

Výzkumná zpráva¹

Nephele: Databáze restaurátorských zpráv s možností vyhledávání podle textové a obrazové informace

Mgr. Miroslav Beneš, RNDr. Barbara Zitová, Ph.D.

ZOI, Ústav teorie informace a automatizace, Akademie věd České republiky

Pod vodárenskou věží 4, Praha 8

¹Tato práce byla podporována MŠMT, projektem č. 1M0572 (výzkumné centrum DAR).

Obsah

1	Úvod	3
2	Nephele: Databáze restaurátorských zpráv	6
2.1	Restaurátorské zprávy	6
2.2	Návrh databáze	8
2.3	Nephele – architektura a funkčnost	12
3	Image retrieval	16
3.1	Co je image retrieval?	16
3.2	Vyhledávání na základě podobnosti	16
3.2.1	Příznaky pro obrazová data	19
3.2.2	Redukce dimenze příznakového prostoru	25
3.2.3	Metriky a indexace	26
3.3	Přehled existujících systémů	27
4	Navrhované a implementované metody	29
4.1	Kookurenční matice	29
4.2	Haralickovy příznaky	30
4.3	Image retrieval v Nephele	33
4.3.1	Zvolené příznaky	33
4.3.2	Vyhledávání v Nephele	35
5	Výsledky na reálných datech	37
6	Závěry	40
A	Implementační detaily	41
A.1	Prostředí	41
A.2	Klient a server	41
A.2.1	Komunikační rozhraní	42
A.2.2	Klient	43
A.2.3	Server	43
A.3	Podpůrné knihovny	44
A.4	Uživatelské rozhraní	45
B	Rejstřík obrázků a tabulek	48

Kapitola 1

Úvod

Digitální technologie v dnešní době vstupují do oblastí, v kterých jsme si je ještě před pár lety nedokázali představit. S jejich použitím se můžeme setkat téměř na každém kroku. Televize, film, hudba a fotografie jsou jen zlomek všech odvětví, které v současnosti procházejí revolucí.

Nové přístupy se uplatňují i v oblastech shromažďování dat o uměleckých dílech. Klasické kartotéky byly vystřídány rozsáhlými databázemi. Ty mohou obsahovat nejen informace o dílech, ale i díla samotná (např. fotografie obrazů a soch, knihy). Shromážděná data mohou sloužit pro další výzkum nebo pro prezentaci těchto děl. Konstrukce 3D modelů soch či obnova poškozených obrazů představují jen část směru, kterým vývoj ubíhá. Nezanedbatelným přínosem pro galerie je pak možnost otevření svých fondů širšímu okruhu zájemců prostřednictvím Internetu. Případný návštěvník má možnost prohlédnout si kolekci a efektivně tak sestavit svůj program. Příkladem mohou být Národní galerie ve Washingtonu a v Londýně ([2, 5]). Podobný projekt na svých webových stránkách připravuje též Národní galerie v Praze ([3]).

Pro databázi uměleckých děl je klíčové vyhledávání, neboť umožňuje jednoduchý a efektivní přístup k datům a další práci s nimi. Běžné bývá vyhledávání textové, kdy se např. obrazy a sochy hledají podle názvu nebo jména autora. Všechny informace se však nedají popsat slovem. Snímky mohou mít větší výpovědní hodnotu než rozsáhlý text. Proto se v současné době začíná uplatňovat možnost vyhledávání podle obrazové informace. Tento způsob se dá ilustrovat na příkladu hledání obrázků na Internetu. Uživatel systému zadá obrázek. Výsledkem hledání jsou pak obrázky, které jsou vizuálně podobné původnímu dotazu.

Databáze uměleckých děl a obrazové vyhledávání ale nemusí být využity jen k prezentačním účelům. Mezi informace, které se o uměleckých dílech uchovávají, patří i zprávy o materiálovém průzkumu (v textu dále také jako restaurátorské zprávy). Zachycují proces restaurování díla v podobě dokumentů, fotografií, mikroskopických řezů a chemických rozborů použitých materiálů. Každá zpráva obsahuje snímky mikroskopických řezů vzorků, které byly z díla odebrány (použitím skalpelu nebo laseru). Jeden ze snímků vyjadřuje optické vlastnosti v bílém světle, další fluorescenci v ultrafialovém záření. Na snímcích 1.1 jsou jasně zřetelné jednotlivé vrstvy materiálů, z kterých se dílo skládá. Tyto snímky mohou být podkladem pro obrazové vyhledávání. Výsledkem jsou pak fotografie řezů

s podobnou strukturou. To může usnadnit pozdější klasifikaci právě zkoumaného díla na základě použitých pigmentů a velikosti a tvaru zrn ve vzorku (heterogenity).

Mikroskopickými řezy a jejich snímky se zabývá i diplomová práce [36]. Jejím cílem je odstranění posunutí kamery při snímání vzorků (metodou registrace). Druhým úkolem je získání základních informací o vrstvách vzorku metodou segmentace.

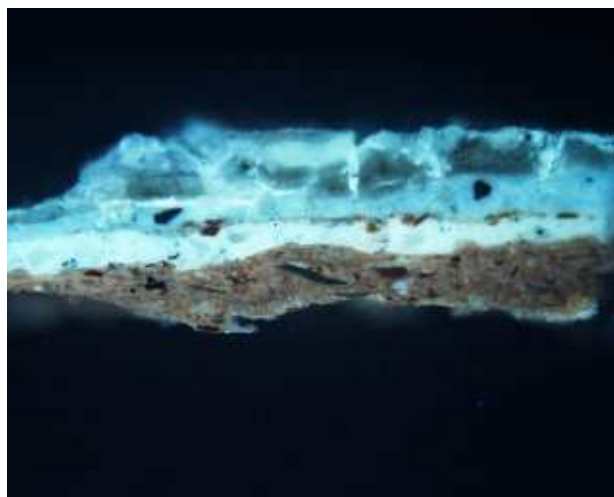
Cílem této práce je návrh databáze zpráv o průzkumu, která by umožňovala kromě klasického textového vyhledávání i hledání podle obrazové informace. Práce je prováděna ve spolupráci s Akademickou laboratoří materiálového průzkumu malířských děl ALMA [1]. Taková databáze by měla představovat znalostní bázi, která by měla jednak informativní charakter a také by umožnila efektivně využívat nabyté zkušenosti při restaurování dalších uměleckých děl. Jelikož jsou archivovaná data potenciálně velmi citlivá a mohou být zneužita, musí návrh systému počítat se zabezpečením přístupu a přístupovými právy.

Stěžejním úkolem bude seznámení se s existujícími postupy obrazového vyhledávání prostřednictvím dostupné literatury. Na základě získaných znalostí budou navrženy metody vyhledávání vhodné pro charakter snímků v restaurátorských zprávách. Ty poté budou také implementovány a jejich úspěšnost testována na dostupných reálných datech.

V kapitole 2 je uveden popis navrženého a implementovaného systému a databáze. Kapitola 3 se věnuje úvodu doobrazového vyhledávání a přehledu základních technik. Navrhované a implementované metody jsou podrobně popsány v kapitole 4. Kapitola 5 obsahuje praktické výsledky metod na reálných datech a kapitola 6 závěry.



(a)



(b)

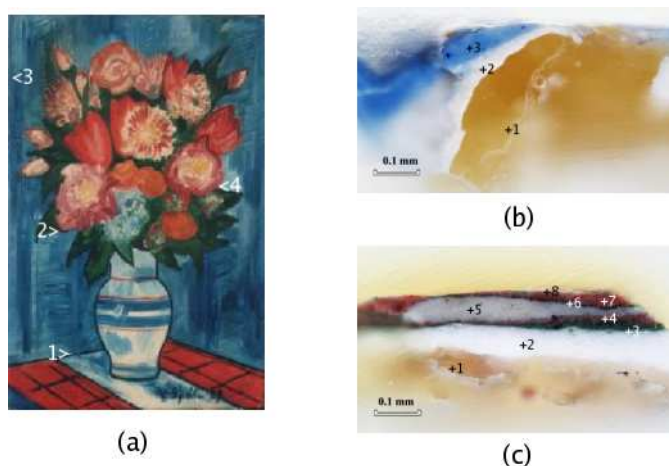
Obrázek 1.1: Snímky mikroskopických řezů v (a) bílém a (b) ultrafialovém světle

Kapitola 2

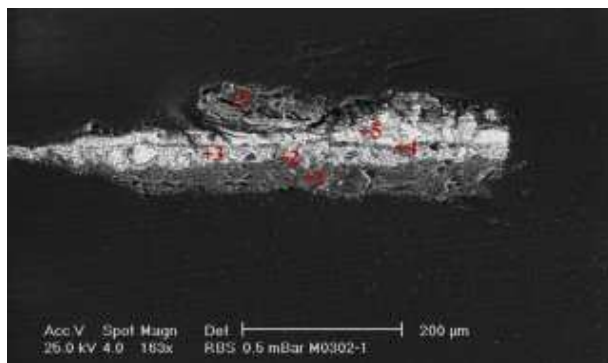
Nephele: Databáze restaurátorských zpráv

2.1 Restaurátorské zprávy

Restaurátorská zpráva zachycuje celý proces restaurace uměleckého díla včetně přípravné fáze (v našem případě nejčastěji díla malířského). Hlavním cílem procesu je především obnova díla, která by ale nebyla možná bez podrobného prostudování jeho materiální podstaty a výstavby. Do přípravné fáze se vedle restaurátorů zapojují i přírodovědci a historici umění. Jejich znalosti umožňují hlouběji pochopit dílo a zařadit ho do širšího historického kontextu. K tomu jsou nezbytné chemické analýzy odebraných vzorků (obrázek 2.1). Zároveň se pořizují snímky příčných řezů v bílém a ultrafialovém světle. Snímky poskytují informace o optických vlastnostech použitých materiálů, resp. fluorescenci látek, které jsou ve vzorku obsaženy. V následujících odstavcích se nejprve seznámíme s procesem získání vzorků a s obsahem zprávy.



