teploměru využívajícího termální napěťový šum rezistoru. Lze ukázat, že střední hodnotu šumového napětí lze vypočítat ze základních fyzikálních principů:

$$U_R = \sqrt{4kTR\Delta f}$$

přičemž k je Boltzmannova konstanta, T je absolutní teplota, R je odpor rezistoru a Δf je šířka frekvenčního pásma. Podrobnější rozbor je uveden v minulé závěrečné zprávě, přičemž měřící elektronika byla v mnoha směrech vylepšena.

2.3.1. Nová verze zesilovačů

Elektronika popsaná v předešlém ÚTR měla zásadní nevýhodu v tom, že z nějakého důvodu nebylo možné realizovat dvoukanálové měření, protože oba zesilovače se vzájemně ovlivňovaly. Při zjišťování příčiny vyšlo najevo, že jeden ze zesilovačů je vážným způsobem poškozen, a z naměřených dat nebylo možné odhalit, kdy k poškození došlo. Zesilovač byl navržen jako čtyřstupňový a na základě předchozích zkušeností lze soudit, že problematický je nejspíše první stupeň, zatímco ostatní stupně, včetně napájecí části, mnohokrát pracovaly bez problémů, a proto není důvod je měnit.

Byla navržena a vyrobena nová verze zesilovačů. První stupeň již nespoléhá na integrovaný obvod, ale je sestaven z diskrétních komponent. Schéma je uvedeno na obrázku 30.



Obrázek 30: Schéma představuje čtyřstupňový zesilovač, přičemž první stupeň (předzesilovač) je tvořen diskrétními JFET tranzistory JFE 2140, což představuje hlavní rozdíl ve srovnání s předchozími verzemi zesilovače.

První stupeň je řešen jako diferenční zesilovač, jehož hlavní zesilující prvek je JFET tranzistor JFE 2140, který je dodáván jako dvojice tranzistorů ve společném pouzdře. Pro snížení šumu byly zapojeny



Počet listů: 30/39

dva tranzistory paralelně. Pokud by byl šum zesilovače stále příliš vysoký, je možné počet tranzistorů zvýšit, což je výhoda diskrétního řešení. Diferenčním zesilovačem protéká konstantní proud, který je řízen tranzistorem Q3 (BC817). Jeho stabilitu zlepšuje červená LED dioda D2, která má podobnou teplotní charakteristiku jako PN přechod báze-emitor, čímž představuje teplotní kompenzaci. Obdobnou funkci má dioda D1, která napomáhá udržovat pracovní bod tranzistorům Q5A a Q5B, což je dvojice DMMT3904 dodávaná v jednom pouzdře. Výstup diferenčního zesilovače má velmi výrazný DC offset, a pro jeho kompenzaci je druhý stupeň oddělen horní propustí tvořenou kondenzátory C18 a C19. Druhý stupeň je tvořen přístrojovým zesilovačem AD8429, dále následuje ADA4898 a poté dolní propust (filtr) LTC1564, stejně jako v předešlých variantách.

Při zprovoznění zesilovačů se ukázalo, že výstupní signál má příliš velký offset, který neodpovídal teoretické předpovědi založené na specifikacích jednotlivých operačních zesilovačů. Navíc, pro prevenci offsetu jsou v obvodu začleněny oddělovací kondenzátory tvořící horní propusti. Jedná se o C9, C18 a C19. Vznik offsetu pravděpodobně souvisí s vysokofrekvenčními oscilacemi, které mohou nastat u prvního stupně s FET tranzistory v případě, že vzniknou podmínky pro vznik Colpittsova oscilátoru, což se projeví především při zkratu vstupů zesilovače. Oscilace se nepodařilo přímo změřit, ale jejich možný vznik též potvrzuje publikace [9], kde autoři popisují, že oscilacím lze zabránit rezistorem zapojeným před vstup zesilovače anebo pomocí tlumivky s feritovým jádrem. Skutečně se potvrdilo, že je vhodné před vstupy zesilovače zapojit rezistor, který oscilace utlumí, ale který současně přispívá svým šumem k celkovému šumu zesilovače. Některé vlastnosti zesilovače se měří při zkratovaných vstupech, což však není možné z důvodu oscilací. Předřazené rezistory $R = 10 \Omega$ je tedy nutno chápat jako nedílnou součást zesilovače. Nejsou uvedeny ve schématu, protože se jedná o dodatečnou úpravu. Při běžném měření nejsou potřeba, protože jejich funkci převezme rezistor, jenž představuje teplotní senzor.

2.3.2. Předběžné testy

Mezi důležité parametry zesilovače patří především jeho zesílení A a úroveň vlastního vstupního šumu $U_{\rm n}$. Oba tyto parametry lze změřit alespoň orientačně, přestože oba mají silnou frekvenční závislost. Pro velmi hrubé přiblížení předpokládejme, že šířka pásma je dána schopností měřící karty PCI-6143, která vzorkuje s frekvencí 250 kHz, přičemž šířka pásma je poloviční, tj. $\Delta f = 125$ kHz. Dále uvažujme, že zesilovač je testován ve dvou režimech. Buď jsou na vstupy připojeny snímací rezistory $R = 1000 \Omega$, anebo jsou vstupy zkratovány. Při zkratu platí, že střední hodnota výstupního šumového napětí $U_{\rm s}$ je rovna

$$U_{\rm s} = AU_{\rm n}$$

přičemž zapojíme-li na vstup snímací rezistory, pak se jejich šum $U_{\rm R}$ geometricky sečte se vstupním šumem zesilovače a na výstupu naměříme šumové napětí U:

$$U = A\sqrt{U_{\rm n}^2 + U_{\rm R}^2}$$

Proudový vstupní šum zesilovače zanedbáváme, protože tranzistory JFET mají mimořádně velký vstupní odpor, a právě z toho důvodu byly zvoleny. Z výše uvedených dvou rovnic můžeme zjistit vlastní šum zesilovače i jeho zesílení, protože známe šumové napětí rezistorů při teplotě 23 °C

$$U_R = \sqrt{4kTR\Delta f} \doteq 1.7\,\mu\text{V}$$

Dále byla naměřena střední hodnota výstupního šumového napětí při zkratu $U_{\rm s} = 128 \,\mathrm{mV}$ a také šum při připojených rezistorech U = 562 mV. Zesílení zesilovače vychází

$$A = \frac{\sqrt{U^2 + U_s^2}}{U_{\rm R}} \doteq 322\,000$$



a jeho šum

$$U_n = \frac{U_s U_R}{\sqrt{U^2 + U_s^2}} \doteq 397 \,\mathrm{nV}$$

Snadno vypočteme hustotu šumového napětí, která vychází $1,12 \text{ nV}/\sqrt{\text{Hz}}$, přičemž tak nízkou hodnotu nemá pravděpodobně žádný z komerčně dostupných integrovaných zesilovačů s JFET vstupy.

Zesílení a šum jsou sice základními parametry každého zesilovače, ale pro praktické účely z nich vyplývají vlastnosti, které můžeme stanovit přímo. Hlavním důsledkem je využití napěťového rozsahu měřící karty a rychlost měření. Výstupní signál musí mít dostatečnou úroveň, přičemž byly naměřeny špičkové hodnoty šumu ± 2.7 V. Rozsah měřící karty je ± 5 V, takže zesílení je velmi dobře zvoleno a lze počítat i se zvýšenou teplotou rezistorů. Z hlediska statistiky a pravděpodobnosti je zajímavé uvést, že špičková hodnota vychází přibližně 5× vyšší než efektivní hodnota.

Šum zesilovače má vliv na rychlost měření, ale vzhledem k tomu, že šum snímacího odporu je výrazně vyšší, nemá valný přínos ještě dále optimalizovat zesilovač. Z měření vychází, že při laboratorní teplotě postačí dvě minuty měření k tomu, aby střední kvadratická chyba klesla na 100 mK, a deset minut pro chybu 50 mK.

Uvedené výsledky byly zjištěny z jednoho zesilovače. Další test měl zjistit, zda mohou souběžně fungovat zesilovače dva a zda se vzájemně neovlivňují, jestliže oba měří tentýž šumový signál ze snímacího rezistoru. Byl proto proveden test, při kterém jsou oba zesilovače v provozu, ale každý zesilovač má svůj vlastní snímací rezistor. Oba snímací rezistory mají stejnou hodnotu, jsou na stejné teplotě, jejich šum má stejnou napěťovou úroveň, ale jejich šum není korelovaný. Každý kanál měří teplotu zvlášť, ale jejich korelace (střední hodnota součinu) by měla vycházet nulová. Tento test odhalí, zda do obou zesilovačů nevstupuje parazitní společný signál, ať už by vznikal společnou oscilací anebo by přicházel z vnějšího prostředí a pronikal nedostatečným stíněním.

Výsledky měření ukázaly, že teplota vypočtená z korelace obou kanálů skutečně klesá k absolutní nule, ale neustálí se. Přibližně po deseti minutách klesne na 100 mK a přibližně po jedné hodině měření vykazuje náhodný charakter, ale nepřesáhne 15 mK. V této souvislosti je vhodné uvést důležitý poznatek, že samotná měřící karta způsobovala v minulosti vznik souhlasného signálu, který vedl k chybě 0,2 stupňů, což silně degradovalo celou měřící metodu. Výsledkem byly marné snahy o vylepšování zesilovačů, což nevedlo k cíli. Příčina na straně měřící karty byla poněkud nečekaná. Situace se velmi zlepšila přepojením na jiné kanály. Ukázalo se, že kanály 0 a 1 mají výrazně vyšší přeslechy než kanály 0 a 7.

Určitý parazitní vliv tedy stále lze naměřit a jistě je zde prostor pro zlepšení, ale v tomto stavu by aparatura již mohla být připravena pro další testy v termální lázni.

Teplotní závislost zesilovačů a jejich kalibrace 2.3.3.

U všech použitých elektronických komponent je nutno počítat s tím, že jejich parametry závisejí na teplotě obvodu. Bývá obvyklé, že měření nezačínáme ihned po zapnutí měřícího zařízení, ale ponecháme aparaturu dostatečnou dobu teplotně stabilizovat. Tato metoda může být dostačující pro operační zesilovače se zpětnou vazbou, jejichž zesílení závisí především na vnějších rezistorech, které lze vybrat na základě jejich teplotního koeficientu. Příznivá situace nastane, závisí-li zesílení na poměru dvou rezistorů. Oba rezistory reagují na teplotu, ale když je jejich teplotní koeficient stejný a zajistíme jejich společnou teplotu, pak je jejich poměr velmi dobře stabilní.

Obtížně lze odhadnout teplotní závislost kondenzátorů, které ovlivňují frekvenční odezvu zesilovače. Ve všech případech jsou zvoleny kondenzátory s velmi nízkým teplotním koeficientem (COG a NPO) anebo X7R.

Velmi problematický z hlediska teploty je zejména předzesilovač. Jeho hlavní prvky jsou čtyři diskrétní JFET tranzistory tvořící zesilovač bez zpětné vazby. Zesílení je tudíž dáno pouze charakteristikami



Počet listů: 32/39

tranzistorů a jejich pracovním bodem, přičemž obojí je teplotně závislé.

Kromě zesílení a frekvenční charakteristiky bude na teplotě záviset též vlastní šum zesilovače, čemuž se nelze vyhnout ani při nejlepším návrhu obvodu či výběru jeho komponent.

Všechny zmiňované faktory vedly v předešlých verzích obvodu k tomu, že teplota elektroniky byla regulována termostatem a jednotlivé integrované obvody byly chlazeny společným hliníkovým blokem, který zároveň zajišťoval rozvod tepla. Tato úprava vedla k výraznému zlepšení, ale přesto není jisté, zda bylo toto řešení dostatečné, a navíc vzniklo podezření, že samotný termostat je zdrojem rušení.

V nové verzi teploměru s diskrétními FET tranzistory se již nespoléhá na stabilní teplotu elektroniky, ale raději na průběžnou kalibraci pomocí šumového generátoru, který představuje referenční vstupní signál. Ukazuje se, že teplota tohoto obvodu má skutečně výrazný vliv. Ihned po zapnutí je vypočtená teplota snímacího rezistoru o 10 stupňů vyšší než po ustálení. Po zapnutí se teplota obvodu nepochybně zvyšuje, zatímco výstupní šum se snižuje. Tento paradox lze vysvětlit například tím, že zesílení zesilovače narůstá s teplotou. Namísto zjišťování konkrétních příčin se raději soustřeďme na kalibrační obvod, který zmiňovaný problém vyřeší.

Mikropočítač ATmega (Arduino) byl naprogramován tak, aby na jednom ze svých výstupních pinů generoval MLS (maximum length sequence) binární sekvenci, která má tu vlastnost, že obsahuje pouze dvě napěťové úrovně a její diskrétní fourierova transformace má všechny koeficienty konstantní. Frekvenční spektrum tedy připomíná hřeben. Rychlost generování jednotlivých bitů a také celková délka opakující se sekvence jsou parametry, které zvolíme tak, aby frekvenční pásmo odpovídalo pásmu, které odpovídá schopnostem měřící karty.



Obrázek 31: Záznam binárního šumu pořízený měřící kartou. Napětí dosahuje pouze dvou diskrétních hodnot, a to 0 V a 4,096 V. Napěťová úroveň je odvozena z napěťové reference ADR4540. Binární šum obsahuje frekvence v celém potřebném frekvenčním rozsahu.

Pro testovací účely byl generátor šumu připojen přímo k měřící kartě, která zaznamenala signál patrný z obr. 31. Samotný zesilovač vůbec nebyl použit, aby se zjistily základní vlastnosti kalibrace, protože stabilitu i šum je potřeba zjistit u všech komponent. Data zpracovával stejný software, který je určen k měření teploty, avšak přepočetní konstanty byly upraveny tak, aby program vyhodnotil teplotu přibližně jako 300 K, což odpovídá 26,85 °C. Po přibližně 45 minutách střední kvadratická chyba klesla na 3 mK a již dále neklesala. Jedná se o příznivý výsledek, byť integrační doba je poněkud dlouhá. Pro ušetření času na úkor přesnosti je vhodné uvést ještě další údaj, kdy po čtyřech minutách již chyba klesla pod 10 mK.

Spolupráce s oddělením 1012 2.3.4.

Předběžně byla konzultována spolupráce s oddělením 1012 – oddělení Primární metrologie tepelnětechnických veličin. Popsaný teploměr pracující na základě Johnsonova šumu (JNT) by mohl být zapůj-



Zpráva úkolu TR č.	Počet příloh: 4
6014-TR-Z0001-22	Počet listů: 33/39

čen pro porovnání s navrhnovanou sestavou primárního akustického plynového teploměru (AGT). Jímky pro snímače byly navrženy s ohledem na rozměry stonku JNT (průměr stonku je 10 mm a délka zhruba 700 mm. Na obr. 32 je vidět detail řezu jímkou. Po realizaci nezbytných prvků bude možné dlouhodobě porovnávat měření AGT a JNT v rozsahu asi $(-40 \text{ až } 160)^{\circ}$ C. Do té doby lze tyto dva systémy porovnat v klimatické kalibrační komoře v rozsahu asi $(-10 \text{ až } 50)^{\circ}$ C.



Obrázek 32: Detail jímky pro snímače teploty v řezu



Počet listů: 34/39

3. DOSAŽENÉ VÝSLEDKY

Většina výsledků řešení úkolu je prezentována v předchozí kapitole. V oblasti hardware a elektroniky se jedná o soubor prvků pro kompenzaci rotace polohovacího systému na různých mikroskopech a o senzor úhlu, který může být využit jako prvek pro přenos metrologické návaznosti úhlu do dalších mikroskopických systémů (které například nemají schopnost kompenzace natáčení). V oblasti teoretické se o různé výsledky metodického charakteru, které nám umožňují modifikovat stávající metodiku pro měření periody a vytvořit novou metodiku pro měření úhlu na planárních strukturách, což proběhne na konci roku 2022, po skončení aktivit v rámci tohoto úkolu technického rozvoje.

V této kapitole proto podáváme souhrn doporučení, které budou základem takových metodických postupů, poté co budou validovány dalšími měřeními na různých nám dostupných mikroskopech.