Obrázek 2.1: (a) umělecké dílo s vyznačenými místy odběru vzorků; (b) snímek řezu vzorkem 1; (c) snímek řezu vzorkem 4



Obrázek 2.2: Snímek z elektronického mikroskopu (SEM snímek)

Proces získávání vzorků

Na počátku restaurace se v uměleckém díle vytipují místa pro odběr vzorku (obrázek 2.1). Vzorků se odebírá více a cílem je co nejlepší rozložení z hlediska zachycení všech vrstev v díle a možnosti jejich vzájemného porovnávání. Velikost vzorků se pohybuje v řádu desetin milimetru. K jejich odběru se používá skalpel nebo laser. Odebrané vzorky se zalévají do polyesterové pryskyřice ve tvaru hranolu a takto vzniklé útvary se leští až k hranici vzorku ve vhodném úhlu [25]. Vzorek se poté fotí v bílém a UV světle (snímky 1.1 (a) a (b)). Na snímcích 1.1 a 2.1 (b), (c) jsou jasně patrné vrstvy, z kterých se zkoumaný obraz skládá. Barevnou vrstvu tvoří pigmenty a jejich pojiva. Vedle optických vlastností látek se zkoumá i změna barvy v důsledku chemické reakce způsobené přidáním kapičky chemického činidla (*mikrochemická* či *histochemická zkouška*). Vzorek je dále analyzován v elektronickém mikroskopu (*spektrochemická analýza*) (obrázek 2.2). Složením pigmentů ve vrstvách, vztahy mezi vrstvami a odlišením původní, autorské malby od pozdějších přemalby se zabývá *stratigrafie*. Vychází se při tom jak z výsledků jednotlivých analýz, tak z techniky malby a stylu autora.

Vedle zkoumání pořízených snímků se provádí i chemické analýzy vzorků. Zjišťuje se přesné materiálové složení jednotlivých vrstev (identifikace anorganických pigmentů, mineralogická analýza, analýza pojiv). Důležitou informací o původu materiálu může nést i tvar a velikost zrn, které jsou viditelné na snímcích vzorků.

Podrobnější popis procesu získávání vzorků spolu s postupem při restaurování konkrétního malířského díla lze nalézt v [21]. Detaily o analýzách a další informace jsou k dispozici na webových stránkách Akademické laboratoře materiálového průzkumu malířských děl [1].

Obsah restaurátorské zprávy

Z poznatků, které jsou pro každé umělecké dílo dostupné (po provedení veškerých analýz), je vypracována zpráva o materiálovém průzkumu. Ta obsahuje jak obecné informace o díle (např. název, autor, použitá technika), tak informace o každém odebraném vzorku. Uvádí se lokalizace vzorku v rámci celkové plochy díla, seznam vrstev a materiálové přiřazení. Přílohy jsou samozřejmě i pořízené snímky. Zpráva dále obsahuje data z provedených chemických analýz a další materiály určené k podrobnějšímu studiu díla.

2.2 Návrh databáze

Cílem práce je návrh databáze restaurátorských zpráv a systému, který by umožnil textové a obrazové vyhledávání v databázi. Ke kvalitnímu návrhu takové databáze je nutná analýza dostupných dat a jejich dalšího použití (zda jsou určena jen k archivaci, nebo nad nimi bude implementováno vyhledávání). V této fázi je též důležitá úzká spolupráce s odborníky a autory zpráv, aby se předešlo zásadním změnám struktury databáze v budoucnosti (např. z důvodu chybějících či neúplných informací).

Analýza dat

Restaurátorskou zprávu (a tedy i data, která máme k dispozici) lze rozdělit do několika odlišných částí:

- **hlavička** – obsahuje obecné informace o díle. Jedná se především o jméno restaurátora, autora zprávy, název díla, dataci, techniku malby apod. Dále je přiložena fotodokumentace díla (celkový náhled na dílo) s vyznačenými místy odběru vzorků. Hlavička také obsahuje informace týkající se vlastních metod odběru a rozboru vzorků (odběr a příprava vzorků, optická mikroskopie, elektronová mikroskopie a mikroanalýza, doplňkové metody mikrochemické a histochemické, prášková RTG – mikrodifrakce) a závěry materiálového průzkumu.
- **rozbor vzorků** – u každého vzorku se nachází místo jeho odběru, snímky v bílém a ultrafialovém světle (případně další snímky, např. SEM snímek). K dispozici je též popis a interpretace jednotlivých barevných vrstev včetně přiřazení materiálů.
- **data chemických analýz** – v současnosti jsou přiložena data těchto analýz (více viz [1]):
 - **Edax** – detailní analýza prvkového složení a heterogenit. Je součástí měření v elektronovém mikroskopu a výsledkem jsou hmotnostní a atomová procenta prvků, které jsou zastoupeny ve vrstvě. U některých látek se nezjišťuje kvantitativní zastoupení, ale pouze kvalitativní.
 - **Difrakce** – přesněji prášková RTG-difrakce a mikrodifrakce. Jedná se o nástroj pro identifikaci minerálních složek barevné vrstvy. Výsledkem je materiálové přiřazení k vrstvě. Každá látka je identifikovatelná svým názvem, katalogovým číslem a chemickým vzorcem.
 - **FTIR** – Fourier Transform Infrared Spectroscopy neboli infračervená spektroskopie. Je součástí analýzy pojiv. Rámcové určení pojiv je nezbytné pro klasifikaci malířské techniky (tempera, olejomalba apod.), přesné určení pak pomáhá při autentifikaci díla. Výsledkem je sada hodnot (peaků), které jsou pro každé pojivo specifické.
 - **Maldi** – **MALDI-TOF MS** – Matrix Assisted Laser Desorption Ionisation Time of Flight Mass Spectroscopy. Rovněž patří k analýze pojiv. Slouží především k odlišení a určení proteinových pojiv (např. vajíčka).
- **další přílohy** – další snímky, textové přílohy

Z hlediska vyhledávání jsou na různé části zprávy kladeny jiné nároky. Podívejme se tedy na jednotlivé části z pohledu vyhledávání:

- **textové údaje** – kromě dat z hlavičky sem můžeme zahrnout i texty, které se týkají rozboru vzorků. Vyhledávání nad těmito informacemi se skládá z jednoduchých dotazů typu: vyhledej zprávy daného autora; vyhledej zprávy, v jejichž názvu je dané slovo apod. Vhodné je i implementování fulltextových dotazů.
- **snímky v bílém a UV světle** – tyto snímky budou základem pro implementaci obrazového vyhledávání, kterému se budeme věnovat v následujících kapitolách. Uživatel zadá systému dotaz, který představuje dvojice snímků v bílém a UV světle (které spolu korespondují). Cílem je pak vyhledání nejpodobnějších snímků. Tomuto přístupu se říká *query-by-example*.
- **data** – přínos vyhledávání nad výsledky chemických analýz spočívá v odhalení děl, které mají podobné materiálové složení jako právě zkoumaný objekt. Druhou možností je cílené vyhledání děl s předem daným materiálovým složením. Příklady dotazů: difrakce – vyhledej díla, které obsahují zadanou látku; FTIR – vyhledej díla, u kterých byla ve vrstvách naměřena daná hodnota (s předem danou tolerancí).

Návrh databáze by měl tyto nároky na vyhledávání respektovat.

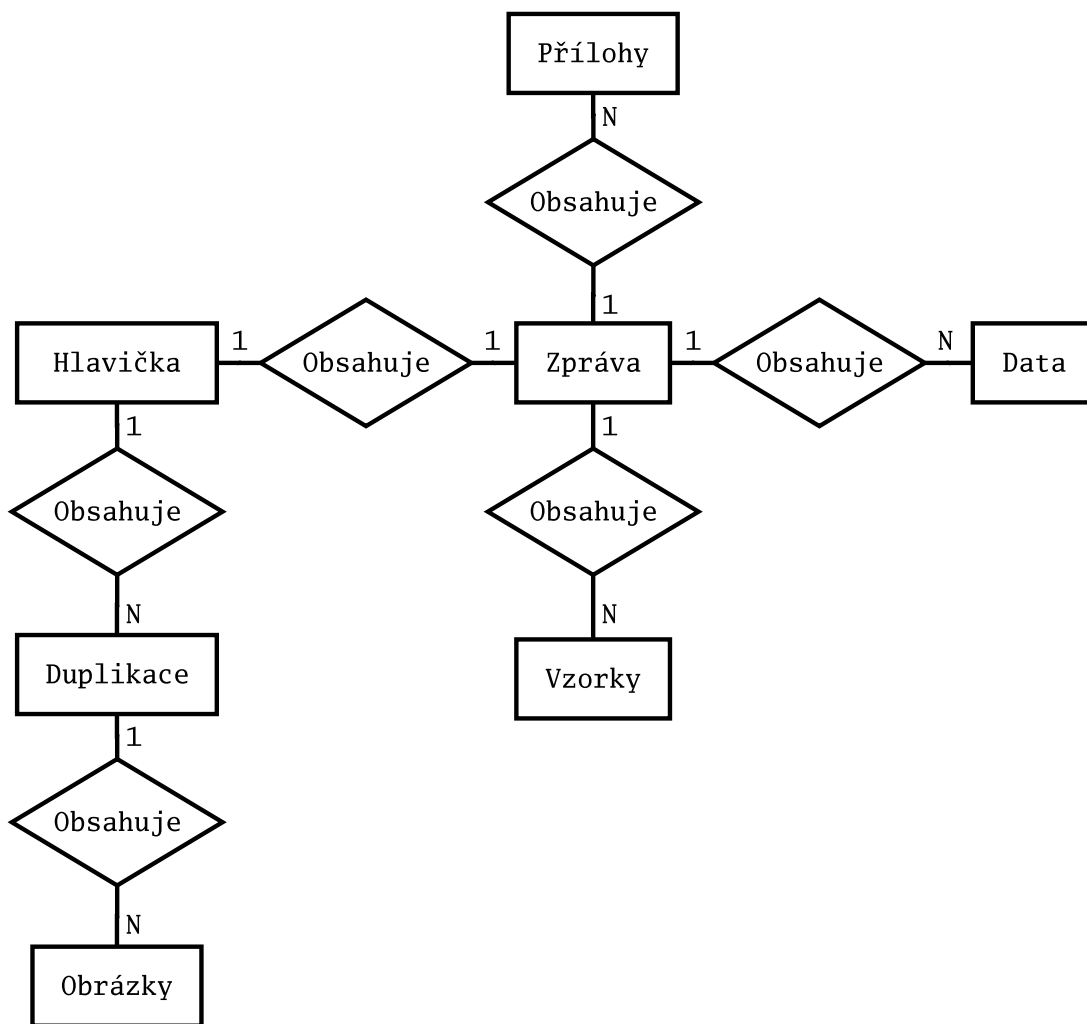
Výsledek analýzy dat – návrh databáze

Analýza dat z předešlé kapitoly byla dále podrobněji rozpracována. V této fázi se již přihlíželo k použití relační databáze s implementovaným dotazovacím jazykem SQL. Výhodou relační databáze je přehledné rozdělení dat do jednotlivých tabulek, jasně definované vztahy mezi daty v tabulkách a rychlost. Jazyk SQL pak umožňuje klást databázi i velmi sofistikované dotazy. Tabulky určené pro uložení restaurátorských zpráv, ke kterým analýza dat vedla, jsou uvedeny v tabulce 2.2. Zjednodušené ER-schéma se nachází na obrázku 2.3.

Nad důležitými sloupci v tabulkách jsou definovány indexy, které podstatným způsobem urychlují výběr záznamů z tabulek (SQL SELECT dotazy). Pro indexy jsou použity B-stromy. B-strom je vyhledávací výškově vyvážený strom, který splňuje jistá pravidla. Podrobnější informace lze nalézt v [37]. Pro fulltextové vyhledávání jsou použity *full-text* indexy. Fulltext index je konstrukce databáze MySQL pro umožnění fulltextového vyhledávání v definovaných sloupcích tabulky (další podrobnosti v manuálu MySQL – [4]).

Název tabulky	Popis obsahu tabulky
headers	Obecné informace z hlavičky, které jsou unikátní. Např. jméno autora zprávy, spolupracovníci, grant, poznámky apod.
headers_dupl	Informace z hlavičky, které mohou být duplikovány. Tzn. mohou se vyskytovat ve více instancích. Restaurované umělecké dílo se může skládat z více částí a údaje se mohou mezi částmi lišit. Patří sem fotodokumentace, název díla, autor díla apod.
headers_dupl_img	Celkové náhledy na dílo.
samples	Údaje o vzorcích, tzn. snímky, interpretace vrstev apod.
data_edax	Výsledky Edax
data_difrakce	Výsledky difrakce
data_difrakce_subst	Látky, které se mohou ve výsledcích difrakce objevit
data_ftir	Výsledky FTIR
data_ftir_subst	Látky, které se mohou ve výsledcích FTIR objevit
data_maldi	Výsledky Maldí
data_maldi_subst	Látky, které se mohou ve výsledcích Maldí objevit
addends	Přílohy restaurátorské zprávy

Tabulka 2.2: Tabulky navržené pro ukládání restaurátorských zpráv



Obrázek 2.3: Zjednodušené ER-schéma pro restaurátorskou zprávu

2.3 Nephele – architektura a funkčnost

Nephele¹ je systém postavený nad navrženou databází. Jelikož systém poběží v reálném prostředí (restaurátorské oddělení Akademie výtvarného umění v Praze), jsou na něj kladeny následující nároky:

- **stabilita**
- **bezpečnost**
- **jednoduchost**
- **přístupnost**
- **rozšiřitelnost**

Aplikace by měla poskytovat jednoduchý přístup ke všem informacím. Na druhé straně, přístup k citlivým údajům musí být chráněn bezpečnostními prvky (přístupová práva). Systém musí být přístupný pro více uživatelů v jeden moment. Důležitá je i rozšiřitelnost. Restaurátorské zprávy se obohacují dalšími analýzami a informacemi a ty musí být možné do systému zahrnout.

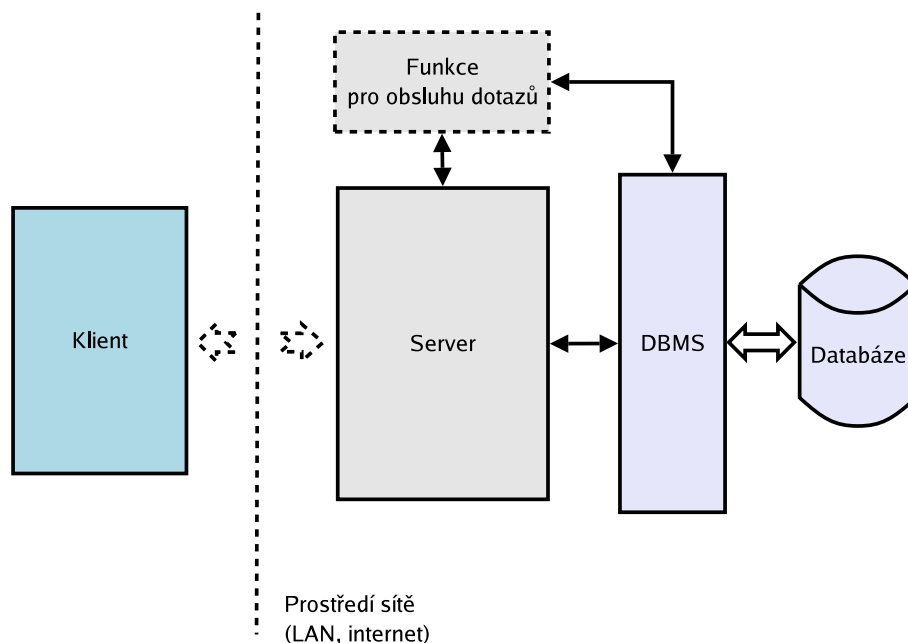
Cílem této práce není implementace všech funkcí, které by měl výsledný produkt obsahovat. Důraz je kladen zejména na obrazové vyhledávání a části, které umožní výsledky hledání zobrazit. Nephele však poskytuje dobrý základ, který se bude dále rozšiřovat o potřebné funkce.

Architektura systému

Systém Nephele je postaven na architektuře klient-server. Ideou této architektury je rozdělení systému na dvě v podstatě autonomní části. Část, které budeme říkat server, se stará o správu dat v databázi, vyhledávání (v našem případě i obrazové vyhledávání), správu uživatelů apod. Úkolem druhé části, kterou budeme nazývat klient, je komunikace s uživatelem. Poskytuje grafické rozhraní (GUI), které vhodnou formou interpretuje data obdržaná od serveru. Komunikace mezi serverem a klientem pak probíhá v relacích dotaz-odpověď na bázi přesně definovaného rozhraní (podrobnosti v příloze A.2.1). Server jakožto pasivní centrální uzel běží v jedné instanci, ke které se připojuje větší počet klientů. Omezení pro množství klientů, které se mohou k serveru připojit je dané konfigurací, která by měla respektovat náročnost serveru a výkonnost stroje, na kterém je zprovozněn (více viz klient-server architektura v [35]).

Na obrázku 2.4 vidíme architekturu Nephele. Klient je od serveru jasně oddělen prostředím sítě. Server obsluhuje dotazy a má přístup k databázi. V následujících odstavcích se podíváme na jednotlivé části podrobněji. Implementační detaily jsou uvedeny v příloze A.

¹*Nephele* - postava z řecké mytologie. Nephele byla nymfa, kterou stvořil Zeus z oblaků do podoby Hery.