3.1. Měření úhlu

Pro měření úhlu používáme metrologický mikroskop, tj. zařízení s odměřováním polohy pomocí interferometrů. V systému musí být využita aktivní kompenzace parazitních rotací. Postup měření a následné zpracování je následující:

- 1. Provedeme orientační měření vzorku a poté natočíme vzorek tak, aby byl natočen ve směru os mikroskopu, v případě potřeby proces opakujeme.
- 2. Odhadneme drift mikroskopu některou jednoduchou metodou, např. měřením dvou po sobě následujících snímků, jednoho směrem shora dolů a druhého zespodu nahoru, s vyhodnocením vzájemných posunů objektů na snímcích. Podle velikosti driftu stanovíme maximální dobu dalších měření, případně necháme vzorek dále mechanicky a teplotně ustálit. Zbytkovou hodnotu driftu můžeme využít následně při analýze nejistot.
- 3. Zvolíme velikost skenů a počet pixelů, přičemž se snažíme abychom zahrnuli alespoň 10 period, počet pixelů v rychlé ose měření byl alespoň 1000 a pixely byly čtvercové, tj. poměr mezi velikostí strany a počtem pixelů v daném směru byl stejný.
- 4. Provedeme měření ve dvou obdélníkových oblastech, tak aby byla zahrnuta jedna řada a následně jeden sloupec zkoumané periodické struktury, rychlá osa měření je vždy orientovaná ve směru delšího rozměru obdélníku, tj. typicky jednou ve směru osy x a jednou ve směru osy y.
- 5. Naměřená data zpracováváme v programu Gwyddion. V datech označíme jednu oblast tvořící periodickou strukturu vzorku, tj. například jeden otvor difrakční mřížky a pomocí kros-korelace nalezeneme polohy všech oblastí se subpixelovou přesností. Polohy nafitujeme linární závislostí, totéž provedeme s oběma soubory. Výsledný úhel určíme jako úhel mezi oběma přímkami. Při stanovení nejistoty využijeme také chybu parametrů získaných z fitování.

3.2. Měření periody

Měření periody provádíme vždy ve směru rychlé osy měření mikroskopu a problém tak efektivně převádíme na měření periody 1D mřížky i v případě že periodická struktura má dva směry, které je nutné vyhodnotit. Na rozdíl od měření úhlu nás v tomto případě nezajímá vztah mezi oběma směry ani v prípadě 2D mřížky. Citlivost měření na neortogonalitu skeneru, drift a obdobné jevy, které mohou ovlivnit měření úhlů je proto nižší. Postup měření a následné zpracování je následující:

1. Provedeme orientační měření vzorku a poté zorientujeme vzorek aby byl natočen ve směru os mikroskopu, v případě potřeby proces opakujeme.



- 2. Zvolíme velikost skenů a počet pixelů, přičemž se snažíme abychom zahrnuli alespoň 20 period a aby počet pixelů v jedné periodě byl alespoň 30. Pokud je to možné, upřednostníme delší profil (tj. více period) nad větším počtem pixelů v rámci jedné periody.
- 3. Měření provedeme podél rychlé osy mikroskopu. Není nutné aby pixely byl čtvercové, tj. jednotlivé profily v rychlé ose mohou být od sebe vzdáleny více, než je samplování. Hlavním smyslem získávání více vůči sobě posunutých profilů je zpřesnění odhadu natočení mřížky, nikoliv určení periody.
- 4. Data zpracujeme některou pokročilejší technikou (např. těžiště, multi-peak ACF, refined FFT) v programu Gwyddion, například nástrojem "Lattice".
- 5. Pokud se škálování rozptylu dat s délkou profilu vymyká předpokládané závislosti, zkontrolujeme homogennost periody v rámci profilu jeho rozdělením na pravidelné úseky, ze kterých periodu vyhodnotíme.

3.3. Komplexní charakterizace mřížky

V případě, že je poždavkem provést komplexní charakterizaci parametrů mřížky, tj. periody ve více směrech, úhlů mezi nimi a případnwě dalších parameterů, jako je výška motivu, můžeme se pokusit kombinovat výše uvedené požadavky tak, aby postačilo provést jen omezené množství měření. V takovém případě bychom měli získat výsledky např. jen z dvou obdélníkových oblastí, tj. z dvou měření. Klíčové je v takovém případě použít alespoň 20 period v rychlé ose měření a rozlišení nastavit tak, abychom měřili čtvercové pixely. V mnohých případech může být rychlejší provést oba typy měření odděleně (omezený počet velmi dlouhých profiů pro periody a kratší obdélníkové oblasti se čtvercovými pixely pro úhel).



ČERPANÉ NÁKLADY NA ÚKOL 4.

		Plán	Odhad čerpání do 31. 12. 2022
Cena hodin	tis. Kč	1430	1430
Materiál	tis. Kč	150	150
Externí kooperace	tis. Kč	130	130
Výdaje na porovnání v zahraničí	tis. Kč		
Interní kooperace	tis. Kč		
Kooperující středisko			
Cestovné zahraniční	tis. Kč		
Celkem	tis. Kč	1710	1710

Tabulka 1: Tabulka plánovaných a čerpaných nákladů (v tis. Kč):

Většina výdajů byla čerpána formou hodin. Externí kooperace se týkala vývoje programu Gwyddion. Materiál byl použit na stavbu jednotlivých experimentálních zařízení zmíněných v technické části zprávy.


Český metrologický institut Oblastní inspektorát Brno

ZÁVĚRY 5.

Projekt se zaměřil na konstrukci hardwarových prvků, elektronických obvodů a metodických doporučení pro zpřesnění měření v oblasti periody a planárního úhlu mikro- a nanostruktur. Podařilo se získat řadu podnětů pro zlepšení metodických postupů, tak aby bylo možné provádět měření úhlu a periody v souladu s nejnovějším vývojem v odborné literatuře a v některých oblastech, jako jsou algoritmy pro vyhodnocení periody jsme se pokusili pochopení problematiky odbornou veřejností dále posunout a uvést na pravou míru některé širece rozšířené omyly.

Jedním z hlavních cílů úkolu bylo snížení nejistoty měření úhlu na planárních strukturách pod 50 μ rad. Z provedených experimentů a metodických postupů je patrné, že dosažení takové nejistoty je v nově vyvinutém uspořádání možné, i když konkrétní nejistota měření na zákazníckém vzorku vždy závisí také na parametrech tohoto vzorku - povrchové drsnosti, přesnosti výroby objektů tvořících např. difrakční mřížku, vlivu spojování polí v litografii (stitching error), a podobných jevech. Tyto faktory ovlivnit nedokážeme, nicméně z pohledu komponent nejistoty, které jsou na naší straně (úhlová stabilita, algoritmy pro vyhodnocení, rozlišení) jsme pro taková měření připraveni.



Český metrologický institut Oblastní inspektorát Brno

Seznam bibliografických citací

- [1] Gaoliang Dai, Frank Pohlenz, Thorsten Dziomba, Min Xu, Alexander Diener, Ludger Koenders, and Hans-Ulrich Danzebrink. Accurate and traceable calibration of two-dimensional gratings. 18(2):415– 421, 2007.
- [2] E. Manske, T. Hausotte, R. Mastylo, T. Machleidt, K.-H. Franke, and G. Jäger. New applications of the nanopositioning and nanomeasuring machine by using advanced tactile and non-tactile probes. Meas. Sci. Technol., 18:520, 2007.
- [3] Jong-Ahn Kim, Jae Wan Kim, Byong Chon Park, Tae Bong Eom, and Chu-Shik Kang. Calibration of two-dimensional nanometer gratings using optical diffractometer and metrological atomic force microscope. Proceedings of SPIE Vol. 5879, page 58790Z, 2005.
- [4] Josef Lazar, Petr Klapetek, Miroslav Valtr, Jan Hrabina, Zdenek Buchta, Onrej Cip, Martin Cizek, Jindrich Oulehla, and Mojmir Sery. Short-range six-axis interferometer controlled positioning for scanning probe microscopy. Sensors, 14(1):877–886, 2014.
- [5] David Nečas and Petr Klapetek. Synthetic data in quantitative scanning probe microscopy. Nanomaterials, 11(7), 2021.
- [6] J.F. Jørgensen, C.P. Jensen, and J. Garnæs. Lateral metrology using scanning probe microscopes, 2D pitch standards and image processing. 66:S847-S852, 1998.
- [7] Gaoliang Dai, Ludger Koenders, Frank Pohlenz, Thorsten Dziomba, and Hans-Ulrich Danzebrink. Accurate and traceable calibration of one-dimensional gratings. 16(6):1241–1249, 2005.
- [8] Leo I. Bluestein. Linear filtering approach to computation of discrete fourier transform. *IEEE T.* Acoust. Speech, AU18(4):451, 1970.
- [9] Nathan Flowers-Jacobs, Alessio Pollarolo, Kevin Coakley, Adam Weis, Anna Fox, Horst Rogalla, Weston Tew, and Samuel Benz. The nist johnson noise thermometry system for the determination of the boltzmann constant, 2017-12-29 00:12:00 2017.



Český metrologický institut Oblastní inspektorát Brno

PŘÍLOHY

Příloha č. 1 Kopie plánovacího listu

Příloha č. 2 Rukopis článku - D. Nečas, A. Yacoot, M. Valtr a P. Klapetek, Demystifying data evaluation in the measurement of periodic structures

- Výsledky optické charakterizace nehomogenních vrstev s drsnými rozhraními Příloha č. 3
- Příloha č. 4 Schéma zapojení obvodu pro analogové zpracování signálů z PSD

PLÁNOVACÍ LIST ÚKOLU TECHNICKÉHO ROZVOJE

Číslo				Termíny	řešení	Neinvestič.	
úkolu*	Název úkolu	Stručná charakteristika úkolu				prostředky	
		Výsledek řešení úkolu		zahájení	ukončení	tis. Kč	
UTR22E601404	AFM měření na velkých plochách	Cílem úkolu je vyvinout nástroje pro měření rozměrů a mikrostrukturách využívaných v optickém a elektronic průmyslu na plochách až 5x5 cm s nejistotou menší 10 měření rozměrů a menší než 50 mikroradiánů pro měře povrchu vzorků.	a úhlů na ckém) nm pro ení úhlů na	1/2022	12/2022	1 710,-	
Předkládá:	I	Bank. spojení / č. účtu: 198139621/0710		1	I	I	
Český metrologický institut.		IČO / DIČ: 00177016/CZ00177016	IČO / DIČ: 00177016/CZ00177016				
úsek FM, Hvožďanská 3, 148 01, Praha 4			-5				
		Odpovědný řešitel: Mgr. Petr Klapetek, Ph.D.	Došlo:				
		Podpis:	Podpis: Zdroj financování: ČMI		4I		
odb. řed. pro FM		Tel. / Fax: 545 555 239 / 545 555 183	Poznámka:				
-		Datum: 26. 10. 2021					

^{*} v případě externího zadavatele (např. ÚNMZ), nezadává toto číslo ČMI Verze 17-001_P02C_011-ZS-C008 1.1 Hlavní cíle úkolu (důvod zadání úkolu, např. zákonné požadavky, plnění koncepce rozvoje NMS, plnění požadavků CIPM MRA):

Mikro- a nanotechnologie se stále častěji využívají v různých průmyslových odvětvích a s tím roste i potřeba speciálních kalibračních vzorků. Ty často obsahují objekty, které je současnými metodami využívanými v nanometrologii obtížné, nebo nemožné měřit. Takovým průmyslovým požadavkem může být například potřeba měření roztečí velkého množství objektů tvořících speciální difrakční mřížku, nebo měření úhlů mezi směry takové difrakční mřížky s velmi malou nejistotou. Metrologická zařízení schopna měřit s dostatečnou přesností (jednotky nm) jsou typicky schopna měřit jen na poměrně malé ploše (například 200x200 mikrometrů v případě státního etalonu využívaného v nanometrologii), což je pro takové úlohy nedostačující. Měření úhlů je ještě obtížnější – obecně se mikroskopy atomárních sil (AFM) využívají jen pro měření v rychlé ose skenování a data v pomalé ose skenování jsou považována za méně přesná. Měření úhlů pomocí těchto zařízení je proto obecně špatně popsané a málo rozvinuté téma metrologie a to navzdory tomu, že požadavky zákazníků na taková měření tvoří cca polovinu poptávek.

Cílem úkolu je využít platformu pro skenování na velké ploše vyvinutou v roce 2021 a vytvořit z ní systém schopný měření s nejvyšší možnou přesností při rozměrových a úhlových měřeních. To zahrnuje po hardwarové stránce vybudování referenčního systému s diferenciálními interferometry a vytvoření kompenzace úhlových chyb lineárních vedení. Po metodické stránce to zahrnuje vytvoření postupu pro měření s izotropní nejistotou aby bylo možné provádět úhlová měření. Dalším cílem bude zajištění vhodných nástrojů pro přenos návaznosti z etalonů úhlu na toto zařízení.

V rámci řešení úkolu bude také pokračováno v dlouhodobých rozvojových aktivitách v oblasti optických měření na drsných površích a v oblasti technického přizpůsobení při měření teploty prostřednictvím šumu (spolupráce s Oddělením primární metrologie tepelně-technických veličin na přípravě experimentu k porovnání prototypu Johnsonova šumového teploměru a sestavy primárního akustického plynového teploměru). Ty jsou na AFM měření, které jsou primárním tématem úkolu navázány prostřednictvím měřených vzorků (etalony drsnosti) a sond (rastrovací termální mikroskopie).

Projekt přímo navazuje na koncepci rozvoje metrologie a cíl roku 2022: Snížení nejistoty měření úhlu na planárních strukturách pod 50 μrad.

1.2 Objektivně ověřitelné výsledky řešení úkolu:

Rozpočet nejistoty pro planární a úhlová měření na velkých plochách.

1.3 Způsob ověření: (uvést neopomenutelné účastníky):

Závěrečná oponentura za účasti oponentů.

1.4 Navrhovaní oponenti úkolu (minimálně 2, podléhají odsouhlasení zadavatelem úkolu): Prof. RNDr. Miloslav Ohlídal, CSc., VUT v Brně

Mgr. Petr Klenovský, Ph. D, Masarykova univerzita v Brně

- 2.1 Dílčí cíle nezbytné k dosažení hlavních cílů úkolu (I. etapa řešení) : Vytvoření systému kompenzací úhlových chyb lineárních vedení včetně elektroniky a ovládacího SW.
- 2.2 Objektivně ověřitelné ukazatele dílčích cílů Výsledky testování systému kompenzací.

2.3 Způsob ověření dílčích cílů

3. Kalkulace úkolu (rozpis nákladů) a ekonomické zhodnocení (přínos ekonomický): odhad jednorázových nákladů (stroje, zařízení, budovy) a odhad nákladů na projekt (pracovní hodiny, materiál, kooperace)

číslo střediska		6014		
č. úkolu	neuvádět			
počet hodin střediska na celý úkol x		746h x 1755 Kč = 1 23		55 Kč = 1 230
sazba				
materiál	tis. Kč	150		
externí kooperace	tis. Kč	250		
výdaje na porovnání v zahraničí	tis. Kč			
interní kooperace	tis. Kč	80		
kooperující středisko				
cestovné zahraniční	tis. Kč			
celkem	tis. Kč	1710		1710

4. Zdroj financování úkolu (v %): 100 % hrazeno z ČMI

5. Praktické využití výsledků řešení a požadavky pro udržovatelnost výsledků (včetně vlivu výsledku na životní prostředí):

Výsledky řešení úkolu budou využity při měření pro zákazníky v oblasti optického a elektronického průmyslu a pro firmy zabývající se vývojem mikroskopických systémů (SEM, SPM).

6. Spoluřešitelé úkolu (uvést rozsah spolupráce-věcný, finanční):
 Mgr. David Nečas, Ph.D. - vývoj software pro zpracování dat, 80 tis. Kč

7. Nároky na zahraniční cesty (stát, organizace, délka pobytu, finanční nároky):

- s ohledem na aktuální situaci nejsou

8. Prohlašujeme, že naše organizace současně nenárokovala, resp. nemá přislíbeny finanční prostředky na řešení úkolu ze státního rozpočtu prostřednictvím jiného orgánu státní správy, popř. nevyužila jiné, duplicitní cesty k financování navrhovaného úkolu. Název, adresa, razítko a podpis statutárního zástupce:

Český metrologický institut Okružní 31 638 00 Brno

doc. RNDr. Jiří Tesař, Ph.D. generální ředitel

Pozn.: Pokud nestačí formulář, je možno příslušný bod rozvést na další přílohu.

Demystifying data evaluation in the measurement of periodic structures

D Nečas¹, A Yacoot², M Valtr^{1,3} and P Klapetek^{1,3}

 1 CEITEC, Br
no University of Technology, Purkyňova 123, 61200 Brno, Czech Republic

² National Physical Laboratory, Hampton Road, Teddington, Middlesex TW11 0LW United Kingdom.