Obrázek 2.4: Architektura Nephele

Klient zprostředkovává komunikaci mezi uživatelem a serverem (potažmo daty v databázi) pomocí grafického prostředí. Chování (resp. některé technické aspekty) jsou ovlivněny konfigurací (v konfiguraci se např. udává adresa a port naslouchajícího serveru). Uživatel se připojí k serveru a klade dotazy (vkládání nové zprávy, vyhledávání). Klient mu odpovědi serveru graficky zobrazí.

Server pasivně poslouchá do chvíle, než se k němu připojí klient. Provede přihlášení uživatele (ověření identity a hesla, zápis do databáze) a čeká na dotazy. Po příchodu dotazu server rozpozná typ a podle údajů v databázi zavolá funkci, která dotaz obslouží. Tato funkce již sama komunikuje s databází a provádí veškeré akce. Po provedení všech náležitostí odešle server odpověď a čeká na další dotazy. Ke své činnosti využívá server řadu podpůrných knihoven, kterým se věnuje samostatná kapitola (A.3). Chování serveru je v jistých směrech ovlivněno konfiguračním souborem.

Databáze zastřešuje veškerá data. Patří sem jak obsah restaurátorských zpráv, tak data, která jsou důležitá pro vlastní běh serveru (kapitola 2.3). Rovněž obsahuje data důležitá pro obrazové vyhledávání (více v kapitole 4.3.2). DBMS (Database Management System) je systém řízení databáze, který poskytuje přístup k databázi.

Přístupová práva

Jelikož budou mít do Nephele možnost přihlášení uživatelé s různými oprávněními (studenti, restaurátoři, administrátor), bylo nutné vymyslet takový systém přístupových práv, který je natolik pružný, aby v něm šlo zachytit různé úrovně přístupu k citlivým údajům. Výsledkem zneužití některých informací z databáze by mohlo být vytvoření velmi dobrých plagiátů. Databáze totiž obsahuje mj. podrobné údaje o materiálovém složení a struktuře, fotodokumentaci. Přístup k nim proto musí být neoprávněným osobám zamezen.

Každý uživatel patří do jisté skupiny práv, kterou mu přidělí administrátor při zakládání účtu. Uživatel má pak přidělena oprávnění této skupiny (ty pro skupinu definuje opět administrátor). Aby nemusela být pro každého uživatele, kterému je třeba mírně upravit přístupová práva, vytvořena zvláštní skupina, má administrátor možnost přidělit mu práva nad rámec jeho základní skupiny. Tato práva mohou být jak rozšiřující, tak restriktivní.

Oprávnění můžeme v zásadě rozdělit do dvou skupin. První skupinu tvoří příznaky u restaurátorských zpráv. Aby měl uživatel ke zprávě nesoucí takový příznak stejný přístup jako v případě, kdy zpráva příznak nenesou, musí mít přidělena odpovídající práva. Druhou skupinu tvoří běžná oprávnění, které se zpráv netýkají. Všechna v současnosti implementovaná uživatelská práva jsou popsána v tabulce 2.4.

Dále se předpokládá zabezpečení na úrovni operačního systému stroje, na kterém databáze a server Nephele poběží. Jedná se zejména o omezení přístupu jen pro stroje z důvěryhodné domény apod.

Podpora v databázi

Pro zajištění běhu serveru a jeho správné funkčnosti musí být v databázi vytvořeny podpůrné tabulky. Server potřebuje ukládat údaje o uživateli a jejich přístupových právech. Tzn. práva jednotlivých skupin, do kterých uživatelé patří, a pak oprávnění, která mohou mít uživatelé nad rámec své skupiny. U zpráv je také nutno vést příznaky, které jim byly nastaveny. Tabulka **res_requests_fnc** je klíčová pro určení funkce, která dotaz klienta obslouží (více v A).

Tabulky jsou uvedeny v přehledu 2.6.

Druh oprávnění	Seznam oprávnění s popisem
Příznaky u zpráv	<i>Zobrazení majitele a umístění Zobrazení inventárního čísla Tisk zprávy Zobrazení snímku v UV světle a SEM snímku Zobrazení interpretace vrstev Zobrazení závěrů v hlavičce Uzamčení zprávy pro další editaci</i>
Další uživatelská práva	<i>Fulltextové vyhledávání Editace záznamů Zobrazení snímků v plném rozlišení Administrátor Vkládání nových záznamů</i>

Tabulka 2.4: Přístupová oprávnění

Název tabulky	Popis obsahu tabulky
res_users	Záznamy o uživateli. Tzn. přihlašovací jméno, heslo, jméno a příjmení, skupina práv atd.
res_requests_fnc	Pro typ dotazu udává jméno funkce, která má dotaz obsloužit.
res_auth_grp_name	Skupiny přístupových práv
res_rights_name	Jména práv
auth_rep_req	Která zpráva má nastaveny jaké příznaky
auth_grp_rights	Práva skupiny
auth_user_rights	Oprávnění, která má uživatel nad rámec skupiny

Tabulka 2.6: Tabulky pro zajištění chodu serveru

Kapitola 3

Image retrieval

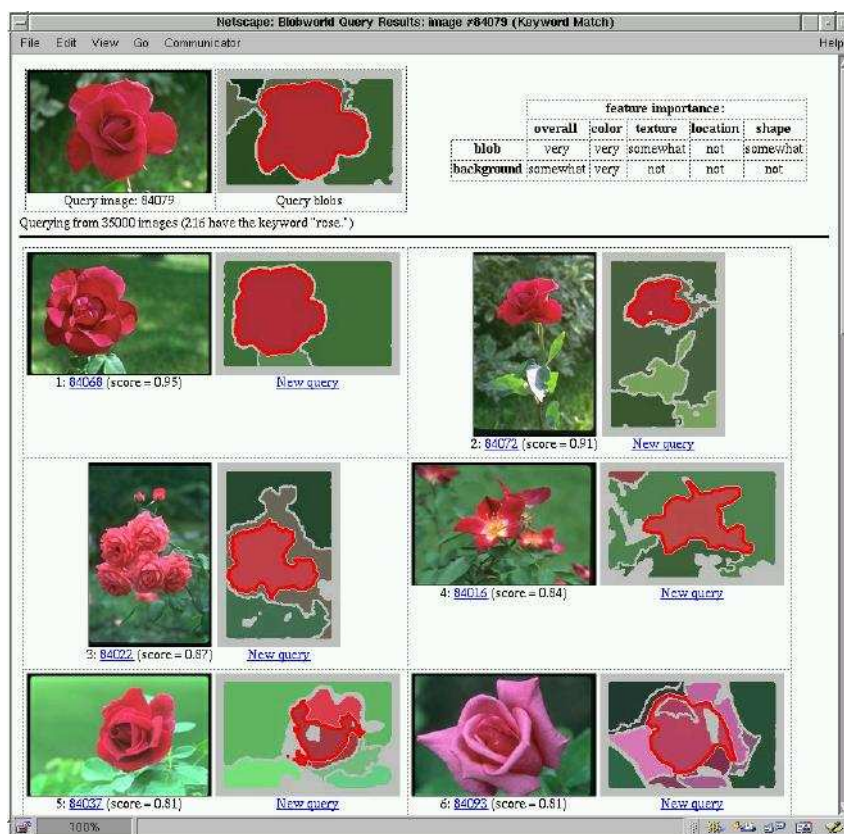
3.1 Co je image retrieval?

Obrazové vyhledávání (*image retrieval*) má kořeny v druhé polovině 70. let 20. století. V té době vznikla potřeba hledání snímků v rozsáhlých databázích. Velmi rozšířenou metodou bylo hledání na základě textové anotace (*text-based image retrieval*). Při vkládání snímku do databáze se připojil krátký popis. O image retrieval se pak v důsledku staral textový vyhledávací systém dané databáze. Nevýhody tohoto přístupu se začaly projevovat při nárůstu objemu databáze. První obtíží byla právě samotná ruční anotace, druhou pak subjektivita lidského vnímání. Dva různí lidé mohou totiž jeden snímek popsat různými způsoby (převzato z [38]).

V 90. letech se začal více rozvíjet jiný přístup: vyhledávání na základě podobnosti obrázků či snímků (*content-based image retrieval*, dále jen CBIR). Textovou anotaci střídá charakteristika založená na vizuálním obsahu snímků. Tou může být barva, textura, tvar objektů, prostorový vztah objektů či celková struktura snímku. V současnosti neexistuje obecná metoda, která by se hodila na vyhledávání obecných snímků. Každá aplikace obrazového vyhledávání si žádá výběr vhodných metod, které dávají nejlepší výsledky pro data v databázi. Do výzkumu těchto metod se zapojují poznatky z počítačového vidění a digitálního zpracování obrazu.

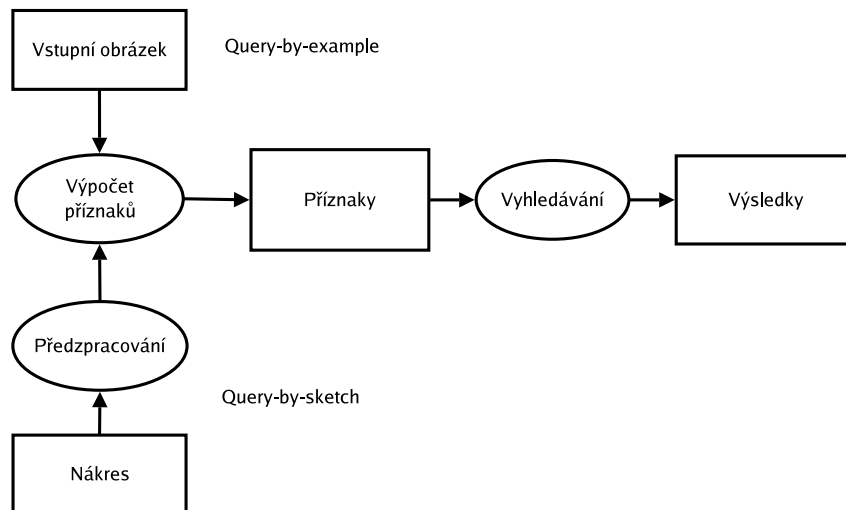
3.2 Vyhledávání na základě podobnosti

Vyhledávání podle vizuální podobnosti je založené na porovnávání příznakových vektorů jednotlivých objektů (snímků). Příznaky se myslí měřitelné veličiny, které lze z každého objektu spočítat. Vektor těchto příznaků pak charakterizuje a zároveň odlišuje snímky v databázi. Výsledkem vyhledávání jsou obrázky s nejpodobnějším vektorem příznaků. Podle typu dotazů můžeme CBIR rozdělit do dvou skupin:



Obrázek 3.1: Příklad vyhledávání podle obrazové podobnosti. Dotaz představuje obrázek růže. Systém vyhledá všechny odpovídající snímky v databázi. Jedná se o systém Blobworld (více v kapitole 3.3 a v [10]).

- *Query-by-sketch* – uživatel poskytne aplikaci hrubý nákres objektu, který má být vyhledán. Systém pak provede obrazové úpravy nákresu (vyhlazení, odstranění přebytečných linií apod.). Z upraveného obrázku se spočítají příznaky.
- *Query-by-example* – uživatel zadá předlohu, která se podobá obrázkům, jež chce uživatel vyhledat. Systém na předloze spočítá příznaky a provede hledání.



Obrázek 3.2: Schéma obrazového vyhledávání (query-by-example, query-by-sketch)

Dále se budeme zabývat pouze vyhledáváním založeném na query-by-example, které je v našem případě perspektivnější. První skupinou se zabývají např. publikace [23, 9]. Formálněji můžeme úlohu query-by-example definovat následovně:

Nechť \mathcal{I} představuje sadu obrázků v databázi ($\mathcal{I} = \{I_i \mid i = 1, \dots, n\}$). Nechť f je funkce, která pro obrázek X vrátí příznakový vektor. Nechť Q je snímek, který zadá uživatel jako dotaz. Nechť $E(f(Q), f(X))$ je metrika. Cílem úlohy je pak vyhledat k nejpodobnějších obrázků z \mathcal{I} . Tzn. že se nejprve setřídí I_i vzestupně podle $E(f(Q), f(I_i))$. Výsledkem vyhledávání je prvních k obrázků ze setříděného seznamu. Metrika $E(f(Q), f(I_i))$ představuje míru podobnosti obrázků Q a I_i .

Pro zajištění co nejlepších výsledků jsou na příznaky (a příznakový vektor) kladeny následující požadavky ([17]):

- **efektivita** – dobrá detekovatelnost, efektivní a rychlý výpočet, malé paměťové nároky
- **robustnost** – vzhledem k aditivnímu šumu (případně k chybám předzpracování)
- **invariance** – podobné obrázky by měly mít podobné příznaky
- **diskriminabilita** – příznaky musí být schopné odlišit různé obrázky

Tyto požadavky jsou dosti obecné a často stojí proti sobě (např. robustnost a diskriminabilita). Pro účely obrazového vyhledávání se proto přistupuje ke kompromisům a požadavky se oslabují. Z požadavků můžeme odvodit nároky na f (viz [24]):

- f rychle spočítatelné
- $|f(I)|$ relativně malé
- $E(f(I), f(I'))$ velké právě tehdy, když I a I' jsou různé

Příznakům pro obrazová data se věnuje kapitola 3.2.1, různým typům E pak mj. kapitola 3.2.3.

3.2.1 Příznaky pro obrazová data

V této kapitole se budeme věnovat přehledu konkrétních příznaků pro obrazová data se zřetelem na použití v obrazovém vyhledávání. Přehled metod lze najít v publikacích [13, 38], ze kterých je také citováno.

Příznaky se mezi sebou liší svou podstatou a přístupem k obrazovým datům. Proto si je pro naše účely rozdělíme do několika skupin. Podrobněji se budeme zabývat příznaky, které se počítají z barevné charakteristiky snímků, textur a tvaru objektů.

Příznaky, k jejichž výpočtu je nějakým způsobem využita barva obrázku, patří k nejjednodušším a nejrozšířenějším. Jsou robustní vzhledem k rozměrům obrázku a rotaci. Asi nejvýznamnějším zástupcem této skupiny příznaků jsou barevné histogramy a příznaky z nich spočítané.

Druhou používanou skupinou jsou texturové příznaky. Textura¹ je pojem, který se běžně užívá v počítačové grafice, digitálním zpracování obrazu a rozpoznávání. Definovat texturu ale není jednoduché. *Encyklopedie Universum* mj. říká, že textura je předpis, který definuje vlastnosti určitého materiálu v počítačové grafice, popisující jeho chování při interakci s paprsky světla. Příklady takové textury jsou na obrázku 3.3.

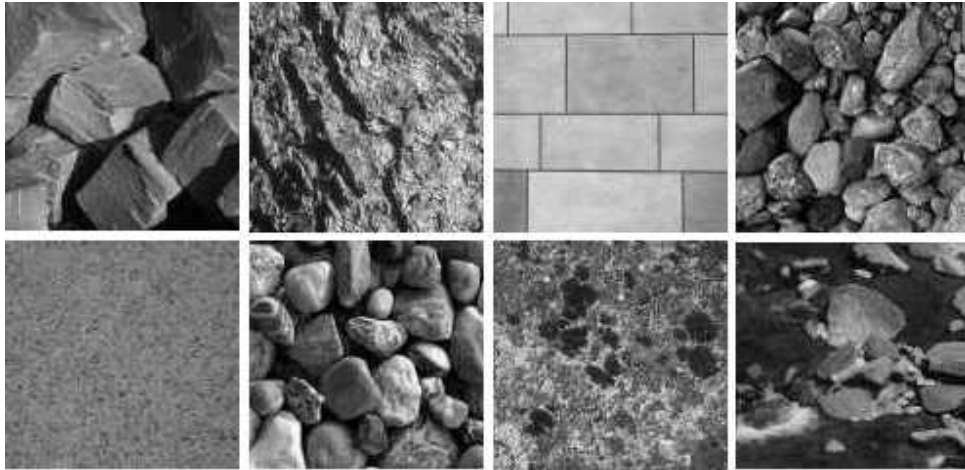
Pro naše účely můžeme texturu definovat jako část obrazu, která vykazuje následující vlastnosti:

- *hladkost (smoothness)*
- *zrnitost (coarseness)*
- *pravidelnost (periodicity)*

Texturové příznaky se pak používají v systémech pro zpracování lékařských snímků (rentgenové, CT apod.), dále v aplikacích, které jsou spojené s uměleckými díly (galerie, mikroskopické řezy), ale třeba i pro charakteristiku hornin. Mezi zástupce patří kookurenční matice, korelogramy, Tamurovy příznaky nebo např. waveletové příznaky.

Do množiny příznaků, jejichž základem je výpočet z tvaru či hranice objektů, můžeme zařadit např. jednoduché geometrické příznaky. Dále sem patří Fourierovy deskriptory, momentové invarianty a další.

¹tkaň; z latiny (*Ottův slovník naučný*)



Obrázek 3.3: Příklady textur

Další skupinu mohou tvořit příznaky, které vyjadřují prostorovou závislost objektů v obrázku. Těmi se dále nebudeme zabývat, neboť pro charakter našich dat nejsou vhodné. Podrobněji se jim např. věnují články [28, 11]. Autoři popisují vztah mezi objekty pomocí jednoduchých textových řetězců.

Barevné histogramy

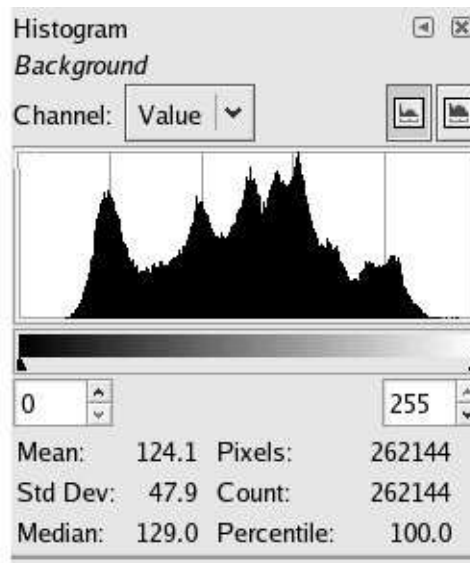
Histogram je statistická veličina², která ve spojitém případě vyjadřuje hustotu pravděpodobnosti. V oblasti počítačové grafiky a digitálního zpracování obrazu histogram představuje grafické znázornění absolutních četností jasových hodnot, které jsou v obrázku zastoupeny. Pro naše účely si histogram zadefinujeme následovně:

Nechť I je obrázek v barevném diskretním prostoru o N barvách (či jasových hodnotách). Pak histogram obrázku $H(I)$ můžeme vyjádřit jako vektor $(h_{c_1}, h_{c_2}, \dots, h_{c_N})$, kde h_{c_i} je četnost pixelů barvy c_i v obrázku I . Pokud má obrázek I velikost $m \times n$ pixelů, tak platí rovnost $\sum_{i=1}^N h_{c_i} = m * n$.