³ Czech Metrology Institute, Okružní 31, 63800 Brno, Czech Republic

E-mail: yeti@physics.muni.cz, pklapetek@cmi.cz, andrew.yacoot@npl.co.uk, mvaltr@cmi.cz

Abstract. Periodic structures are often found in various areas of nanoscience and nanotechnology with many of them being used for metrological purposes either to calibrate instruments, or forming the basis of measuring devices such as encoders. Evaluating the period of one or two-dimensional periodic structures from topography measurements, *e.g.* performed using scanning probe microscopy (SPM) methods, can be achieved using different methodologies with many grating evaluation methods having been proposed in the past and applied to a handful of examples. The optimum methodology for determining the grating pitch is not immediately obvious. This paper reports the results of extensive large-scale simulations and analysis to evaluate the performance of both direct and Fourier space data processing methods. Many thousands of simulations have been performed on a variety of different gratings under different measurement conditions and including the simulation of defects encountered in real life situations. The paper concludes with a summary of the merits and disadvantages of the methods together with practical recommendations for the measurements of periodic structures and for developing algorithms for processing them.

Keywords: Scanning Probe Microscopy, traceability, grating pitch, nanometrology, uncertainty

1. Introduction

Surface topography measurements are one of the important tools in the area of nanoscience and nanotechnology. The more complex the properties that are measured, the more important is the selection of the data processing methodology, as surface topography measurements do not provide the result directly. In this paper we deal with evaluation of grating parameters and similar periodic structures, which can be used for a variety of applications. Periodic structures in the nanometre to micrometre range are frequently found in the areas of surface science, nanoscience and nanotechnology. Starting from the nanoscale, the atomic lattices, that are providing insight into the atomic arrangement of matter are periodic and their properties can be evaluated from very high spatial resolution measurements of surface topography [1, 2]. The lattice parameter of silicon is even recognised as a secondary realisation of the metre for dimensional nanometrology [3]. At larger scale, surface topography is being used to analyse artificially created 2D periodic structures that can be used as metamaterials [4], photonic crystals [5–7] or phononic structures [8], controlling the way the energy passes through the matter. In all these cases the period is one of the key characteristics, directly affecting the material function.

When it comes to manufacturing methods themselves and related measurement science, there are also different potential roles for periodic structures. First of all, they can be used as key components in the manufacturing process, such as gratings used for mask overlay adjustment in the semiconductor industry [9]. Second, they can be a critical part of the measurement device itself. Most of the area sensors, e.g. CCD chips are also periodic structures with an arrangement similar to 1D or 2D gratings and position of individual features may have direct impact on the measurement accuracy, e.g. when sub-pixel accuracy needs to be achieved, as in astronomy, and when both the geometrical errors and electronic performance defects of individual pixels need to be considered [10], or, at larger scale, when making accurate X-ray tomography measurements [11]. Shack-Hartmann wavefront sensors combine a 2D periodic microlens array with a CCD chip, mounted in the focal plane of the microlens array (usually this is calibrated by using an ideal wavefront). Grating structures are used as the basis of optical encoders and 2D gratings are being adapted for multi-axis position sensing [12]. Self-assembled periodic particle or hole arrays can be used as substrates for surface enhanced Raman scattering measurements [13], providing the plasmonic field enhancement. All the above mentioned structures need to be either measured or calibrated at some stage and one of the approaches is to use some of the surface topography measurement methods, like scanning probe microscopy (SPM) or surface profilometry which in turn use gratings as calibration standards. This has led to the publication of ISO guidelines e.g [14] ISO 11952 Surface chemical analysis — Scanning-probe microscopy Determination of geometric quantities using SPM: Calibration of measuring systems outlining methods for the evaluation of grating periods. The methods are not restricted the to calibration standards but are generically applicable to all periodic structures.

This raises the question, why there is a need to revisit the measurement of grating period? Experience shows that in the main, people predominantly use only the simplest possible methods, even if the ISO standard recommends more sophisticated methods. Based both on our experiences with SPM data processing software development over last 20 years and on the experiences from an interlaboratory comparison [15], we see that many SPM users evaluate grating period only from a distance of two crossing points on a line profile, or, base their uncertainty in the grating pitch solely on the standard deviation, from multiple distance measurements. So, many users still struggle to calibrate their AFMs using gratings, thereby limiting their ability to make basic

3

dimensional measurements. Moreover, many users tend to systematically underestimate the area that is needed for obtaining statistically significant results, as shown already also for roughness measurements [16]. The intuitive choice of the scan area is not optimal and without a detailed analysis it is hard to understand why this is the case.

One of the reasons why people are struggling with periodic structures analysis might be the wide variety of choices available. There are many methods in the literature and standardisation documents that can be used for evaluation of periodic structures and their explanation is not always clear. The Fast Fourier Transform (FFT) method, gravity centre method and a combined method (FFT+cross-correlation) are at present recommended by ISO 11952. Even though there is guidance on the number of grating periods that should be measured on the sample and suggested resolution (e.q. more than 5 for gravity centre method and more than 7 for FFT methods), there is no detailed analysis in the standard or referenced papers of the dependence of resulting uncertainty on the size of the area scanned. It can be seen that preferably a large number of periods should be chosen for achieving the smallest uncertainties [17, 18]. However, it is hard to guess how the uncertainty will grow when these ideal conditions are not met, as is likely to be the case in reality. When a grating is measured, the main limitations are the scanning range and grating pitch. The scanning ranges can be varied as well as the choice of grating. The typical maximum scanning range for an AFM is 100 µm. The range of most grating pitches is between 100 nm and $5\,\mu\text{m}$, for SPM calibration gratings, but can be very different in other cases. The typical number of pixels in an SPM image is 500 to 5000 pixels. This gives users a wide choice of parameter combinations and from the AFM comparison [15], it was seen that users chose a variety of pitch/scan range parameters. In this paper we consider the effect of this poor choice. Moreover, it is not clear how the different aspects of non-ideal measurement (feedback loop effects, noise, tip convolution) or non-ideal grating parameters (form errors, roughness) and of course limited scan range, affect the result. Special cases are the lateral imperfections of the grating such as misplaced grating pits or stitching errors introduced in the fabrication process. Even if the calibration gratings available on the market are very good from this point of view, the uncertainties of present metrological SPMs are at the nanometre level so the measurements should also address this aspect. Moreover, grating analysis methods are used for analysis of other periodic structures used in industry (CCD chips, microlens arrays) where the positioning error can be much larger.

There are different types of periodic structures that could be evaluated using surface measurements. Figure 1 shows synthetic images of some typical periodic structures, including some typical experimental errors and sample imperfections.

2. Methods

The goal of this paper is to systematically assess the performance of various implementations of the periodic structure evaluation methods. Therefore, the methodology is purely based on numerical simulations, using synthetic data with known parameters [19]. Using



Figure 1. Examples of periodic structures: (a) an ideal 1D grating; (b) a measurement of a too small area on a 2D grating; (c) nanowires; (d) a poor quality 1D grating; (e) an atomic lattice — Si 7×7 surface reconstruction; (f) self-assembled spherical particles.

the data synthesis tools in the open access AFM data analysis software Gwyddion [20], grating surfaces with different properties and deterministic distortions can be generated. In detail, for all the simulations many gratings with slightly different parameters were generated in each simulation to suppress coincidences and aliasing effects. The interinstance parameter variation range was 5%. If, for example, the nominal grating period was 50 pixels (and this value would be shown in a figure), individual generated gratings would in fact have periods from 47.5 to 52.5 pixels. Various data processing methods can be then applied to the generated gratings and their performance can be evaluated statistically. In contrast to using real data this can directly provide a quantification of the errors as the true parameters of generated surfaces are known. It also permits a wider variation in the value of parameters that would be possible with experimental work. The analysis is performed initially on 1D gratings, to introduce the methodology and different error source types. Then, the differences obtained when evaluating 2D gratings are presented and discussed.

Methods suitable for analysis of gratings and other periodic structures that can be found in the literature and in ISO 11952 can be divided into the categories as presented below together with details of their implementation. Eight period evaluation methods were implemented. Their principles are illustrated in figure 2.



Figure 2. Principles of methods for one-dimensional grating period evaluation: (a) Fourier transform based methods; (b) autocorrelation function based methods; (c) gravity centres; (d) zero crossing; (e) fitting of a model function.

2.1. Direct space measurement

As mentioned earlier, direct space measurement can be the most intuitive approach, namely when a profile is drawn across the periodic structure and the lateral distance between two or more similar features on the profile is evaluated [21,22]. More advanced approaches use more elaborate methods to determine the period, *e.g.* by evaluating the position of the profile at some height, interpolating or fitting the points near to this position to obtain the position more accurately. This is called the zero crossing method and is also often used [22–24]. To use only points related to a special feature and throw away the rest of the structure might be not very efficient, so the centre of gravity evaluated from whole structures can be used in order to include all of the data [18,23–25].

Two methods working in the direct domain were implemented for this study:

Gravity centre (GC) — analysis of gravity centres of grating bars [17, 24]. The measured profile is plotted as a function, a threshold is chosen and the profile shifted to make the threshold line z = 0. The centre of gravity of each bar is then defined as the center of gravity of the area under the curve (figure 2c). Its horizontal coordinate is $x_c = \int xz \, dx / \int z \, dx$ (both integrals are over the bar interval). We obtain a set of centres $x_{c,n}$, indexed by integer bar number n, and fit them with a linear function

$$x_{\mathbf{c},n} = nT + c \tag{1}$$

with parameters T (period) and c (offset). This can be repeated with areas above the curve for the negative parts and the results averaged.

Zero crossing (ZC) — analysis of positions where the profile crosses the zero line [22, 24, 26]. A zero line is chosen as in the gravity centre method. Data around each zero line crossing are fitted with straight line z = a + bx to estimate precisely

the crossing coordinate x_0 (figure 2d). We obtain the coordinates of up-crossing positions $x_{0,n}$, indexed by integers n, and fit them with a linear function

$$x_{0,n} = nT + c \tag{2}$$

with parameters T (period) and c (offset). This can be repeated with down-crossings and the results averaged.

The implementation of some methods is straightforward, whereas others require more care to work reliably. Some comments on our experience with their implementation are made after the results of numerical simulations have been presented.

2.2. Fourier transform

Next class of methods is based on the spectral density of spatial frequencies that can be obtained using Fourier transform (FT). A 1D or 2D discrete Fourier transform (DFT) is run on the measured topography and from the frequencies corresponding to the peaks the period is evaluated [23, 27]. As the lateral size of the scan is limited, the frequency resolution can be very coarse, which can be handled using different approaches for calculating a refined FT [18,28], *e.g.* calculating the spectral density also for non-integer components.

The following methods working in the frequency domain were implemented for this study:

Simple FFT — an elementary Fourier transform based estimation. The DFT of the measured profile data z_n (both n and ν take values 0, 1, 2, ..., N-1)

$$Z_{\nu} = \sum_{n=0}^{N-1} z_n \exp\left(-2\pi \mathrm{i}\frac{n\nu}{N}\right) \tag{3}$$

is computed using an FFT. We then find the index ν where the spectral density $|Z_{\nu}|^2$ attains its maximum and take the corresponding spatial frequency $f = \nu/(Nh)$ as the pitch value, where h is the sampling step. The period is obtained using the relation T = 1/f (figure 2a). Prior to the Fourier transform, data are multiplied by a windowing function. The simple raised cosine Hann window was used (in all frequency domain methods).

- **Dai05 FT** refined Fourier transform [17, 28]. A coarse estimate is computed using the previous method and then refined by allowing non-integer values of ν . The search for the precise maximum starts with the interval $[\nu - 1, \nu + 1]$ around the integer coarse maximum ν . This interval is then progressively refined using a simple grid search, until it becomes shorter than a prescribed length. Fourier coefficients are computed by a direct evaluation of expression (3).
- **Zoom FFT** Zoom-FFT refinement is, in principle, equivalent to the preceding method but computed in a different way. Fourier coefficients for non-integer ν are not computed individually; Instead Bluestein's algorithm [29] is used (discussed in more detail in section 4.3). It computes Fourier coefficients corresponding to an

arithmetic progression of frequencies $f, f + \Delta f, f + 2\Delta f, \ldots, f + n\Delta f$ using an FFT. Therefore, it is possible to zoom into the interval around the coarse maximum and refine its position using a simple search. The results presented here show the effect of zooming in the interval $[\nu - 1, \nu + 1]$, computing again N Fourier coefficients, which corresponded to an N/2-times refinement.

2.3. Autocorrelation

A direct domain parallel to the spectral density is the auto correlation function (ACF). The methods implemented for this study can be considered hybrid as the measurement is made in the direct domain, but the ACF is obtained using the FFT:

Simple ACF — an elementary autocorrelation-based estimation. The discrete ACF

$$G_k = \frac{1}{N-k} \sum_{n=0}^{N-1-k} z_n z_{n+k}$$
(4)

is computed using FFT, utilising the discrete cross-correlation theorem. Integer indices, k, are related to real distances τ via the sampling step h: $\tau = hk$. Its first maximum always lies at zero. The next one corresponds to the grating period T(figure 2b) and is directly used as the estimate.

Multi-peak ACF — multiple ACF maxima are used to improve accuracy. The position of a single ACF maximum cannot be determined very precisely. However, the function has many maxima corresponding to integer multiples of T (figure 2b). We locate as many of them as feasible, obtaining a set of horizontal distance τ_n , indexed by integers n, and fit them with a linear function

$$\tau_n = nT \tag{5}$$

with a single unknown parameter T.

2.4. Fitting

Most of the methods used for grating analysis provide only pitch and angle as a result. There are, however, more parameters that could be evaluated on a grating; roundness of corners or fill ratio, that cannot be determined using the standard methods but still could provide some information for practical use of the grating, *e.g.* when one wants to characterise the shape of the AFM tip using the grating. Although tip shape is not strictly necessary when evaluating grating periods, it would be useful if the sample had other features to be measured that were non-periodic. The grating height, if constant throughout the grating structure could be used for z calibration and extracted background could, in principle, be used instead of basing it on separate measurements of a flatness standard. However, care should be taken when calibrating the z axis as the optimum measurement strategies for pitch and height are different. Fitting the data using a model is used in many areas of measurement science and using this approach for grating analysis could be understood as a straightforward approach. Fitting grating

parameters can be already done in SPM data processing software, such as Gwyddion [30] which has the capability to evaluate many other parameters rather than just grating pitch.

For least-squares fitting, a suitable function describing the grating shape must be first chosen, for instance, a rectangular or sine wave. The function has several unknown free parameters: period T, height, offsets in x and z, and possibly others such as slope width, describing the shape in more detail. An initial estimate of their values is necessary, for example this can be obtained using the simple FFT method. Precise values are then obtained by non-linear least-squares fitting of the model function using the Marquardt–Levenberg algorithm [31]. In this study a rectangular wave with sloped walls was used as the model function (figure 2e).

Least-squares fitting of long periodic data is sensitive to local phase variation. Therefore, as an alternative, the model function can be fitted separately to each repeating unit (*e.g.* grating bar), obtaining individual parameters for each. Parameters representing positions are then processed as in GC and ZC to obtain the period T. This approach will be denoted **Piecewise fitting**, whereas **Model fitting** will denote fitting the entire data with the model.

If only position estimates of individual features are required, **Cross-correlation** [?] can be used instead [?,28]. A template of the repeated unit can be created either using an explicit model or from experimental data. A synthetic rectangular wave template was used in this work, constructed based on a coarse estimate of T. The cross-correlation of the template with the data has maxima where they match best, albeit the best match criterion differs from fitting and generally the positions from the two methods do not coincide. Positions of these maxima are then refined using parabolic interpolation and again processed as in GC or ZC to obtain the period T [28].

2.5. One-dimensional gratings

One-dimensional gratings were modelled as rectangular waves with slightly sloped walls (5% of length). The following random and scanning artefact types could be added to the ideal grating data, individually or in combinations (a graphical illustration of the artefacts can be seen in section 3.1 in figure 5):

- Waviness the deviation of the grating substrate from ideally flat surface was added as a multi-scale locally smooth random additive background.
- **Unevenness** the grating geometry imperfections were added using uncorrelated random variation of individual bar parameters, including position, height and fill ratio.
- **Broken bars** another grating imperfection was introduced by random removal of top parts of individual bars (up to complete removal).
- **Particles** the presence of dust particles was added using random bumps with size typically comparable to or somewhat smaller than one grating bar.

- **Noise** the impact of SPM noise was added using independent random Gaussian noise of each sample.
- **Tip convolution** the impact of SPM probe-sample convolution was added by convolution with an ideal parabolic tip.
- **PID loop** the impact of the feedback loop imperfections was added using a simple proportional-integral-derivative feedback loop simulation. [?]