V případě barevného obrázku je pak histogram vícerozměrný (n -rozměrný, kde n udává počet barevných kanalů). Histogram se pro další použití často normalizuje celkovým počtem pixelů v obrázku ($m * n$).

Pro účely vyhledávání mohou být porovnávány samotné histogramy. Systém popsany v publikaci [39] používá histogramy a euklidovskou metriku pro předtřídění obrázků v databázi. Na vyhovujících snímcích se počítají další příznaky. Z histogramu lze však také snadno spočítat sadu hodnot, které mohou být použity přímo jako příznaky pro obrazové vyhledávání. Příkladem může být *průměrná hodnota jasu*, *kontrast* (čili rozptyl hodnot) apod. Příklad histogramu se nachází na obrázku 3.4.

²*Encyklopedie Universum* definuje histogram jako grafické znázornění četností hodnot statistického znaku ve tvaru sloupků.



Obrázek 3.4: Histogram obrázku v grafickém programu Gimp (www.gimp.org)

Kookurenční matice

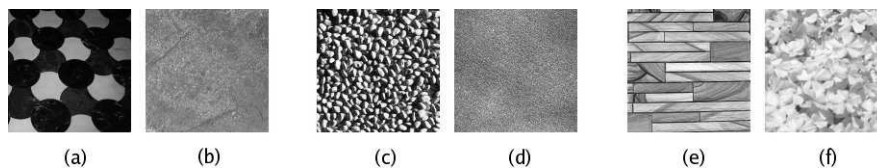
Využití kookurenčních matic v digitálním zpracování obrazu bylo poprvé publikováno v roce 1973 v článku [19] kolektivem kolem Roberta Haralicka. Matice jsou definované pro určitou orientaci a vzdálenost. $M(i, j)$ udává četnost výskytu pixelu s jasnem j ve vzdálenosti d a s orientací θ od pixelu s jasnem i (podrobné vysvětlení kookurenčních matic s příklady v kapitole 4.1). Pro lepší představu uvedeme ještě jinou definici v již diskretním prostoru: $M(i, j)$ udává četnost výskytu dvou pixelů s odstíny i a j . Dvojice pixelů je dána tzv. **vektorem posunu** (*displacement vector*) (dx, dy) . dx představuje posunutí v pixelech po x -ové souřadnici, dy obdobně pro y -ovou osu.

Matice tedy vyjadřují prostorovou závislost v textuře. Z kookurenčních matic lze dále spočítat různé příznaky. Haralick ve svém článku navrhl sadu asi 14 příznaků (Haralickovy příznaky), které charakterizují předloženou texturu. Patří sem například kontrast, entropie, korelace a další (solistikovanější) veličiny. Haralickovými příznaky se zabývá kapitola 4.2.

Kookurenční matice spolu s příznaky zůstávají jednou z nejpoužívanějších metod. Matice, které byly původně definovány pouze pro šedotónové obrázky, byly rozšířeny i pro barevné kanály (např. [26]). Variací barevných kookurenčních matic jsou modifikované barevné kookurenční matice (uvedeny v [40]). Narozdíl od barevného prostoru RGB v předešlých metodách vycházejí z prostoru HSV (Hue, Saturation, Value) a zejména složky Hue, která podle článku nese nejdůležitější informaci. Článek rovněž redukuje množinu přípustných orientací a vzdáleností pro samotnou kookurenční matici.

Barevné korelogramy

Na podobném principu jako kookurenční matice jsou založeny barevné korelogramy (představeny v publikaci [24]). Korelogramy zachycují lokální prostorovou korelaci barev v obrázku a jsou to svou podstatou barevné histogramy rozšířené o prostorovou informaci.



Obrázek 3.5: Texturey ilustrující tři z Tamurových příznaků (texturey pochází z databáze [6]): (a) vysoká hodnota *coarseness*, (b) nízká hodnota *coarseness*, (c) vysoký *kontrast*, (d) nízký *kontrast*, (e) vysoká hodnota *directionality*, (f) nízká hodnota *directionality*

Barevný korelogram můžeme definovat jako tabulku indexovanou párem barev (i, j) , kde k -tá položka na pozici (i, j) udává pravděpodobnost, že se pixel barvy j nachází ve vzdálenosti k od pixelu barvy i . Formálněji můžeme definici zapsat následovně:

Nechť I je obrázek o rozměrech $m \times n$, p je pixel v obrázku I ($p = (x, y)$) a I_i je množina pixelů barvy i . Nechť vzdálenost dvou pixelů $p_1 = (x_1, y_1)$ a $p_2 = (x_2, y_2)$ je definována jako $|p_1 - p_2| = \max\{|x_1 - x_2|, |y_1 - y_2|\}$. Pak barevný korelogram C pro vzdálenost k je definován vztahem:

$$C^{(k)}(i, j) = \Pr_{p_1 \in I_i, p_2 \in I} [p_2 \in I_j \mid |p_1 - p_2| = k]$$

Autokorelogram zachycuje prostorovou informaci pouze mezi identickými barvami:

$$AC^{(k)}(i) = C^{(k)}(i, i)$$

Článek [24] dále stanovuje vhodné metriky pro porovnávání dvou korelogramů, resp. autokorelogramů. Pozornost je též věnována efektivnímu výpočtu korelogramu, k němuž se používá techniky dynamického programování a rychlého násobení matic.

Tamurovy příznaky

V roce 1978 navrhl Tamura a kolektiv v článku [41] texturové příznaky, které vycházejí z principů lidského vizuálního vnímání. Jednalo se o těchto šest příznaků: *coarseness*, *contrast*, *directionality*, *line-likeness*, *regularity* a *roughness*. Experimentálně ověřili, že zejména první tři velmi dobře korelují se způsobem lidského vnímání. Na rozdíl od Haralickových příznaků pak tyto vyjadřují dobře představitelné vlastnosti textur. Zmíněným třem příznakům se budeme věnovat podrobněji (více informací dále v [20, 13]). Texturey, které ilustrují jednotlivé příznaky, se nacházejí na obrázku 3.5.

Coarseness Tento příznak nese informaci o velikosti elementů (zrn), které se v textuře pravidelně opakují. Pokud je v textuře přítomno více takových vzorů, bere se v úvahu ten větší z nich. Výpočet příznaku probíhá v několika krocích:

1. Pro každý bod (x, y) v obrázku I se spočítá průměrná hodnota v jeho okolí. V úvahu se berou okolí velikosti mocniny dvou, tzn. $2^k \times 2^k$ pro $k = 0, 1, \dots$

$$A_k(x, y) = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} I(i, j)/2^{2k}$$

2. Pro každý bod (x, y) se spočítá rozdíl průměrných hodnot nepřekrývajících se okolí bodu (x, y) , které se nacházejí na opačných stranách vůči zvolenému bodu a to v horizontálním a vertikálním směru

$$E_k^h(x, y) = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)|$$

$$E_k^v(x, y) = |A_k(x, y + 2^{k-1}) - A_k(x, y - 2^{k-1})|$$

3. V každém bodě (x, y) se zvolí okolí s maximální hodnotou rozdílu (najdeme optimální k)

$$S(x, y) = \operatorname{argmax}_{k=1 \dots \delta} \max_{d=v, h} E_k^d(x, y)$$

4. Hodnota coarseness je pak průměr přes 2^S

$$F_{crs} = \frac{1}{m * n} \sum_{i=1}^m \sum_{j=1}^n 2^{S(i, j)}$$

Contrast Vyjadřuje dynamický rozsah jasových hodnot v textuře a polarizaci distribuce bílé a černé barvy v histogramu.

$$F_{con} = \frac{\sigma}{\alpha_4^z}, \quad \alpha_4 = \frac{\mu_4}{\sigma^4}$$

μ_4 je čtvrtý moment střední hodnoty jasu a σ^2 je rozptyl jasových hodnot. Parametr z byl experimentálně stanoven na hodnotu $\frac{1}{4}$.

Directionality Zachycuje míru existence orientace v textuře. Na texturu jsou nejprve v každém bodě aplikovány jednoduché konvoluční masky pro zjištění derivací v horizontálním a vertikálním směru (Δ_H a Δ_V). V každém bodě se následně spočte hodnota $\theta = \frac{\pi}{2} + \tan^{-1} \frac{\Delta_V(x, y)}{\Delta_H(x, y)}$. Z hodnot θ je sestaven histogram. Příznakem je pak suma druhých momentů z hodnot kolem vrcholů histogramu.

Gaborovy příznaky

Rozvoj teorie wavelet na počátku 90. let 20. století přinesl jejich využití i do obrazového vyhledávání. Známé jsou zejména příznaky založené na Gaborových filtrech a Gaborově waveletové transformaci. Gaborovy funkce pro analýzu textur poprvé použil Turner ve své publikaci [42]. Praktickým použitím se též zabývá práce [30], ze které je dále citováno.

Dvourozměrnou Gaborovu funkci definujeme následovně

$$g(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y} \right) \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi j W x \right]$$

W je modulační frekvence. Aplikací dilatace a rotace můžeme z Gaborovy funkce získat sadu Gaborových filtrů

$$g_{mn}(x, y) = a^{-m} G(u, v)$$

$$a > 1, m, n \in \mathbb{N}, u = a^{-m}(x\cos\theta + y\sin\theta), v = a^{-m}(-x\sin\theta + y\cos\theta), \theta = \frac{n\pi}{K}$$

$G(u, v)$ je Fourierova transformace Gaborovy funkce, K je celkový počet rotací.

Nechť $I(x, y)$ je obrázek. Gaborova waveletová transformace je pak definována jako

$$W_{mn}(x, y) = \int I(x', y') g_{mn}^*(x - x', y - y') dx' dy'$$

g_{mn}^* je funkce komplexně sdružená ke g_{mn} . W v této definici není totožné s W v definici Gaborovy funkce.

Gaborovy příznaky tvoří sada středních hodnot a směrodatných odchylek $|W_{mn}|$.

Geometrické příznaky

Do skupiny geometrických příznaků můžeme zařadit jednoduché veličiny, které je možné spočítat z tvaru objektu či jeho hranice. Patří sem například následující příznaky:

- *konvexnost* – poměr plochy objektu k jeho konvexnímu obalu
- *kompaktnost* – míra podobnosti objektu ke kruhu
- *podobnost obdélníku* – poměr plochy objektu a plochy minimálního opsaného obdélníka
- *podlounlost* – poměr stran minimálního opsaného obdélníka
- *Eulerovo číslo* – rozdíl počtu komponent a počtu děr

Slabinou těchto příznaků je jejich nízká diskriminabilita. Lepší vlastnosti mají jiné dva příznaky:

- *shape vector* – vektor délek paprsků, které jsou vyslány z těžiště k hranici objektu. Tento příznak je invariantní vůči rotaci. Pokud vektor znormujeme délkou nejdelšího paprsku, získáme příznak, který je invariantní i vůči měřítku. Nevýhodou tohoto příznaku je jeho určení pouze pro hvězdicové objekty.

- *shape matrix* – v nejvzdálenějším bodě od těžiště opíšeme kolem objektu kružnici. Mezi těžištěm a touto kružnicí pak zkonstruujeme soustředné kružnice ve stejné vzdálenosti. Následným vytyčením paprsků od těžiště k opsané kružnici získáme síť se segmenty. Ze sítě vznikne matice již jednoduše. Na souřadnicích je 1, pokud je odpovídající segment alespoň z poloviny překryt objektem, v opačném případě 0.

Fourierovy deskriptory

Fourierovy deskriptory patří k příznakům, které se počítají přímo z hranice objektu. Z hranice je sestavena radiální funkce (např. vzdálenost hranice od těžiště). Tato funkce je 2π -periodická, a proto na ni můžeme provést Fourierovu transformaci. Absolutní hodnoty Fourierových koeficientů tvoří jednoduché příznaky. Tyto příznaky jsou invariantní vůči rotaci. To je důsledkem Fourier shift theoremu. Pokud se objekt otočí, posune se radiální funkce o konstantu. Z theoremu pak plyne, že Fourierova transformace se změní pouze ve fázi a ne v amplitudě. Dalšími úpravami lze dosáhnout i invariance vůči změně měřítka či posunu. Více informací o Fourierových deskriptorech je možné nalézt např. v článku [45].

Momentové invarianty

Momentové invarianty tvoří další skupinu příznaků, které můžeme zařadit do množiny veličin, které se počítají z tvaru objektu.

- *základní (geometrický) moment*: $m_{pq} = \int \int x^p y^q f(x, y) dx dy$, kde $p, q \geq 0$
- *centrální moment*: $\mu_{pq} = \int \int (x - x_t)^p (y - y_t)^q f(x, y) dx dy$, kde x_t a y_t jsou souřadnice těžiště; $p, q \geq 0$
- $p + q$ nazýváme řádem momentu

Základní moment nultého řádu (m_{00}) pak představuje nenormovanou střední hodnotu. Centrální momenty druhého řádu (μ_{20} a μ_{02}) vyjadřují rozptyl v dané ose (směru). Centrální momenty jsou invariantní vůči posunu. Invarianci ke změně měřítka lze dosáhnout např. vydělením mocninou momentu nultého řádu (podobný vztah platí i pro základní momenty):

$$\nu_{pq} = \frac{\mu_{pq}}{\mu_{00}^\omega}, \quad \omega = \frac{p+q}{2} + 1$$

Dalšími výpočty lze získat invarianty k dalším geometrickým transformacím. V roce 1962 představil Hu sedm rotačních invariantů ([22]). Postupně byly objeveny invarianty k afinní transformaci atd. Afinním momentovým invariantům je věnován článek [16].

3.2.2 Redukce dimenze příznakového prostoru

Motivace redukce dimenze příznakového vektoru je zřejmá. Příliš velký vektor (v řádu mnoha desítek až stovek příznaků) může zpomalovat a zhoršovat vyhledávání. Proto

vznikly metody na snížení velikosti příznakového vektoru resp. dimenze $f(I)$. Existují dva přístupy: *feature extraction* a *feature selection*.

Feature extraction Principem je transformace (projekce) existujícího příznakového prostoru do prostoru s nižší dimenzí. Tím vzniknou nové, umělé příznaky, které nemají význam původních příznaků. Zástupcem této skupiny je Karhunen-Loevova transformace (KLT), která je podrobněji vysvětlena v knize [14].

Feature selection Tento přístup spočívá ve výběru podmnožiny z množiny již existujících příznaků. Zbylé příznaky se zahodí. Výhodou oproti předchozímu přístupu je zachování původního významu příznaků, nevýhodou pak ztráta informace zahazením příznaků. Cílem je vybrat příznaky s nejlepší diskriminabilitou. Optimální metoda výběru (vyzkoušení všech možných d-tic) je prakticky nerealizovatelná. Proto se používají metody suboptimální. Vybírají se postupně příznaky takové, že již vybraná skupina je lokálně nejlepší (případně se zahazují nejhorší příznaky). Různé metody a algoritmy včetně jejich srovnání je možné nalézt např. v článku [15].

3.2.3 Metriky a indexace

Pro úspěšnou implementaci obrazového vyhledávání je mj. nutné zvolit správnou metriku pro porovnání příznakových vektorů a vyhledávací (či indexační) datovou strukturu. Metrika by měla význam jednotlivých příznaků a respektovat jejich diskriminabilitu. Indexační struktury se implementují z důvodu efektivního vyhledávání. Porovnávání příznakového vektoru dotazu se všemi vektory snímků v databázi by bylo zdlouhavé a neefektivní.

Metriky

V této kapitole si představíme nejčastěji používané metriky. Další vzdálenosti jsou uvedené např. v publikacích [13, 14]. Jednou z nejběžnějších metrik je **euklidovská vzdálenost**. Euklidovská vzdálenost mezi dvěma body x a y v n -rozměrném prostoru je definována následovně:

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n |x_i - y_i|^2}$$

případně

$$d(x, y) = \sqrt{(x - y)^T(x - y)}$$

Pro zohlednění důležitosti některých příznaků lze euklidovskou vzdálenost upravit použitím vektoru váh, které se každému příznaku přidělí. Vzniklá **vážená euklidovská vzdálenost** je pak definována takto:

$$d(x, y) = \sqrt{(x - y)^T \text{diag}(w)(x - y)}$$

Další možnou metrikou, kterou lze používat je tzv. **city-block (Manhattan) vzdálenost**:

$$d(x, y) = \sum_{i=1}^n |y_i - x_i|$$

Tuto vzdálenost si lze jednoduše představit v dvourozměrném prostoru. Pokud máme danou síť (městské bloky), tak city-block vzdálenost říká, kolik hran (ulic) musíme projít, abychom se dostali z jednoho bodu sítě do jiného.

Další používanou metrikou je **Mahalanobisova vzdálenost**:

$$d(x, y) = \sqrt{(x - y)^T \Sigma^{-1} (x - y)}$$

, kde Σ je kovarianční matice.

Indexační struktury

Pro urychlení a zefektivnění vyhledávání je výhodné použít datovou strukturu, která je přímo určená pro ukládání a práci s vícerozměrnými daty a umožňuje klást prostorové dotazy. Takovou strukturou může být třeba vícerozměrná mřížka ([37]). V praxi se ale velmi často používají různé varianty vyhledávacích stromů, např. k-d stromy, kterým je věnována kapitola v článku [8].