Particles and broken bars were always used together as one 'local defects' artefact. Other important systematic error sources exist: erroneous calibrations, drift, Abbe error and cosine error. Their linear components change the measured data to give a slightly different value for the grating period. Given such data, even a hypothetical ideal evaluation would compute the changed period. From the data processing standpoint they are, therefore, not interesting as all methods are affected in exactly the same way. The conclusion is similar for the non-linear components. Non-linearity can be detected in data processing and some methods can deal with local phase variation better than others [?]. However, correction requires additional information and is outside the scope of the evaluation method.

The evaluation also included a preprocessing step with two main goals: suppression of long-wavelength background (waviness) and shifting the profile mid-height to z = 0. The latter is required mainly by the zero crossing method (ZC) and to a lesser degree the gravity centre method (GC), but the same preprocessing was used for all methods. Waviness was removed using a custom envelope method (see section 4.5 for an explanation as to why it was chosen). Upper and lower profile envelopes were found as the local maximum and minimum within 1.5*T* interval (with *T* estimated using Simple FFT). Their average was processed using a low-pass Gaussian filter (0.5*T*) and subtracted from the data. The mid-height was located by finding the two main peaks in the height distribution and taking their midpoint [23].

2.6. Two-dimensional gratings

Two-dimensional gratings were generated with slightly rounded rectangular holes, half the period wide, corresponding to 3/4 of upper surface and 1/4 of lower surface. The two lattice vectors could differ in length, but they were always orthogonal. The orientation and offsets (phases) in the plane were random. A typical image is shown in figure 3a. As in the 1D case, individual grating instances varied within 5% of the nominal values.

The simulated images could include three artefacts analogous to 1D: noise, waviness and uneven positions. In addition, random tilt could be added as it is ubiquitous in AFM images. Images usually cover much smaller areas than scan designed to capture what is effectively a line profile, specified by the number of periods and several scans very close together perpendicular to the profile. Therefore, only this simple background was considered. Also the preprocessing was simpler and only included tilt removal. It was implemented as an initial plane levelling, followed by splitting the surface to upper



Figure 3. (a) An example of typical grating image used in the simulation, with dimensions 500×500 pixels, period T = 50 pixels and a combination of selected artefacts. All values are rescaled to pixels, *i.e.* unitless. (b) Point lattice which is the usual intermediate step of 2D grating evaluation. (c) Recursion in the 2D GC computation.

and lower portion using Otsu's threshold [32] and final plane levelling using only the upper portion of the surface (as defined by the threshold).

A subset of 1D evaluation methods was implemented: simple FFT and ACF, refined FT, multi-peak ACF, GC and model fitting. Zero crossing was not implemented because it is not clear how it generalises to 2D. Both FT and ACF based methods generalise directly to 2D. Refined FT was only implemented using Zoom-FFT as the two refinement methods are equivalent (this was already verified for the 1D case).

Both FT and ACF methods produce sets of points in a more or less regular lattice (figure 3b) and two lattice vectors have to be selected [28]. The procedure can be outlined in two steps:

- (i) Find the point closest to the origin, but not at the origin. Use its position as lattice vector \mathbf{a}_1 .
- (ii) Find the point which is closest to the origin and linearly independent (sufficiently small scalar product with \mathbf{a}_1). Use it as \mathbf{a}_2 .

Integer indices of any other point \mathbf{v} are then determined (where necessary) by solving $\mathbf{v} = m\mathbf{a}_1 + n\mathbf{a}_2$ for m and n and rounding them to integers. Then the positions are fitted by linear least squares model similar to expressions (1), (2) and (5), only 2D:

$$x_{m,n} = ma_{1x} + na_{2x}; \quad y_{m,n} = ma_{1y} + na_{2y} \tag{6}$$

with free parameters a_{1x} , a_{1y} , a_{2x} and a_{2y} .

GC has two steps, identification of holes and computation of their centres. Otsu's threshold was again used to choose the z = 0 plane and mark the holes. Holes touching image borders were filtered out as well as holes that were too small (single-pixel holes).

The GC was defined exactly the same as in the 1D cases, even though its computation was more involved, requiring integration over the region where interpolated data lie below the z = 0 plane. A simple recursive quadrature was used (figure 3c). We started with pixel-sized rectangles formed by 2×2 neighbour values and then:

- (i) If the values at all rectangle corners were negative it was considered completely covered. The integral over the rectangle was computed analytically using the bilinear interpolation of the four corners.
- (ii) If all corners were negative the rectangle was skipped.
- (iii) If some corners were positive and some negative, the rectangle was split into four, with corners computed by interpolation, which were then evaluated recursively.
- (iv) If a rectangle became too small the recursion was terminated.

The image has N^2 pixels but only O(N) are at hole boundaries, requiring recursion. Therefore, a more efficient integration method was not necessary. After finding all centres, one close to image centre was chosen as the initial origin. The analysis then proceeded as for the case of a multi-peak ACF, except that the origin position was also a free parameter and it was updated after each fitting step.

Model fitting was implemented as in the 1D case, again relying on simple FFT for the initial lattice vector estimates. The model function was similar to the grating generation function. However, it was more general, allowing non-orthogonal lattice vectors. Not doing so would give fitting an unfair advantage over the other methods since the generated gratings had orthogonal lattice vectors.

2.7. Gratings with small number of periods

Data with just a couple of periods are evaluated differently than data with a thousand. The profile is likely to be scanned more slowly with respect to feature size, the entire data can be inspected and ensured they are defect-free, levelling and zero line can be checked manually. It is not entirely fair to use the models and methods outlined in sections 2.5 and 2.6, focused on automated processing of long profiles and large areas, to study this case. Therefore, they were modified for the 1D case as follows:

- The grating model was a perfect rectangular wave, distorted by the convolution with a rounded triangular tip (which was the only systematic artefact considered).
- There would be no significant local defects and the background would be mainly tilt, corrected by the user. Hence, we considered only one random artefact: noise.
- The zero line was set exactly at mid-height.
- A 'manual' evaluation method was included.

The typical choice of two points for manual measurement is zero crossings, so that is what the 'manual' method used (it was still carried out automatically). The two most distant zero crossings of the same type (up or down) were found, located with



Figure 4. Distributions of relative errors of the evaluated pitch for almost ideal gratings with approximately 1000 periods. Note the order-of-magnitude factors in the bottom right corners of the graphs — the errors can differ by several orders. For some methods errors much larger than typical were occasionally encountered (as indicated by the maximum error δ_{max}). In these cases the abscissae do not cover the outliers and corresponds roughly to the mean square error δ_{rms} .

subpixel precision by linearly interpolating the two adjacent points, and their distance was divided by the number of periods between.

Other methods were tweaked to stretch their applicability. Usually, in GC and ZC both 'up' and 'down' features are analysed and the results averaged. This was still done when possible, but if the method could find at least one measurable T, the evaluation was considered successful. Zoom FFT started the initial coarse estimate from $3 \times$ zoomed FFT, instead of plain FFT, because the peak could be indistinguishable from the peak at origin otherwise.

3. Results and discussion

3.1. Methods performance on one-dimensional gratings

The distributions of relative errors δ_T of the period are illustrated in figure 4 for a typical calibration grating with 1000 periods, 50 samples per period and waviness background with noise to signal ratio of 7% (approximately corresponding to [17], by visual comparison). The distributions were obtained by running the evaluation on 25000 random grating instances. The error distributions are anything but Gaussian. Simple FFT has a uniform error distribution, which is expected because the error is basically a rounding error. Simple ACF has an odd bimodal error distribution, which is probably related to the use of parabolic interpolation to improve the maximum location. Its shape seems partially preserved in multi-peak ACF error distribution, which is also rather asymmetrical. All the seemingly Gaussian distributions have in fact heavier tails. The error distributions become more conventional when waviness is replaced by simple noise — for instance multi-peak ACF asymmetry disappears. However, most simulated artefacts led to odd error distributions and heavy tails.

Both less accurate methods, simple FFT and ACF, consistently give results with bounded errors; the maximum error encountered is a small multiple of the mean error. The more accurate methods occasionally give a value with much larger error than typical, causing the heavy tails of the distributions. For FFT-based and cross-correlation methods this occurs when the period is very close, but not exactly equal, to an integer number of pixels. Grating edges can then align with sampling points in a way that makes edge positions less certain than if there was no relation between the grating period and sampling step. Direct space methods are affected for a similar reason. In the following we will refer to the eight more accurate methods as 'the good methods' for brevity.

Figure 5 compares how the methods behave when the number of periods P in the profile change. For additive disturbances the noise ratio was kept at 7%, likely to be the worst case encountered in a real situation. For uneven bars the relative standard deviation of parameters was 2.5%. Particles and broken bars covered each randomly and independently 2% of the profile. Tip convolution and PID loop are non-random effects and their parameters were chosen to obtain roughly comparable disturbance of the profile shape. The accuracy $\delta_{\rm rms}$ is measured as the mean square relative error.

There are some differences in sensitivity to different artefacts. Multi-peak ACF seem the most difficult to thwart by local defects (particles & broken bars), but it is more susceptible to asymmetry of the shape (PID loop) than others. The rapid deterioration of accuracy of GC, ZC correlation search and fitting in the case of local defects and low P is driven by occasional cases when an unfortunate constellation of defects manages to derail the method entirely. In most cases the accuracies of the good methods are comparable, although the logarithmic scale in figure 5 is deceptive as curves that appear close to each other can still differ by factor 2 or 3.

3.2. Super-linear scaling

For most methods, the accuracy in figure 5 clearly follows a power law. The scaling exponents are listed in table 1. As expected, the accuracy of simple FFT scales linearly with P, the number of periods. Simple ACF shows almost no improvement with an increased number of sampled periods beyond a certain point. It is limited in precision



Figure 5. Scaling of the accuracy (mean square relative error) $\delta_{\rm rms}$ of grating period with the number of measured periods P. Curves for the two refined FT methods cannot be visually distinguished. Each plots illustrates the corresponding type of artefact with features disturbed in scale to the simulation.

by the sampling inverval h which does not change. The most interesting observation, however, is that the accuracy of all the good methods scales super-linearly with P. The exponent somewhat varies among them; it also varies somewhat with simulation settings. However, it is consistently around 3/2 or larger.

Several factors contribute to the super-linear scaling. The easiest case to analyse is the multi-peak ACF. The period is obtained by least-squares fitting of peak positions with the model equation (5). Assuming for simplicity uniform uncorrelated Gaussian errors σ_{τ} of positions τ_n , the estimated period is $\hat{T} = S_{n\tau}/S_{nn}$ and its variance $\operatorname{Var}[\hat{T}] = \sigma_{\tau}^2/S_{nn}$, where $S_{nn} = \sum_n n^2$ and $S_{n\tau} = \sum_n n\tau_n$. The sum S_{nn} is just the sum of squares of the first P natural numbers $S_{nn} = P(P+1)(2P+1)/6 \approx P^3/3$ (for

Waviness	Noise	Uneven	Defects	Tip conv.	PID loop
1.00	1.00	1.00	1.00	1.00	1.00
1.97	1.50	1.50	1.56	1.49	1.46
1.99	1.50	1.50	1.56	1.50	1.51
0.05	0.27	0.57	0.32	0.09	0.13
1.86	1.54	1.49	1.54	1.79	1.81
2.00	1.52	1.51	1.57	1.51	1.57
2.05	1.51	1.51	1.56	1.55	1.73
2.05	1.50	1.50	1.57	1.61	1.54
2.06	1.52	1.52	1.57	1.51	1.54
1.83	1.52	1.52	1.57	1.51	1.56
	Waviness 1.00 1.97 1.99 0.05 1.86 2.00 2.05 2.05 2.06 1.83	WavinessNoise1.001.001.971.501.991.500.050.271.861.542.001.522.051.512.051.502.061.521.831.52	WavinessNoiseUneven1.001.001.001.971.501.501.991.501.500.050.270.571.861.541.492.001.521.512.051.511.512.061.521.501.831.521.52	WavinessNoiseUnevenDefects1.001.001.001.001.971.501.501.561.991.501.501.560.050.270.570.321.861.541.491.542.001.521.511.572.051.511.511.562.051.501.501.572.061.521.521.571.831.521.521.57	WavinessNoiseUnevenDefectsTip conv.1.001.001.001.001.001.971.501.501.561.491.991.501.501.561.500.050.270.570.320.091.861.541.491.541.792.001.521.511.571.512.051.511.511.561.552.051.501.501.571.612.061.521.521.571.511.831.521.521.571.51

Table 1. Estimated accuracy scaling exponents for 1D methods and different artefact types.

large P). Therefore, the standard deviation of T is $\sigma_T \approx \sqrt{3} \sigma_\tau / P^{3/2}$, giving scaling exponent 3/2.

The same effect is in play in GC, ZC, cross-correlation and both fitting methods. If the ideal profile is disturbed by adding Gaussian noise, position errors are close to uncorrelated and normally distributed, all three indeed scale in line with the theoretical expression. In fact all the good methods scale similarly with exponent $\approx 3/2$ in the white noise case. Deviations from 3/2 are, therefore, influenced by spectral properties of the disturbance as they scale non-trivially with the number of data points. For instance, the simulated waviness has a certain frequency spectrum and effectively disturbs short profiles more than long ones.

In refined FT methods the scaling comprises two factors. A factor P comes with the increased frequency domain resolution, as in unrefined FFT. The additional $P^{1/2}$ is determined by how precisely the peak can be located inside one frequency step, *i.e.* how deep we can 'zoom'. Intuitively, the DFT concentrates all direct space data to a few peaks in the spectrum. The peak width (measured in DFT frequency steps) does not depend on the profile length. So for longer profiles more data contribute to one peak and thus better define its shape.

3.3. Optimal number of samples per period

The super-linear scaling leads to an important conclusion concerning measurement strategies. Assume we measured a profile with P periods. Now we want to increase the precision by measuring $5\times$ more data. Keeping the sampling step h, we can either repeat the measurement five times or measure a single $5\times$ longer profile. Although the first option is useful if representativeness can be an issue, the second is vastly more precise. Measuring five times and averaging reduces the standard deviation by factor $\sqrt{5} \approx 2.2$. However, measuring a five times longer profile reduces the standard deviation by factor $5^{3/2} \approx 11.2$ for scaling exponent 3/2.

Consider now that the maximum number of samples is limited, but the sampling



Figure 6. Optimisation of the number of samples per period for 1D gratings. The mean square error $\delta_{\rm rms}$ of grating period is plotted as a function of the number of measured periods P for three fixed numbers of samples in the profile (480, 4800 and 48000).

step can be chosen freely. Should we measure many periods, a few, or is there a Goldilocks zone? If $\sigma_T \propto \sigma_x/P^{3/2}$ and the position error σ_x is proportional to sampling step, *i.e.* $\sigma_x \propto h \propto P$, it can be seen that $\sigma_T \propto 1/\sqrt{P}$. Therefore, it seems as many periods as possible should be crammed into to the measured profile because the error decreases monotonically with P.

This suggestion may sound counter-intuitive. When performing SPM measurements, one often cares about pixel size, assuming that it is the key parameter limiting the accuracy of the result. However, in contrast to simple manual evaluation, all the presented methods are substantially sub-pixel by nature, by using all the available data.

Of course, if the sampling step becomes too long, and in particular when reaches or exceeds T/2, individual bars become impossible to distinguish and the scaling relation breaks. Therefore, there will be an optimal sampling step which was estimated using

a numerical simulation. Its results for fixed N = 480, 4800 and 48000 and varying number of periods P (that we again assume can be chosen freely) are shown in figure 6. The three cases roughly correspond to a standard profile (possibly read from an image) using routine settings for pixel resolution on a commercial instrument, a profile obtained using commercial AFM at the limits of the possible pixel resolution, and a measurement using a specialized long-range metrological AFM capable of producing scans of virtually unlimited pixel resolution at the cost of low speed. A combination of random artefacts was used in this simulation, noise and waviness, each with 1% noise, and slightly uneven bar parameters with 10^{-5} relative standard deviation.

All the good methods improve down to about 20–30 samples per period (N/P)before the errors level off and then become erratic, varying depending on how the sampling step and period align. The accuracy of piecewise fitting deteriorates quickly around 10 samples per period, indicating the implementation might be insufficiently robust for small N/P. The simple methods behave differently. For simple ACF the optimum is $P \approx \sqrt{N}$. Simple FFT improves steadily up to 3-4 samples per period. However, neither reaches the accuracy of the good methods. The optimum of 20–30 samples per period may depend on the defects present in the data. Nevertheless, it seems quite consistent over two orders of magnitude of N. It also agrees with the study of scanning speed influence [23] where the profile length was kept fixed and N decreased with increasing scanning speed. The variance of results did not change much for more than 20 samples per period, but it started to increase sharply when less than 20 points were measured (one needs to combine tables 1 and 2 in [23] to compute N/P).

To summarise, measuring more periods is better than measuring each position more precisely — provided the sampling does not become too coarse and some other error does not grow too large. Still, the optimum settings can be quite counter-intuitive. This is a similar situation to roughness measurements where a measured area which 'feels right' is often way too small [16, 33]. If we are limited by the maximum profile length, measuring with a shorter sampling step is still useful, but the precision gain is slower.

3.4. Methods performance on two-dimensional gratings

The overall accuracy of the two lattice vectors \mathbf{a}_1 and \mathbf{a}_2 was measured as

$$\delta^2 = \frac{|\hat{\mathbf{a}}_1 - \mathbf{a}_1|^2}{|\mathbf{a}_1|^2} + \frac{|\hat{\mathbf{a}}_2 - \mathbf{a}_2|^2}{|\mathbf{a}_2|^2},\tag{7}$$

a natural extension of the 1D relative error to the 4D space formed by \mathbf{a}_1 and \mathbf{a}_2 .

Results for 2D gratings are plotted in figure 7 for combined random artefacts (noise + waviness + uneven + tilt). The scaling exponents for all four good methods are around two. They vary slightly, but this is expected based on the 1D results. The abscissa in figure 7a is the linear image size in pixels instead of number of periods P



Figure 7. Scaling and optimisation for 2D gratings: (a) accuracy scaling with image size, with estimated scaling exponents listed in the legend; (b) accuracy as a function of the number of measured periods for fixed image size 1000×1000 pixels.

used in figure 5. An effective P can be defined

$$P = \sqrt{\frac{|\mathbf{a}_1 \times \mathbf{a}_2|}{h_x h_y}}, \qquad (8)$$

where $h_x h_y$ is the area of one pixel with sides h_x and h_y . In the typical case $|\mathbf{a}_2| \approx |\mathbf{a}_1|$ and $h_y = h_x$. However, image size is easier to imagine.

Following the analysis of 1D scaling, the explanation of the scaling exponent is simple. Two vectors need to be determined. The accuracy of each vector scales with $P_{\rm along}^{3/2}$ for the number of periods along its direction $P_{\rm along}$. It also scales with $P_{\rm across}^{1/2}$ for the number of periods across its direction $P_{\rm along}$ because sufficiently distant profiles behave like independent measurements. For a square image $P_{\rm along} \approx P_{\rm across} \approx P$. Therefore, the overall accuracy scales with

$$P_{\rm along}^{3/2} P_{\rm across}^{1/2} \approx P^{3/2} P^{1/2} = P^2 , \qquad (9)$$

which is proportional to the number of image pixels N^2 . Scaling with P^2 is again much better than with $\sqrt{P^2} = P$ which would follow from simply measuring more data. In higher dimensions D the expected scaling is with $P^{D/2+1}$. In the 1D case we commonly saw mean relative errors below 10^{-6} with 50 000 samples, whereas here we do not reach them even with $2000 \times 2000 = 2\,000\,000$ image pixels (for comparable noise levels). This is a direct consequence of the smaller number of periods in images and slower scaling



Figure 8. Distributions of relative errors of 2D lattice vector \mathbf{a}_1 along the vector and in the perpendicular direction.

with P. The error (7) is also larger because it is the total relative error of four vector components, instead of a single parameter T.

These observations together support the measurement strategy which converts 2D evaluations to 1D evaluations [17]. The two lattice vectors are first found using a 2D measurement. They are then improved by measuring long thing stripes along each vector and evaluating them separately as this maximises the number of periods measured along each vector. Of course, this strategy can only be realised with a long-range metrological AFM.

As in one dimension, we can ask what is the optimal number of samples per period, measured using the effective P defined by equation (8)? The result, analogous to figure 6, is plotted in figure 7b for images with 1000×1000 pixels. The overall behaviour is similar to the 1D case. The optimal number of samples per period for the good methods is lower than in 1D, about N/P = 10. A possible reason is that the number of data points per grating unit is $(N/P)^2$, not N/P.

Even though the generated hole patterns were always perfectly orthogonal, the evaluation did not impose any orthogonality constraint. The measured angle between \mathbf{a}_1 and \mathbf{a}_2 could thus deviate from 90°. In figure 8 we can see the 2D error distribution of \mathbf{a}_1 for 500 × 500 images. The simple ACF method underestimates the length of \mathbf{a}_1 slightly and the error distribution for simple FFT resembles more a uniform distribution in square than a Gaussian distribution. Distributions for the good methods are isotropic and the errors of angle and length can be thus estimated using the simple error propagation rule.

2D measurements are also used for 1D structures. In principle we then have two options, process the entire image using a 2D method, or extract a set of profiles and evaluate them using a 1D method. Only frequency domain methods and model fitting work identically as in the pure 1D and 2D cases. The rest are not directly applicable to images of 1D structures. As for profiles, they have to be taken along the lattice vector to prevent cosine errors, creating a chicken and egg problem since the lattice vector is what we are trying to determine. Even though the vector can be estimated using 2D FFT, if we are to compute 2D FFT a sensible strategy is to employ a refined FT method in 2D instead of returning back to profile extraction.

The analysis of scaling with image size remains unchanged from the pure 2D case. Now there is only one lattice vector, but the analysis considered each vector separately anyway. One may think that a 1024×1024 image would be equivalent to 1024 independent profiles, so the standard deviation would be reduced by factor $\sqrt{1024} = 32$ compared to a single 1024-point profile. Although this can be the case, often such estimate is too optimistic. Individual scan lines and artefacts in them, such as line roughness, can be highly correlated and each line thus adds less independent information than the simple estimate suggests. This is analogous to roughness measurement which is plagued by the same problem [16, 33, 34].

3.5. Small number of periods

Two main reasons for measuring a small number of periods are instrumental limitations, *i.e.* scanner range, and the preconception that it is necessary to measure the grating in fine detail for accurate results. We hope this work helps to dispel the latter, but the former is much harder to deal with. Measurements of short profiles/small areas are, and will be, common as most AFM scans are limited to $100 \,\mu\text{m}$.

A profile cannot be shorter than one T (P = 1) to measure the period. The manual method or ZC in principle require profiles only slightly longer than T, whereas GC needs profiles longer than 3T/2 (P = 3/2) to find two gravity centres of the same type. FFT methods are limited by the ability to distinguish the correct peak from the one at origin. Model fitting is interesting, in particular in 2D, by its ability to utilise information which is not along the two lattice vectors. The image in figure 1b would not be the best measurement of the grating, but despite being 'too small' and not accommodating measurements along the two lattice vectors, it can still be easily evaluated by model fitting (admittedly, the example is a bit contrived in order to illustrate the point). Piecewise fitting and cross-correlation both require profiles longer than 2T and possibly up to 3T, depending on the phase. Therefore, their usefulness for short data is limited and were not considered here.

1D simulation results are shown in figure 9. The profile always had 1000 samples while T and P were varied. The additive noise was relatively low, 0.7%. Since direct space methods may not find any usable points, the figure shows the accuracy $\delta_{\rm rms}$ for successful evaluations and also the success rate. Success was defined as (a) the method itself did not report failure, and (b) the result was not an obvious outlier.

As predicted, the success rate of GC drops rapidly below P = 2, whereas other methods can work closer to P = 1. Frequency domain methods never failed according to the criterion, but of course their accuracy is poor for a small P. The manual method appears to work well up to about P = 5, at least in the low-noise case when two points suffice to find the intersection precisely. It is, in essence, a worse version of ZC, but for



Figure 9. Results for small number of periods.

just a couple of periods they can behave similarly. Still, already in the range P = 5 to 10 it is clear that the accuracy of the manual method remains constant, whereas the good methods start to scale with $P^{3/2}$ (as in figure 5). We must stress that even in this case all the methods were most accurate for the *coarsest* sampling (large P), not the finest sampling step.

Representativeness is a major concern when only a couple of features are measured. For periodic structures such as those in figure 1c or 1f the position and shape variation of individual features can be much larger than the precision with which we can measure the lattice vectors. Measurements at many different locations and statistical analysis are then essential.

3.6. Error estimates

Figure 5 demonstrated that different artefacts with similar noise to signal ratios can result in order-of-magnitude differences in accuracy. It is, therefore, difficult to assign a universally valid accuracy estimate to each method. Frequency-domain methods do not directly give error estimates, but direct-domain and hybrid do.

GC, ZC multi-peak ACF, correlation search and piecewise fitting employ linear regression to obtain \hat{T} . In most the individual values can be considered independent and have finite variances. Therefore, the estimate of result standard deviation $\hat{\sigma}_T$ from linear regression should be usable. In ACF all the peaks are computed from the same data, so the independence assumption may not be justified.

Model fitting using non-linear least squares also gives standard deviation estimates $\hat{\sigma}_T$. They should be correct if data are disturbed by uncorrelated noise, although this



Figure 10. Comparison of mean square relative error estimated from the method with the true error. The dashed line corresponds to the errors being equal.

is rarely the case. For large correlated artefacts the situation is more complicated and the standard deviations can be severely underestimated. They are also affected by systematic differences between the model and experimental data.

The estimates provided by these four good methods were compared with the true error. The simulation was run for 1D profiles (4800 samples), T of 50 pixels, and disrupted by the combined artefact (noise + waviness + placement). The noise to signal ratio was varied over a few orders of magnitude to obtain different $\delta_{\rm rms}$ and for each ratio 12 000 grating instances were evaluated.

The results are plotted in figure 10. For all the methods the mean estimated relative standard deviation matched the true $\delta_{\rm rms}$ quite well. Only multi-peak ACF overestimated the error somewhat when it became small. For a large number of degrees of freedom the quantity $(\hat{T} - T)/\hat{\sigma}_T$ should be normally distributed. Inspection of the distributions showed that even though this was not entirely true, the distributions were not far from standard normal. The estimates from individual methods thus seem usable, with caution. Comparison of several different methods can also be helpful [23], despite not being directly usable for standard deviation estimates if we do not know how the errors are correlated.

Usually the accuracy is limited by artefacts in the experimental data and calibration uncertainties. However, one can also ask if the evaluation method can ever become the limiting factor and at what accuracy level? The two simple methods are obviously limited by frequency and sampling steps. For the good methods the answer is more interesting. It can be, unfortunately, also implementation-dependent. In order to investigate these intrinsic errors our implementations were run on ideal data. With no defects, two parameters remained to be chosen at random, the true period T (within 5% of nominal value, as usual) and grating phase.

The results are summarised in table 2. Simple FFT and ACF behave as expected. For 1D data, ZC and both types of least-squares fitting could achieve more or less the full double precision, *i.e.* they were limited by rounding errors. Interestingly, neither GC nor cross-correlation achieved the same precision, most likely because of subtle discretisation errors. They have been studied in detail under different assumptions than

Method	Profile 480	Profile 4800	Profile 48 000	Image 250	Image 1000
Simple FFT Dai05 FT Zoom FFT	3.1×10^{-2} 3.4×10^{-5} 3.4×10^{-5}	3.0×10^{-3} 1.1×10^{-6} 1.1×10^{-6}	$\begin{array}{c} 3.0\times 10^{-4} \\ 3.1\times 10^{-8} \\ 2.5\times 10^{-8} \end{array}$	$\begin{array}{c} 1.4 \times 10^{-2} \\ - \\ 6.1 \times 10^{-5} \end{array}$	3.6×10^{-3}
<mark>Simple</mark> ACF Multi-peak ACF	3.3×10^{-4} 2.5×10^{-4}	1.3×10^{-4} 4.2×10^{-6}	1.2×10^{-4} 6.1×10^{-8}	$\begin{array}{c} 1.0 \times 10^{-2} \\ 6.3 \times 10^{-5} \end{array}$	4.6×10^{-3} 5.3×10^{-6}
Gravity centre Zero crossing Model fitting Piecewise fitting Cross-correlation	$\begin{array}{c} 2.6\times10^{-5}\\ 2.8\times10^{-16}\\ 9.9\times10^{-17}\\ 5.0\times10^{-16}\\ 4.6\times10^{-5} \end{array}$	$\begin{array}{c} 6.6 \times 10^{-7} \\ 7.0 \times 10^{-16} \\ 7.8 \times 10^{-17} \\ 9.8 \times 10^{-16} \\ 8.4 \times 10^{-7} \end{array}$	$\begin{array}{c} 1.6\times10^{-8}\\ 2.1\times10^{-15}\\ 8.1\times10^{-17}\\ 3.0\times10^{-15}\\ 2.7\times10^{-8} \end{array}$	8.9×10^{-5}	$ \begin{array}{c} 1.1 \times 10^{-5} \\ $

Table 2. Upper accuracy limit of individual methods, measured as $\delta_{\rm rms}$ for ideal data. Profiles had nominal period 50 samples; images 25 pixels.

hold in SPM [35], but are present also here. Other good methods also behave more or less similarly to GC. Peak apex locations in frequency domain methods can be subtly affected by windowing. For 2D data, model fitting did not achieve the same precision as in 1D and was not even improving much with image size. Although a portion of fit results was exact or within rounding errors, not all were — unlike in 1D. The reasons are not clear; possibly the Gwyddion fitter has convergence problems for huge data sets.

4. Good and bad practices

About 15 methods were implemented (for 1D and 2D) and were run hundreds of thousands of times on a variety of data, ranging from one period to thousands, and with different simulated artefacts. This enabled us to draw more general conclusions and remark on the merits and pitfalls in comparison to what would be achievable from the evaluation of a small set of measured gratings.

4.1. Evaluation of a sequence of points

When one has a sequence of key points on the profile $x_1, x_2, x_3, \ldots, x_n$ (for instance zero crossings) it is tempting to compute the distances $x_2 - x_1, x_3 - x_2, \ldots, x_n - x_{n-1}$, and average them. This would be counterproductive because only the first and last positions contribute to the average (as already noted in [17]):

$$\hat{T} = \frac{1}{n-1} \sum_{i=1}^{n-1} (x_{i+1} - x_i) = \frac{x_n - x_1}{n-1} .$$
(10)

Instead we have to fit the sequence with a straight line $x_i = iT + c$. A least squares fit gives the best unbiased linear estimate of T for homoscedastic x_n . Nevertheless, it still gives much more weight to points close to the edges. The effective weight is proportional to the distance from the centre. Weighted fit should be considered if data close to the

edges can have larger errors. This is in an interesting contrast with frequency domain methods. Although DFT itself acts uniformly, windowing suppresses data close to the edges, giving larger effective weight to data in the centre.

4.2. Resampling

It has been suggested to interpolate the data to K times larger number of points, with K possibly being as high as 20 [17]. However, there is no theoretical reason for it. The added data do not bring any new independent information, meaning it cannot improve evaluation methods based on fitting. The computation of quadratures, intersections with zero or locations of maxima already gives answers corresponding to a continuous function interpolating the discrete data. Therefore, unless the computation utilises discrete operations which cannot be easily expressed in this manner, preinterpolating the data is pointless.

Resampling was also suggested for frequency-domain methods [17]. This technique is sometimes called enhanced DFT (as opposed to refined). It is a waste of computational resources and not recommended. As both interpolation and DFT are linear operations, the final result of the procedure is a summation over the original data z_n as in eq. (3) but with modified weights. If the signal is sampled densely (satisfying the conditions of Nyquist–Shannon sampling theorem [36, 37]) then a perfect interpolation exists, the Whittaker–Shannon formula [36, 37]. If we simultaneously consider $K \to \infty$, *i.e.* a very fine subdivision, equivalent to analytical evaluation of integrals, the final result simply reproduces the DFT. Worse interpolations and less precise quadratures will deviate from it (generally Z_{ν} would be multiplied by a slowly varying function of ν), but that does not constitute an improvement. On the other hand, if the signal is undersampled then interpolation generally is not justified. It certainly cannot magically correct aliasing we have to measure with a shorter sampling step.

Only ACF-based methods can benefit from resampling — moderately. They are direct space methods but average over many periods and resampling can help locating the maxima in the averaged data slightly more precisely. ACF-based methods have not been utilised in other works.

The theoretical conclusions are demonstrated numerically for 1D data in figure 11a (N = 4800, 50 samples per period, combined noise). We calculated the improvement achieved by resampling K times for all the implemented methods. Only a few curves actually differ visibly from improvement factor of 1, *i.e.* no improvement at all. Only the two ACF methods and data fitting differ systematically. Simple ACF shows the largest improvement, but of course from a poor base accuracy. Data fitting seems to be affected negatively. In the end, multi-peak ACF is the only method for which we might suggest resampling if greatly increasing computation time for a moderate improvement is an acceptable trade-off.



Figure 11. (a) Accuracy improvement achieved by resampling the data to K times larger number of points (most curves are indistinguishable from 1); (b) Scaling law for experimental data; (c) reduction of accuracy in GC caused by moving the threshold away from the mid-height; (d) aliasing effects in waviness removal by linear filtering.

4.3. Refined Fourier transform

Fourier transform refinement is the evaluation of DFT expression (3) for non-integer ν (sometimes called a bit confusingly 'FFT-FT'), as already briefly introduced in section 2.5. It enables a more precise location of peaks in the power spectrum and thus more precise measurement of T [17,23,28,38]. It has been usually implemented using brute force computation [17,28]. The Goertzel algorithm [39] can compute Fourier coefficients one by one more efficiently in such case [23]. However, the computation of each coefficient still costs O(N) operations.

The now standard FFT-based refinement method is Zoom-FFT, a specialisation of the Chirp z-transform (CZT) [40–42], based on Bluestein's algorithm [29]. The algorithm computes $Z_{\beta} = \sum_{n} z_{n} s_{\beta}^{-n}$ for a geometric sequence of complex numbers s_{β} , $\beta = 0, 1, 2, \ldots, M$ (usually s is denoted z, in line with the name z-transform, but it clashes with our z coordinate). It does so by expressing the transform as a convolution and computing the convolution efficiently by FFT, utilising the convolution theorem. In total three FFTs of size O(M + N) are needed to compute the M values of Z_{β} . If we chose $s_{\beta} = \exp(-2\pi i\beta/M)$, *i.e.* uniformly covering the unit circle, we would recover DFT. But other choices are possible. In particular, we can cover only a small segment on the unit circle. This special case is called Zoom FFT because it zooms into a small interval of frequencies (although some other refinements methods are also called Zoom FFT).

Even though it may seem counter-intuitive, generally it is not recommended to locate the peak maximum using a smart search method if it comes with costly evaluation of each spectral density value. It can be efficient if the search is guaranteed to converge in $O(\log N)$ steps. However, one N/2 times refinement around the coarse peak and a simple scan for the maximum is much more straightforward. A parabolic refinement of the maximum can be added as a final step. Zoom FFT takes only a couple of extra FFTs and is commonly available as a function in numerical software and libraries. In higher dimensions it is more efficient to zoom twice by \sqrt{N} instead of once by N, but in 1D there is no benefit.

Finally, we note that refined FT has been presented as substantially different from DFT [17] and it was even suggested that it cannot be done by interpolating in the frequency domain [23]. However, this is misleading (the connection and interpolation procedure are discussed in Supplementary Information S2).

4.4. Zero line selection

Zero line selection in GC and ZC attracted considerable attention in existing works. It appears as a tunable parameter [17] or is located in a sophisticated manner [23]. When only the upper halves are utilised it was suggested the zero line should be below the profile mid-height as larger integrated areas reduce the relative errors slightly [26]. Yet the symmetrical GC should be rather insensitive to the zero line height. The accuracy is symmetrical with respect to the optimum which occurs around the mid-height (the exact optimum is slightly different for each grating profile) and its first derivative by zero line height is zero at the optimum. Hence, as long as the zero line is approximately correct the accuracy should stay basically the same. This conclusion can already be made from [26] and is confirmed by numerical results illustrated in figure 11c. We see that the zero line has to be moved quite far from the optimum for a substantial reduction of accuracy. It should be kept roughly around mid-height, but no great accuracy is necessary.

A similar path of reasoning can be followed for ZC. Nevertheless, we found it more sensitive to the zero line level. In addition, if the zero line is far from mid-height ZC, choosing a good segment around the crossing to fit can become more involved. The zero line was found similarly to Ortlepp *et al.* [23], *i.e.* by locating the two main peaks in height distribution and choosing the midpoint. However, this was mainly for the sake of simplicity in ZC implementation. Subtraction of the mean value was sufficient for the other methods.

4.5. Background subtraction

Subtraction of a slowly varying background on the substrate (waviness) from oscillatory data is not trivial. A moving average [17] can create a wavy pattern because the number of samples averaged is an integer, but the period is not. Sometimes the upper part of

the profile contributes more to the average, sometimes the lower part, and this varies along the profile. The same conclusion can be made from analysis in the frequency domain [23]. One instance of the effect is illustrated in figure 11d. It can clearly distort an already perfectly levelled profile.

A Butterworth filter [43], which approximates a rectangular frequency-domain filter, was suggested as a replacement because it has monotonous frequency response [23]. Unfortunately, it is not a good choice either because of its poor response to edges, where it exhibits a considerable overshoot and ringing. Furthermore, if the data processing is off-line and probably involves FFT anyway, there is no reason to limit it to filters originating in classical signal processing such as Butterworth. Filtering can be done in the frequency domain by modifying the Fourier coefficients, even using a perfect rectangular high-pass filter, for instance. However, it would not entirely solve the poor edge response. A possible result of rectangular filter processing is also illustrated in figure 11d.

Other linear filters, such as Gaussian or Bessel, have a better response. However, any linear filter is just multiplication by some function in the frequency domain and involves trade-offs between not disturbing the profile shape and removing waviness on sufficiently short length scale. In our opinion non-linear filtering may be necessary. Empirical mode decomposition (EMD) [44] has already been proposed to improve grating evaluation [38]. The highly non-linear envelope method used here to remove the background somewhat resembles one step of EMD. However, it extracts the low-frequency background, not the highest frequency component. Its key feature is that it preserves an ideal rectangular wave exactly. However, selection of the optimum filter requires further investigation.

4.6. The lock-in method

A lock-in method was also suggested for grating evaluation [23]. We did not implement and do not recommend it because it is basically a worse version of the refined FT. It proceeds as follows

- (i) Multiply the measured data by the model response a sine or rectangular wave is used.
- (ii) Compute the average value. This is described in a somewhat complicated manner as 'low-pass filtering', but the end result is the mean value of the multiplied data.
- (iii) Find the maximum of this average over a domain of model response parameters, period T and phase φ .

For a sine wave the first two steps are equivalent to the computation of a Fourier coefficient. The quantity to maximise is

$$\sum_{n=0}^{N-1} z_n \cos\left(2\pi \frac{hn}{T} + \varphi\right) = \Re\left(e^{i\varphi} Z_{\nu}^*\right) , \qquad (11)$$

0.2 ms

100 ms

Table 3. Typical execution time of various methods, single-threaded on a	standard
PC. Simple FFT and ACF are grouped under 'Simple'; more precise FFT,	ACF and
FFT-based correlation search methods under 'ACF & FFT'; and both fitti	ng under
'Fitting'.	
Data GC & ZC Simple FFT & ACF Dai05 FFT Fitt	ing

5 ms

 $0.5 \ s$

30 ms

300 ms

 $10 \mathrm{s}$

2 ms

40 ms

FT.

4.7. Robustness and speed

profile, 50k points

image 1000×1000

One practical concern is evaluation speed, in particular in a high-throughput context. Execution times reported in [17] may seem worrying, even considering the advances in computer performance. For our implementations in C using FFTW [45] and Gwyddion [30] libraries we can fortunately report much more encouraging data, summarised in table 3. The straightforward direct-domain methods, GC and ZC, can evaluate 50 000 samples long profiles in a fraction of a millisecond. In 2D the two simple methods were fastest, but they are not sufficiently accurate. The next fastest was again GC. Model fitting was the slowest and could take over 10 s for a 1000×1000 image. The Gwyddion fitter has a large overhead so the fitting execution time may not be entirely representative. Still, the number of arithmetic operations per data point is invariably high in non-linear least-squares fitting.

Perhaps an even more important property is method robustness, *i.e.* ability to behave correctly for a wide range of input data. In the simulations all methods were run hundreds of thousands times on generated data without human intervention, demanding perhaps a bit more robust implementations than is typical. With FFT and ACF based computations this was easy to achieve since they work with transformed ('summary') data and are insensitive to local defects. If there is a peak where the algorithm is looking for it, it is found and measured correctly. This makes them very reliable.

Model fitting is decidedly less robust. Its known Achilles' heel is initial parameter

estimation and the existence of multiple local minima of the sum of squared residuals. Simple FFT was used for initial estimation of T and a scalar product similar to the lockin method (section 4.6) for phase estimation. Such an estimate is still insufficient for long gratings — when the estimated T is 1% off the model gets completely out of phase after 50 periods. Therefore, only several periods were initially fitted and the fitted segment was increased in a geometric progression until the entire data were covered. Multiple local minima did not seem a major problem with long periodic data, such as gratings. Although fitting can definitely fail when the initial estimate is not sufficiently close, we observed more or less only two possible outcomes. It either succeeded and gave an accurate T value — or failed rather obviously. We should also mention that model fitting does not work without a reasonable parametric description of the data. Gratings are relatively simple to describe, but other periodic structures may be more complicated. Evaluation of a different type of sample may require the construction of a new model.

Bootstrapping was necessary also in other methods. All tested background removal procedures require at least an approximate *a priori* knowledge of the period. Furthermore, any refined FT method starts from a coarse estimate. Simple FFT is reliable and serves well for this purpose, unless a very small number of periods is measured. In such case the profile may have to be zero-extended (after windowing) to two or three times the size to increase the frequency resolution.

GC, ZC and piecewise fitting require the most care to work reliably. Particles and other local defects can lead to incorrect local fits and bar gravity centres [17]. When N/P is high, noisy data can cross the threshold more than once, creating very short segments which need to be filtered out [23]. Our GC implementation used a twostage filtering. Most incorrect segments come from multiple threshold crossings and are too short, whereas too long are rare. Therefore, the first stage computed the 90 th percentile of segment lengths and segments more than $10 \times$ shorter were discarded (and segments that were too long). The second stage found the median inter-centre distance and only kept centres whose distance to the closest good centre was close to an integer, avoiding possible problems illustrated in figure 13 in [17]. The initial set of good centres was identified as three consecutive centres with distances close to the median. ZC points were processed in the same manner, just without the first stage because an entire region around a crossing is fitted and so each crossing only gives one value of x_0 . This approach made the procedures self-contained. If a FFT-based estimate of T is available, the filtering can be simplified. A modification of random sample consensus (RANSAC) [46] may also be suitable. In both the GC and the ZC cases, the positive and negative features were separately fitted and their average taken in order to reduce the sensitivity of the methods to the choice of threshold.

GC and ZC employ various thresholds, interval/point filtering parameters, fitting ranges and other similar algorithm tunables. All could be made user-controllable, as was suggested for the thresholds [17, 26]. One has to resist the temptation do so. In addition to the usual problems that follow [47] we noticed one specific to highly accurate
measurements. It is not difficult to make the evaluation accuracy worse by an order of magnitude or two by a poor parameter choice (or a subtle implementation error). What may be difficult for the user, is spotting that the relative error jumped from 3×10^{-7} to 3×10^{-5} , even in a situation when the former is negligible and the latter a major error source. Extensive numerical verification can ensure a method achieves the accuracy it should when used for samples within defined parameters. With half a dozen tunables the user can tweak freely this becomes impossible. From this point of view refined FT based methods are preferable.

5. Experimental example

To demonstrate the validity of discussed scaling also on experimental data, we measured a commercially available grating VGRP-UM from Bruker using self-sensing Akiyama probe, a custom built electronics and Nanopositioning and Nanomeasuring Machine (NMM) [48]. Use of NMM enabled measurement to be made over large area, in a similar way to that used to make the most accurate grating measurements that have been reported in the literature [18].

In order to obtain data suitable for validation of the scaling law, the measured primary profile was split into several independent segments and each was evaluated separately. The spread of obtained values was characterised using an estimated standard deviation (with Student coefficient), divided by the estimated value to obtain a quantity similar to $\delta_{\rm rms}$ in the simulations. The results were averaged for twelve independent measured profiles. Since the spread of values, not differences from the correct value, were considered, systematic errors did not enter the analysis. This meant the standard deviations should follow the $P^{3/2}$ scaling law, which holds for random errors, and scale with inverse 3/2-th power of the profile length.

The results are plotted in figure 11b for the eight good methods. The two simple methods, strongly affected by discretisation, are not shown. The scaling rule holds quite well for short segments as data follow the lines corresponding to the 3/2-th power. The scaling exponents are in fact somewhat higher than 3/2, like for waviness in section 3.2, and for the same reason. For long segments the curves start to deviate and the spread seems to almost saturate, decreasing much more slowly if at all. This apparent violation of the power scaling is caused by manufacturing errors (a more detailed analysis is in Supplementary Information S3). In summary, the general conclusions about behaviour of the data evaluation methods, which were derived from simulations and theoretical reasoning, apply also to processing of experimental data. The apparent violation of the power scaling serves as a reminder that odd and distinctive properties of the sample and measurement system must be always considered — in addition to the general principles studied here.

Method	Advantages	Disadvantages
Manual	easy to understand	poor accuracy upper limit poor representativeness user-dependent laborious
Simple FFT	insensitive to local defects insensitive to background insensitive to odd shapes relatively fast	poor accuracy no error estimate
Refined FT	good accuracy insensitive to local defects insensitive to background insensitive to odd shapes	no error estimate
Simple ACF	insensitive to local defects insensitive to background relatively fast	poor accuracy no error estimate sensitive to shape asymmetry
Multi-peak ACF	good accuracy insensitive to local defects insensitive to background provides error estimate relatively fast	sensitive to shape asymmetry relatively low accuracy upper limit
Gravity centres	good accuracy provides error estimate fast	difficult to make robust depends on background subtraction
Zero crossings	good accuracy provides error estimate high accuracy upper limit fast	difficult to make robust depends on background subtraction only one-dimensional
Model fitting	good accuracy provides error estimate high accuracy upper limit possibly multiple parameters	requires a model depends on parameter estimation sensitive to local phase slow
Piecewise fitting	good accuracy provides error estimate parameters for each unit high accuracy upper limit	requires a model depends on parameter estimation difficult to make robust slow
Cross-correlation	good accuracy provides error estimate relatively fast	requires a template

 Table 4. Summary of advantages and disadvantages of various methods.

6. Conclusion

In this paper we report on the use of both direct space and Fourier space based methods for evaluating the periodic structure parameters. In addition to gravity centres, refined FT and cross-correlation methods described in the ISO standard, this includes zero crossing, multi-peak ACF and global and piecewise model fitting. We concluded that these seven evaluation methods can be recommended, if implemented properly. Overall they all behave similarly, although some differences in sensitivity to various artefacts exist; see table 4. Importantly, their accuracy scales super-linearly with the number of periods P, typically with $P^{3/2}$ for profiles and P^2 for images. As a side effect of this analysis, more can be said about the overall benefits and drawbacks of different methods, as also shown in table 4.

There are also more general conclusions and recommendations that can be used when designing the experiment and processing the measured data, these are explained in the previous text and are summarized here:

- Since all the good methods are sub-pixel, pixel size/sampling step is not the limiting factor for accuracy.
- Although the ISO standard recommends measuring more than 5 or 7 samples per period, ideally at least 20 points should be measured. If at least 20 pixels per period can be measured, measure as many periods in the profile as possible.
- Resampling measured data offers no benefits.
- Accuracy scaling with the 3/2-th power of number of periods is quite difficult to beat using alternative measurement strategies. The accuracy scales only linearly with decreasing sampling step and only with the square root of the number of repetitions.
- The GC and ZC methods readily provide an estimate of the statistical error.
- If these guidelines are followed, the uncertainty contribution from the numerical procedure used is insignificant compared to the other uncertainty components related to measurement (at least in the case of a standard SPM).

As a simple example of applying the above recommendations, for most typical gratings used in a commercial AFM calibration, which are $1-5 \,\mu\text{m}$ pitch, use full range of the microscope (typically 100 μm) as long as the positioning errors of the scanning stage are not significant at the periphery. These settings should provide 20–100 periods. Keeping at least 20 pixels per period then means collecting about 2000 pixels per line, which is achievable by standard SPMs.

It is hoped that the results presented in this paper will give users confidence in evaluating periodic structures and calibrating SPMs.

Acknowledgements

This work was supported by Technology Agency of the Czech Republic project No. TN01000008 and by the project 20IND08 MetExSPM that received funding from the

EMPIR programme co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme. Part of this work was funded by the National Measurement System Programme from the Department of Business Energy and Industrial Strategy, UK.

References

- [1] Kim S, Moon D, Jeon B R, Yeon J, Li X and Kim S 2022 Nanomaterials 12 1542
- [2] Li J, Li S, Higashi T, Kawai K, Inagaki K, Yamamura K and Arima K 2021 Phys. Rev. B 103(24) 245433
- [3] Yacoot A, Bosse H and Dixson R 2020 Meas. Sci. Technol. 31 121001
- [4] Kondratov A, Rogov O and Gainutdinov R 2017 Ultramicroscopy 181 81-85
- [5] Romano S, Lamberti A, Masullo M, Penzo E, Cabrini S, Rendina I and Mocella V 2018 Materials 11 526
- [6] Xiong L, Forsythe C, Jung M, McLeod A S, Sunku S, Shao Y, Ni G, Sternbach A J, Liu S, Edgar J, Mele E J, Fogler M M, Shvets G, Dean C and Basov D 2019 Nat. Commun. 10 4780
- [7] Panfilova E, Syritskii A and Ibragimov A 2019 IOP Conference Series: Materials Science and Engineering 699 012034
- [8] Alegre T P M, Safavi-Naeini A, Winger M and Painter O 2011 Opt. Express 19 5658–5669
- [9] Schmidt R H 2012 Philos. T. Roy. Soc. A 370 3950–3972
- [10] Peterson J R, O'Connor P, Nomerotski A, Magnier E, Jernigan J G, Cheng J, Cui W, Peng E, Rasmussen A and Sembroski G 2020 arXiv:2001.03134 [astro-ph.IM]
- [11] Lüthi M, Bircher B A, Meli F, Küng A and Thalmann R 2019 Meas. Sci. Technol. 31 035002
- [12] Li X, Gao W, Muto H, Shimizu Y, Ito S and Dian S 2013 Precis. Eng. 37 771–781
- [13] Mosier-Boss P A 2017 Nanomaterials 7 142
- [14] ISO 11952:2019 2019 Surface chemical analysis scanning probe microscopy determination of geometric quantities using SPM: Calibration of measuring systems
- [15] Yacoot A, Lawn M, Korpelainen V, Dongmo H, O'Neill J, Hermann J, Miles J, Coleman V, Seppä J and Lassila A 2020 Microscopy and Analysis 34 S4–S9
- [16] Nečas D, Valtr M and Klapetek P 2020 Sci. Rep. 10 15294
- [17] Dai G, Koenders L, Pohlenz F, Dziomba T and Danzebrink H U 2005 Meas. Sci. Technol. 16 1241–1249
- [18] Dai G, Pohlenz F, Dziomba T, Xu M, Diener A, Koenders L and Danzebrink H U 2007 Meas. Sci. Technol. 18 415–421
- [19] Nečas D and Klapetek P 2021 Nanomaterials 11 1746
- [20] Gwyddion developers Gwyddion http://gwyddion.net/
- [21] Misumi I, Gonda S, Kurosawa T and Takamasu K 2003 Meas. Sci. Technol. 14 463
- [22] Huang Q, Misumi I, Gonda S, Sato O and Kurosawa T 2006 Pitch calibration of one-dimensional grating standard by tapping mode nanometrological atomic force microscope Sixth International Symposium on Instrumentation and Control Technology: Signal Analysis, Measurement Theory, Photo-Electronic Technology, and Artificial Intelligence (Proc. SPIE vol 6357) p 635730
- [23] Ortlepp I, Stauffenberg J and Manske E 2021 Sensors 5862
- [24] Chen J, Liu J, Wang X, Zhang L, Deng X, Cheng X and Li T 2017 Meas. Sci. Rev. 17 264–268
- [25] Misumi I, Gonda S, Kurosawa T, Tanimura Y, Ochiai N, ichiro Kitta J, Kubota F, Yamada M, Fujiwara Y, Nakayama Y and Takamasu K 2003 Meas. Sci. Technol. 14 2065
- [26] Huang Q, Gonda S, Misumi I, Keem T and Kurosawa T 2006 Proc. of SPIE 6280 628007
- [27] Dixson R, Chernoff D A, Wang S, Vorburger T V, Tan S L, Orji N G and Fu J 2010 Interlaboratory comparison of traceable atomic force microscope pitch measurements *Scanning Microscopy 2010* (*Proc. SPIE* vol 7729) p 77290M
- [28] Jørgensen J, Jensen C and Garnæs J 1998 Appl. Phys. A 66 S847–S852
- [29] Bluestein L I 1970 IEEE T. Acoust. Speech AU18 451

- [30] Nečas D and Klapetek P 2012 Cent. Eur. J. Phys. 10 181–188
- [31] More J J 1978 The Levenberg–Marquardt algorithm: Implementation and theory Lecture Notes in Mathematics vol 630 ed Watson G A (New York: Springer-Verlag) pp 106–116
- [32] Otsu N 1979 IEEE Trans. Sys., Man., Cyber. 9 62–66
- [33] Nečas D, Klapetek P and Valtr M 2020 Meas. Sci. Technol. 31 094010
- [34] Zhao Y, Wang G C and Lu T M 2000 Characterization of Amorphous and Crystalline Rough Surface – Principles and Applications (Experimental Methods in the Physical Sciences vol 37) (San Diego: Academic Press)
- [35] Landi G 2002 Nucl. Instrum. Meth. A 485 698-719
- [36] Shannon C E 1949 P. IRE **30** 10–21
- [37] Whittaker J M 1935 Interpolatory Function Theory (New York: Cambridge University Press)
- [38] Li C, Yang S, Wang C and Jiang Z 2016 Appl. Surf. Sci. 380 274–280
- [39] Goertzel G 1958 Am. Math. Mon. 65 34–35
- [40] Shilling S A 1970 A study of the chirp z-transform and its applications Ph.D. thesis Kansas State University Manhattan
- [41] Rabiner L R, Schafer R W and Rader C M 1969 IEEE T. Acoust. Speech AU17 86
- [42] Rabiner L R and Gold B 1975 Theory and Applications of Digital Signal Processing (New Jersey: Prentice-Hall)
- [43] Butterworth S 1930 Experimental Wireless and the Wireless Engineer 7 536-541
- [44] Huang N E, Shen Z, Long S R, Wu M C, Shih H H, Zheng Q, Yen N C, Tung C C and Liu H H 1998 Proc. R. Soc. A 454
- [45] Frigo M and Johnson S G 2005 Proceedings of the IEEE 93 216–231
- [46] Fischler M A and Bolles R C 1981 Commun. ACM 24 381–395
- [47] Nečas D and Klapetek P 2017 Meas. Sci. Technol. 28 034014
- [48] Hausotte T, Jäger G, Manske E, Hofmann N and Mastylo R 2004 Traceable nanometrology with a nanopositioning and nanomeasuring machine. Proc. International Symposium on Precision Mechanical Measurements (ISPMM'2004) pp 399–404

Výsledky optické charakterizace nehomogenních vrstev s drsnými rozhraními – závěrečná zpráva

Vypracovali: Mgr. Jiří Vohánka, Ph.D., prof. RNDr. Ivan Ohlídal, DrSc.

Příparava vzorku

Nehomogenní vrstva a drsnými rozhraními byla připravena plazmochemickou depozicí z plynné fáze na jednostranně leštěný křemíkový substrát se zdrsněným povrchem.

Strukturní model

V případě křemíkového substrátu se zdrsněným povrchem byla předpokládána následující struktura:

- Křemíkový substrát je v takřka celé měřené spektrální oblasti neprůhledný a jeho zadní strana je zdrsněna natolik, že veškeré prošlé světlo se rozptýlí. Proto byl pro popis substrátu užit model polonekonečného prostředí.
- Na povrchu substrátu byla předpokládána tenká nativní oxidová vrstva tloušťky $d_{\rm nol}$, která byla modelována jako tenká homogenní vrstva s optickými konstantami zvolenými v hodnotách odpovídajících amorfnímu SiO₂.
- Při popisu náhodné drsnosti byl užit model předpokládající identickou drsnost u všech rozhraní (t.j. rozhraní substrát/nativní oxidová vrstva a rozhraní nativní oxidová vrstva/okolí).

V případě křemíkového substrátu se zdrsněným povrchem a deponovanou vrstvou byla předpokládána následující struktura:

- Pro popis křemíkového substrátu byl užit model polonekonečného prostředí.
- U nehomogenní vrstvy bylo předpokládáno, že optické konstanty závisí na souřadnici z, jejíž hodnota určuje vzdálenost od spodního rozhraní nehomogenní vrstvy.
- Mezi substrátem a nehomogenní vrstvou byla předpokládána tenká přechodová vrstva tloušťky d_{tl} , která byla modelována jako tenká homogenní vrstva.
- Při popisu náhodné drsnosti byl užit model předpokládající identickou drsnost u všech rozhraní (t.j. rozhraní substrát/přechodová vrstva, rozhraní přechodová vrstva/nehomogenní vrstva a rozhraní nehomogenní vrstva/okolí).

Disperzní model

Optické konstanty nehomogenní vrstvy a přechodové vrstvy byly popsány pomocí Campi–Coriassova disperzního modelu [1]. Imaginární část dielektrické funkce je dána jako

$$\varepsilon_{\rm i}(E) = \frac{2N_{\rm vc}}{\pi E} \frac{B(E-E_g)^2 \Theta(E-E_g)}{\left[(E_c - E_g)^2 - (E-E_g)^2\right]^2 + B^2(E-E_g)^2},$$

kde E je energie fotonů, $N_{\rm vc}$ určuje sílu mezipásových přechodů, Θ je Heavisideova skoková funkce, E_g je energie šířky zakázaného pásu, E_c a B ($E_c > E_g$) jsou parametry určující polohu a šířku oblasti kde má $\varepsilon_i(E)$ maximum. Reálná část dielektrické funkce je pak dopočtena s užitím Kramers– Kronigových relací.

V případě nehomogenní vrstvy byla nehomogenita popsána s tak, že u jednotlivých disperzních parametrů byla předpokládána lineární závislost na souřadnici z. Hodnota parametru $p_{\alpha}(z)$ pro danou hodnotu souřadnice z je dána jako

$$p_{\alpha}(z) = p_{\alpha}^{\mathrm{L}} + \left(p_{\alpha}^{\mathrm{U}} - p_{\alpha}^{\mathrm{L}}\right) \frac{z}{d},$$

kde index α rozlišuje disperzní parametry, p_{α}^{U} je hodnota disperzního parametru u horního rozhraní, p_{α}^{L} je hodnota disperzního parametru u spodního rozhraní a symbol d značí střední tloušťku nehomogenní vrstvy. Z hodnot disperzních parametrů odpovídajících dané hodnotě souřadnice z pak byly vypočteny optické konstanty s užitím Campi–Coriassova modelu.

Optické konstanty nativní oxidové vrstvy byly zvoleny v hodnotách určených v [2] pro amorfní SiO₂. Pro křemíkový substrát byly užity optické konstanty určené v [3].

Započtení vlivu náhodné drsnosti

V případě odrazivosti měřené v blízké a viditelné oblasti a elipsometrických veličin byla pro započtení vlivu náhodné drsnosti rozhraní užita kombinace skalární difrakční teorie (SDT) a Rayleigh–Riceovy teorie (RRT). V případě odrazivosti měřené v ultrafialové oblasti tento přístup nevyhovoval protože RRT již nebylo možné kvůli vysokým výškám drsnosti vzhledem k vlnové délce užít. Pro započtení vlivu drsnosti byla v tomto případě užita pouze SDT. Pro interpretaci dat odrazivosti v ultrafialové oblasti bylo rovněž nutné zahrnout příspěvek od světla rozptýleného vzorkem a zachyceným detektorem spektrofotometru. To bylo rovněž provedeno s užitím SDT.

Kombinace SDT a RRT

Skalární difrakční teorie je užita pro započtení části drsnosti s nízkými prostorovými frekvencemi, zatím co Rayleigh–Riceova teorie je užita pro střední a vysoké prostorové frekvence.

Skalární difrakční teorie vychází z Kirchoffova–Helmholtzova integrálu. V případě drsných povrchů, vrstev a systémů vrstev s identicky drsnými rozhraními dostaneme následující vztah pro reflexní koeficient

$$r_q = \int_{-\infty}^{\infty} r_q^{\text{RRT}} e^{iv_z \eta} w(\eta) \mathrm{d}\eta = \chi(v_z) r_q^{\text{RRT}},\tag{1}$$

kde index q = p, s rozlišuje p a s polarizované vlny, $v_z = -4\pi n_0 \cos \varphi_0 / \lambda$, přičemž λ je vlnová délka světla, n_0 je index lomu okolního prostředí a φ_0 je úhel dopadu. Proměnná η reprezentuje hodnoty náhodné funkce popisující část drsnosti s nízkými prostorovými frekvencemi. Hustota pravděpodobnosti pro rozdělení výšek této části drsnosti je pak označena symbolem $w(\cdot)$ a odpovídající charakteristická funkce je označena $\chi(\cdot)$. Symbol r_q^{RRT} značí lokální koeficient odrazu, který již zahrnuje vliv drsnosti se středními a vysokými prostorovými frekvencemi. Při popisu drsnosti jsme užili předpokladu normálního rozdělení výšek s charakteristickou funkcí

$$\chi(v_z) = \exp\left(-\sigma_{\text{SDT}}^2 v_z^2/2\right).$$

U drsnosti s vysokými a středními prostorovými frekvencemi, které popisujeme pomocí RRT, se korelace mezi drsnostmi jednotlivých rozhraní uplatní pouze v případě velmi tenkých vrstev. Proto můžeme při výpočtu reflexního koeficientu r_q^{RRT} užít přístupu, kdy vliv drsnosti započteme při vyjádření koeficientů odrazu a průchodu horního a spodního rozhraní nehomogenní vrstvy. Koeficient odrazu odpovídající nehomogenní vrstvě pak vypočteme s užitím takto získaných koeficientů postupem, který bychom užili v případě nehomogenní vrstvy s hladkými rozhraními.

V případě koeficientu odrazu od horního rozhraní dostaneme

$$r_{\mathrm{U}q} = r_{\mathrm{U}q}^0 + \Delta r_{\mathrm{U}q}^{\mathrm{RRT}}, \qquad \Delta r_{\mathrm{U}q}^{\mathrm{RRT}} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_q(K_x, K_y) W(K_x - \bar{K}_x, K_y) dK_x dK_y, \qquad (2)$$

kde r_{Uq}^0 je koeficient odrazu odpovídající hladkému rozhraní a $\Delta r_{Uq}^{\text{RRT}}$ je korekce druhého řádu Rayleigh–Riceovy teorie. Integrace je provedena přes prostorové frekvence K_x a K_y , přičemž $W(K_x, K_y)$ je PSDF (power spectral density function) a $g_q(K_x, K_y)$ je komplikovaná funkce, která kromě K_x , K_y , závisí také na vlnové délce, úhlu dopadu a optických konstantách prostředí, ale nezávisí na drsnosti rozhraní. Symbol $\bar{K}_x = (2\pi/\lambda)n_0 \sin \varphi_0$ značí složku vlnového vektoru dopadající vlny, která je tečná ke střední rovině rozhraní. Vzorec (2) je pro koeficient odrazu a světlo dopadající z horní strany. Stejným způsobem, ale s jinak vyjádřenými koeficienty pro hladké rozhraní a odlišnými funkcemi $g_q(K_x, K_y)$, lze vyjádřit i koeficient průchodu a koeficienty odpovídající dopadu světla ze spodní strany. Detailní popis výpočtu lze najít v [4, 5]. Při popisu drsnosti jsme uvažovali PSDF danou Gaussovou funkcí

$$W(K_x, K_y) = \pi \sigma_{\text{RRT}}^2 \tau_{\text{RRT}}^2 \exp\left(-\frac{1}{4}\tau_{\text{RRT}}^2 (K_x^2 + K_y^2)\right),$$
(3)

kde $\sigma_{\rm RRT}$ je rms hodnota výšek dr
snosti se středními a vysokými prostorovými frekvencemi a
 $\tau_{\rm RRT}$ je autokorelační vzdálenost.

V případě spodního rozhraní nehomogenní vrstvy je třeba vzít v úvahu že koeficienty odrazu a průchodu musí zahrnout vliv tenké přechodové vrstvy. Jelikož přechodová vrstva je velmi tenká, nelze korelaci mezi drsnými rozhraními zanedbat a v rámci RRT je třeba užít vztahy odpovídající vrstvě s identicky drsnými rozhraními. Vztahy pro výpočet koeficientů odrazu a průchodu jsou analogické (2), avšak se složitějšími funkcemi $g_q(K_x, K_y)$, ve kterých kromě výše uvedených závislostí rovněž vystupuje tloušťka přechodové vrstvy. Detailní popis výpočtu lze najít v [4, 5].

Poté co byly vyjádřeny koeficienty odrazu a průchodu u horního a spodního rozhraní, byl koeficient odrazu nehomogenní vrstvy r_q^{RRT} vypočten pomocí metody popsané v [6], která je založena na WKBJ aproximaci a korekcích na vnitřní odrazy v nehomogenní vrstvě. Pro dosažení dostatečně přesných výsledků bylo potřeba započíst nejméně jednonásobný a dvojnásobný odraz uvnitř nehomogenní vrstvy.

Kombinace SDT se započtením příspěvku od rozptýleného světla

Celková odrazivost měřená spektrofotometrem je v tomto případě vyjádřena jako součet koherentní odrazivosti, která odpovídá zrcadlovému odrazu, a nekoherentní odrazivosti, která reprezentuje příspěvek od rozptýleného světla. Koeficient odrazu je vyjádřen stejným způsobem jako v (1), s tím rozdílem že veškerá drsnost je započtena pomocí SDT. To znamená, že lokální koeficient odrazu odpovídá nehomogenní vrstvě s hladkými rozhraními a při vyjádření hustoty pravděpodobnosti a charakteristické funkce musíme užít celkovou rms hodnotu výšek drsnosti $\bar{\sigma}_{\text{SDT}}$ místo hodnoty odpovídající části drsnosti s nízkými prostorovými frekvencemi. Výsledný koeficient odrazu je tedy roven

$$r_q = \bar{\chi}(v_z) r_q^0, \qquad \qquad \text{kde} \qquad \qquad \bar{\chi}(v_z) = \exp\left(-\bar{\sigma}_{\text{SDT}}^2 v_z^2/2\right),$$

kde r_q^0 je koeficient odrazu nehomogenní vrstvy v případě, kdy jsou rozhraní hladká. Koherentní složka odrazivosti je pak rovna

$$R_{\rm c} = |\bar{\chi}(v_z)|^2 R_0,$$

kde R_0 je odrazivost odpovídající případu s hladkými rozhraními.

Pro vyjádření nekoherentní odrazivosti je třeba znát hustotu pravděpodobnosti pro rozdělení výšek ve dvou odlišných bodech, která je zvolna ve tvaru

$$w_2(\eta_1, \eta_2; t) = w_1(\eta_1)C(t)\delta(\eta_1 - \eta_2) + w_1(\eta_1)w_1(\eta_2)[1 - C(t)],$$
(4)

kde $\delta(\eta_1 - \eta_2)$ je Diracova delta funkce, η_1 a η_2 jsou odchylky od středních rovin a t je vzdálenost těmito mezi body. Pro autokorelační koeficient C(t) je užit model s Gaussovou funkcí

$$C(t) = \exp\left(-t^2/\bar{\tau}_{\rm SDT}^2\right),\tag{5}$$

kde $\bar{\tau}_{\text{SDT}}$ je autokorelační vzdálenost. Pro spektrofotometr, který zachytí světlo rozptýlené do kužele s polovinou vrcholového úhlu rovnou $\bar{\alpha}_{\text{SDT}}$, lze nekoherentní složku odrazivosti vyjádřit jako [7]

$$R_i = R_0 \left[1 - \exp\left(-\frac{16\pi^2 n_0^2 \bar{\sigma}_{\text{SDT}}^2}{\lambda^2}\right) \right] \left[1 - \exp\left(-\frac{\pi^2 n_0^2 (\bar{\alpha}_{\text{SDT}} \bar{\tau}_{\text{SDT}})^2}{\lambda^2}\right) \right].$$
(6)

Experimentální data a jejich zpracování

Spektrální závislosti elipsometrických parametrů $I_{\rm s}$, $I_{\rm c}$ and $I_{\rm n}$, které odpovídají nezávislým prvkům normalizované Muellerovy matice izotropních systémů [8], byly měřeny pro 5 úhlů dopadu v rozsahu 55–75° a energie fotonů 0.6–6.5 eV (190–2066 nm) elipsometrem Horiba Jobin Yvon UVISEL. Odrazivost byla měřena při úhlu odpadu 6° pro energie fotonů 1.44–6.5 eV (190–860 nm) spektrofotometrem Perkin Elmer Lambda 1050.

Experimentální data měřená na zdrsněném povrchu a pro nehomogenní vrstvu s drsnými rozhraními byla zpracovávána současně. Při zpracování bylo předpokládáno, že rozhraní nehomogenní vrstvy vykazují stejnou drsnost jako zdrsněný povrch substrátu, a jsou tedy popsány společnými parametry pro drsnost. Zpracování experimentálních data bylo provedeno s užitím metody nejmenších čtverců.

Výsledky



Obrázek 1: Spektrální závislosti odrazivosti (nalevo) a elipsometrických parametrů při úhlu odpadu 70° (napravo) měřených na zdrsněném povrchu křemíku. Body reprezentují experimentální data, křivky odpovídají vypočteným hodnotám.



Obrázek 2: Spektrální závislosti odrazivosti (nalevo) a elipsometrických parametrů při úhlu odpadu 70° (napravo) měřených u vzorku s nehomogenní vrstvou a drsnými rozhraními. Body reprezentují experimentální data, křivky odpovídají vypočteným hodnotám.



Obrázek 3: Spektrální závislosti optických konstant (nalevo) a profil optických konstant pro energii fotonů $E = 3 \,\mathrm{eV}$ (napravo).



Obrázek 4: Spektrální závislosti indexu lomu (nalevo) a extinkčního koeficientu (napravo) přechodové vrstvy. Čerchované křivky odpovídají optickým konstantám krystalického křemíku.

parametr		hodnota
zdrsněný povrch křemíku		
tloušťka nativní oxidové vrstvy	$d_{\rm nol} \ [\rm nm]$	2.66 ± 0.01
nehomogenní vrstva s drsnými rozhraními		
tloušťka nehomogenní vrstvy (určeno z elipsometrie)	$d_{\rm e} \ [{\rm nm}]$	379.1 ± 0.3
tloušťka nehomogenní vrstvy (určeno z odrazivosti)	$d_{\rm r} [{\rm nm}]$	379.4 ± 0.3
tloušťka přechodové vrstvy	$d_{\rm tl} \; [{\rm nm}]$	18.8 ± 0.2
parametry drsnosti pro kombinaci RRT a SDT		
rms hodnota výšek drsnosti (RRT)	$\sigma_{\rm RRT}$ [nm]	10.34 ± 0.07
autokorelační vzdálenost (RRT)	$\tau_{\rm RRT}$ [nm]	92.1 ± 0.7
rms hodnota výšek drsnosti (SDT)	$\sigma_{\rm SDT}$ [nm]	12.5 ± 0.3
celková rms hodnota výšek drsnosti	$\sigma_{\rm total} \ [\rm nm]$	16.2
parametry drsnosti pro SDT se započtením rozptýleného světla		
rms hodnota výšek drsnosti (SDT)	$\bar{\sigma}_{\rm SDT}$ [nm]	15.8 ± 0.1
součin autokorelační vzdálenosti a poloviny přijímacího úhlu $\bar{\alpha}_{\rm SE}$	$_{\rm DT} \bar{\tau}_{\rm SDT} \ [\rm nm]$	18 ± 1

Tabulka 1: Hodnoty strukturních parametrů a parametrů popisujících náhodnou drsnost rozhraní.

		v			
			nehomog	enní vrstva	přechodová
parametr			horní rozhraní	spodní rozhraní	vrstva
celková síla mezipásových přechodů	$N_{\rm vc}$	$[eV^2]$	934 ± 14	142 ± 3	165 ± 3
energie šířky zakázaného pásu	$E_{\rm g}$	[eV]	1.89 ± 0.01	1.05 ± 0.02	1.70 ± 0.03
	$E_{\rm c}$	[eV]	13.04 ± 0.08	8.97 ± 0.07	3.68 ± 0.01
	B	[eV]	20.2 ± 0.3	0.01 (fixed)	1.01 ± 0.03

Tabulka 2: Hodnoty disperzních parametrů.

Výsledky – vrstva s hladkými rozhraními

Kromě nehomogenní vrstvy s drsnými rozhraními byla provedena i optická charakterizace nehomogenní vrstvy připravené identickým způsobem na substrát s hladkým povrchem. Byly užity stejné strukturní a disperzní modely jako u nehomogenní vrstvy s drsnými rozhraními, s tím rozdílem, že byla uvažována hladká rozhraní.



Obrázek 5: Spektrální závislosti odrazivosti (nalevo) a elipsometrických parametrů při úhlu odpadu 70° (napravo) nehomogenní vrstvy s hladkými rozhraními. Body reprezentují experimentální data, křivky odpovídají vypočteným hodnotám.

Tabulka 3: Hodnoty strukturních parametrů.	
parametr	hodnota
tloušťka nehomogenní vrstvy (určeno z elipsometrie) $d_{\rm e}$ [nn	n] 378.3 ± 0.4
tloušťka nehomogenní vrstvy (určeno z odrazivosti) $d_{\rm r}$ [n m	n] 378.9 ± 0.3
tloušťka přechodové vrstvy $d_{\rm tl}$ [nr	m] $ 16.2 \pm 0.3 $

			nehomoge	přechodová	
parametr			horní rozhraní	spodní rozhraní	vrstva
celková síla mezipásových přechodů	$N_{\rm vc}$ [e	eV^2]	1078 ± 16	201 ± 4	142 ± 3
energie šířky zakázaného pásu	$E_{\rm g}$ [[eV]	1.68 ± 0.01	1.29 ± 0.02	1.63 ± 0.03
	$E_{\rm c}$ [[eV]	14.02 ± 0.09	10.12 ± 0.07	3.39 ± 0.01
	B [[eV]	23.8 ± 0.4	1 (fixed)	0.96 ± 0.03

Tabulka 4: Hodnoty disperzních parametrů.



Obrázek 6: Spektrální závislosti optických konstant (nalevo) a profil optických konstant pro energii fotonů $E = 3 \,\mathrm{eV}$ (napravo). Čárkované křivky odpovídají závislostem určeným pro nehomogenní vrstvu s drsnými rozhraními.



Obrázek 7: Spektrální závislosti indexu lomu (nalevo) a extinkčního koeficientu (napravo) přechodové vrstvy. Čárkované křivky odpovídají závislostem určeným pro nehomogenní vrstvu s drsnými rozhraními. Čerchované křivky odpovídají optickým konstantám krystalického křemíku.

Závěr

Byla provedena úplná optická charakterizace nehomogenní vrstvy s náhodně drsnými rozhraními. Pro zahrnutí vlivu drsnosti na odrazivost v blízké infračervené a viditelné oblasti byla užita kombinace skalární difrakční teorie a Rayleigh–Riceovy teorie. Pro zahrnutí vlivu drsnosti na odrazivost v ultrafialové oblasti byla užita pouze skalární difrakční teorie, přičemž pro správnou interpretaci experimentálních dat bylo nutné zahrnou i to, že detektor spektrofotometru zachytí i část světla rozptýleného na vzorku. Mezi substrátem a nehomogenní vrstvou bylo nutné předpokládat tenkou přechodovou vrstvu, která zřejmě zahrnuje i tenkou vrstvu na povrchu substrátu porušenou při čištění substrátu v argonovém výboji provedeném před depozicí vrstvy. Výsledkem optické charakterizace jsou optické konstanty nehomogenní vrstvy a přechodové vrstvy, tloušťky nehomogenní vrstvy a přechodové vrstvy a parametry popisující náhodnou drsnost rozhraní.

Výsledky získané při optické charakterizaci nehomogenní vrstvy s náhodně drsnými rozhraními byly porovnány s výsledky získanými optickou charakterizací nehomogenní vrstvy s hladkými rozhraními připravenou za užití stejných technologických podmínek. Optické konstanty u horního a spodního rozhraní a profil indexu lomu určené pro nehomogenní vrstvu s drsnými rozhraními a pro nehomogenní vrstvu s hladkými rozhraními jsou téměř shodné, což potvrzuje správnost výsledků získaných při optické charakterizaci nehomogenní tenké vrstvy s drsnými rozhraními.

Reference

- D. Campi, C. Coriasso, Prediction of optical properties of amorphous tetrahedrally bonded materials, J. Appl. Phys 64 (1988) 4128–4134.
- [2] D. Franta, D. Nečas, I. Ohlídal, A. Giglia, Optical characterization of SiO₂ thin films using universal dispersion model over wide spectral range, Proc. SPIE 9890 (2016) 253–267.
- [3] D. Franta, A. Dubroka, C. Wang, A. Giglia, J. Vohánka, P. Franta, I. Ohlídal, Temperaturedependent dispersion model of float zone crystalline silicon, Applied Surface Science 421 (2017) 405–419.
- [4] D. Franta, I. Ohlídal, Ellipsometric parameters and reflectances of thin films with slightly rough boundaries, Journal of Modern Optics 45 (1998) 903–934.
- [5] J. Vohánka, M. Čermák, D. Franta, I. Ohlídal, Efficient method to calculate the optical quantities of multi-layer systems with randomly rough boundaries using the Rayleigh–Rice theory, Physica Scripta 94 (2019) 045502.
- [6] I. Ohlídal, J. Vohánka, J. Mistrík, M. Čermák, F. Vižd'a, D. Franta, Approximations of reflection and transmission coefficients of inhomogeneous thin films based on multiple-beam interference model, Thin Solid Films 692 (2019) 137189.
- [7] J. Porteus, Relation between the height distribution of a rough surface and the reflectance at normal incidence, J. Opt. Soc. Am. 53 (1963) 1394–1402.
- [8] I. Ohlídal, J. Vohánka, M. Čermák, D. Franta, Ellipsometry of layered systems, in: O. Stenzel, M. Ohlídal (Eds.), Optical Characterization of Thin Solid Films, Springer, International Publishing, Cham, 2018, pp. 233–267.



Table 3-14 USB Type-C to USB 3.1 Standard-B Cable Assembly Wiring

USB Type-C Plug			Wire	USB 3.1 Standard-B plug		
Pin	Signal Name	Wire Signal Number Name		Pin	Signal Name	
A1, B1, A12, B12	GND	1 7, 10	GND_PWRrt1 SDP1_Drain, SDP2_Drain	4 7	GND GND_DRAIN	
A4, B4, A9, B9	VBUS	2	PWR_VBUS1	1	VBUS	
A5	сс	See Note 1				
B5	VCONN					
A6	Dp1	3	UTP_Dp	3	D+	
A7	Dn1	4	UTP_Dn	2	D-	
A2	TXp1	5	SDPp1	9	StdB_SSRX+	
A3	TXn1	6	SDPn1	8	StdB_SSRX-	
B11	RXp1	8	SDPp2	6	StdB_SSTX+	
B10	RXn1	9	SDPn2	5	StdB_SSTX-	
Shell	Shield	Outer Shield	Shield	Shell	Shield	



S5980 TIA

CMI Petr Dvor	dK				
Sheet: /				CMI	NV
File: CMIAFMTIA0	2B.kicad_sc	h		CM	
Title: 4CH T	IA for S5	980			
Size: A4	Date: 20	22-07-05		Rev: A	
KiCad E.D.A. ki	cad (6.0.7-	1)-1		ld: 1/1	
4			5		

Table 3-14 USB Type-C to USB 3.1 Standard-B Cable Assembly Wiring

USB Type-C Plug			Wire	USB 3.1 Standard-B plug	
Pin	Signal Name	Wire Signal Number Name		Pin	Signal Name
A1, B1, A12, B12	GND	1 7, 10	GND_PWRrt1 SDP1_Drain, SDP2_Drain	4 7	GND GND_DRAIN
A4, B4, A9, B9	VBUS	2	PWR_VBUS1	1	VBUS
A5	CC	See Note 1			
B5	VCONN				
A6	Dp1	3	UTP_Dp	3	D+
A7	Dn1	4	UTP_Dn	2	D-
A2	TXp1	5	SDPp1	9	StdB_SSRX+
A3	TXn1	6	SDPn1	8	StdB_SSRX-
B11	RXp1	8	SDPp2	6	StdB_SSTX+
B10	RXn1	9	SDPn2	5	StdB_SSTX-
Shell	Shield	Outer Shield	Shield	Shell	Shield

J1 USB3_0483930003









C22

Ą

GND

10u_35\

C18

GND

-12\

C14

10u_35V





U1C TLE2074CDW

U1B

LE2074CDW

U1A TLE2074CDW

U1D

TLE2074CDW

IN1_BUF

IN2_BUF

IN3 BUF

IN4_BUF









ČMI | Petr Dvořák Sheet: / File: CMIAFMDIV03.kicad_sch Title: 4CH divider

GNE

U1

X_VP

X_Y1

GNE

_Y1

-0^{TP7}

X_Y1

È.

R23 ST7ETA104

R20

0603

10k



O TP16 GND

↓ GND

+12V

♦

-0 TP12 IN1

TP13^{GND} IN2

C2

0603

+

GND

0603

 \uparrow

GND

IN1

<u>IN2</u>

C1 22p_50_C0G

22p_50_C0G





• H1 M3.2 • H2 M3.2 H3 M3.2 H4 M3.2 $\stackrel{\frown}{\rightarrow}$ \uparrow $\stackrel{\frown}{\rightarrow}$ \uparrow GND GND GND GND