Guttman v roce 1984 představil R-stromy [18]. R-stromy jsou modifikací vyhledávacích B-stromů (viz [37]). Existuje několik možných modifikací těchto stromů. Oblíbenou variantou pro obrazové vyhledávání jsou R*-stromy (více v [7]). Jedná se o nejlepší dynamickou variantu R-stromu, která se velmi hodí pro indexaci příznakového prostoru.

Výzkum vhodných datových struktur však stále pokračuje. Příkladem relativně nové datové struktury mohou být SS-stromy ([44]). Vycházejí z R*-stromů a byly vymyšleny výhradně pro indexaci v systémech s obrazovým vyhledáváním.

3.3 Přehled existujících systémů

S výzkumem metod a technik obrazového vyhledávání souvisí i vývoj systémů, které tyto techniky využívají v praxi. Zmíníme si pár existujících aplikací (komerčních i výzkumných), které jsou v současné době k dispozici. Pro hlubší zájem lze doporučit přehled [43], který zahrnuje popis více než čtyřiceti systémů. U jednotlivých publikací jsou uvedeny materiály, které se projektům věnují detailněji. K některým referencím jsou rovněž přiloženy odkazy na webové stránky, které obsahují fungující demoverze systémů.

QBIC QBIC (Query By Image Content) patří mezi nejznámější systémy s obrazovým vyhledáváním vůbec. Byl vyvinut v laboratořích IBM Almaden v Silicon Valley. QBIC podporuje jak query-by-sketches, tak query-by-example. Používá různé typy příznaků: barevné histogramy v několika barevných prostorech, texturové příznaky založené na původních Tamurových příznacích, momentové invarianty a další. Metrikou je vážená euklidovská vzdálenost, pro indexaci je použit R*-strom. Podrobnosti o systému QBIC lze též nalézt v článku [31].

Chabot Chabot byl vymyšlen na univerzitě v Berkeley v Kalifornii. Spojuje vyhledávání podle textového popisku s obrazovou informací. Příkladem dotazu může být „najdi všechny západy slunce nad jezerem“. Západ slunce může být v systému definován jako převažující oranžová a žlutá na snímku. Souvislost s jezerem pak představuje vyjádření prostorové závislosti mezi objekty na obrázku (slunce je vždy nad jezerem). Veškerá data se ukládají do relační databáze POSTGRES, která rovněž pochází z Berkeley. Systém je podrobně popsán v publikaci [32].

Quicklook2 Systém původem z Itálie. Vedle textových popisků implementuje příznaky založené na barevných histogramech, Cannyho detektoru hran, momentové invarianty apod. Rovněž obsahuje zpětnou vazbu od uživatele. Uživatel má možnost rozhodnout, které obrázky ve výsledcích vyhledávání jsou žádané a který typ není. Systém pak na tomto základě může provést nové, přesnější hledání. Více viz článek [12].

NeTra Aplikace NeTra byla vytvořena na kalifornské univerzitě v Santa Barbaře. Mj. implementuje příznaky z Gaborovy wavelety. Podrobnosti v [29].

Blobworld Další systém vyvinutý na univerzitě v Berkeley. Každý snímek je segmentován do regionů (blobů). Z tvaru těchto oblastí jsou pak počítány příznaky. Uživatel má též možnost označit oblasti jeho zájmu a určit jejich význam. K indexaci je použit R*-strom. Podrobnostem je věnován článek [10].

Photobook Systém vytvořený na MIT v Cambridge, Massachusetts. Implementuje mj. rozpoznávání lidských tváří. Texturové příznaky jsou založené na Woldově dekompozici. Rovněž obsahuje zpětnou vazbu od uživatele. Systém je rozebrán v publikaci [34].

Kapitola 4

Navrhované a implementované metody

Po úvodu do problematiky obrazového vyhledávání přikročíme v této kapitole k volbě konkrétních metod, které jsou vhodné pro naše specifické zadání. Jedná se především o zvolení sady příznaků, které dobře charakterizují a vzájemně odlišují mikroskopické řezy restaurovaných děl. Zaměříme se zejména na kookurenční matice a Haralickovy příznaky. Stranou zájmu však nezůstanou ani podrobnější detaily systému vyhledávání v naší aplikaci.

V každé kapitole je pak odůvodnění použití metod, které jsou také náležitě rozebrány (včetně ilustračních příkladů).

4.1 Kookurenční matice

Kookurenční matice byly již stručně představeny v předešlé části této práce (3.2.1). V této kapitole si podrobně vysvětlíme jejich význam a výpočet prošetotónové obrázky.

Definice *Kookurenční matice M (jak je definována v [27]) je dvourozměrná matice o rozměrech $n \times n$, kde n je počet šedých odstínů v původním obrázku I (typicky 256). $M(i, j)$ pak udává četnost výskytu dvou pixelů s odstíny i a j . Dvojice pixelů je dána tzv. **vektorem posunu** (displacement vector) (dx, dy) , kde dx představuje posunutí v pixelech po x -ové souřadnici, dy obdobně pro y -ovou osu.*

Kookurenční matici s vektorem (dx, dy) budeme dále v textu značit jako $M_{dx,dy}$. Na obrázku 4.1 je ukázán výpočet kookurenční matice pro vektor posunutí $(1, 0)$ na obrázku o rozměrech 4×4 .

V praxi používané kookurenční matice jsou normované a symetrické. Symetrie znamená, že $M(i, j) = M(j, i)$. Pro vektory posunu (dx, dy) a $(-dx, -dy)$ tedy vznikne totožná matice. Vlastnosti textury mají také zajímavý vliv na řídkost kookurenční matice. Pokud

1	3	4	4
3	1	1	2
4	3	2	3
2	1	4	4

(a)

	1	2	3	4
1			+1	
2				
3	+1			
4				

(b)

	1	2	3	4
1	2	2	2	1
2	2	0	2	0
3	2	2	0	2
4	1	0	2	4

(c)

Obrázek 4.1: Ukázka výpočtu symetrické kookurenční matice s vektorem $(1,0)$ na obrázku s rozměrem 4×4 . V (a) jsou v původním obrázku vyznačeny dva pixely určené vektorem. (b) ukazuje přičtení 1 na příslušná místa v kookurenční matici. V (c) je výsledná kookurenční matice.

je textura hodně homogenní, je matice velmi řídká (malé vzdálenosti řídkost ještě umocňují), protože zkoumané páry pixelů mají podobnou barvu. V opačném případě je matice hustší.

Kookurenční matice lze s výhodou využít i pro barevné obrázky. Matice se pak může indexovat přímo barvou složenou z jednotlivých barevných kanálů. Další možností je počítat kookurenční matice pro všechny kanály zvlášť a s výsledky dále pracovat.

Samotné kookurenční matice se pro obrazové vyhledávání vzhledem ke své velikosti příliš nehodí (přestože podobnost dvou snímků na základě přímého porovnávání jejich matic je předmětem teoretického bádání, např. viz [25]). Na základě matic se proto dále počítá sada příznaků ([19]), kterým se věnuje následující kapitola.

Kookurenční matice mají vlastnosti, které jsou vhodné pro řešení naší úlohy. Zachycení prostorových závislostí mezi barvami v řezu má potenciál zaznamenat vztahy mezi jednotlivými vrstvami a též rozdílnou zrnitost a hrubost vzorků. Výběr je dále podpořen invariancí jednotlivých snímků vzhledem k měřítku. Dalším důvodem byly dobré výsledky kookurenčních matic na podobných datech. Práce věnovaná texturám hornin ([33]), které jsou podobné našim mikroskopickým řezům, dokonce na základě experimentů upřednostňuje kookurenční matice před Gaborovými příznaky (viz kapitola 3.2.1).

4.2 Haralickovy příznaky

Nyní si představíme příznaky navržené Haralickem, které je možné z kookurenčních matic vypočítat ([19, 27, 33]). Dále předpokládejme, že máme danou kookurenční matici M o rozměrech $N \times N$.

Angular Second Moment (Energy) Vyjadřuje míru hladkosti či uniformity textury. ASM je velká, pokud je textura homogenní. Z hlediska kookurenční matice to znamená

vysoký výskyt opakujících se dvojic pixelů v textuře.

$$\sum_{i=1}^N \sum_{j=1}^N M^2(i, j)$$

Contrast Lokální kontrast původního obrázku. Roste s počtem lokálních změn jasové funkce obrázku.

$$\sum_{i=1}^N \sum_{j=1}^N (i - j)^2 M(i, j)$$

Correlation Korelace. μ_x a μ_y jsou střední hodnoty vektoru, který vznikne posčítáním hodnot v řádcích resp. ve sloupcích. σ_x a σ_y jsou směrodatné odchylky.

$$\frac{\sum_{i=1}^N \sum_{j=1}^N (i * j) M(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

Inverse Difference Moment Inverzní ke kontrastu. Nabývá velkých hodnot, pokud je většina výskytů v kookurenční matici situována kolem hlavní diagonály. Tato situace nastává v případě malých lokálních změn.

$$\sum_{i=1}^N \sum_{j=1}^N \frac{1}{1 + (i - j)^2} M(i, j)$$

Entropy Entropie vyjadřuje míru náhodnosti v textuře. Nabývá vysokých hodnot, pokud je textura hrubá (všechny hodnoty v kookurenční matici jsou víceméně shodné). Entropie je tak opakem energie (ASM).

$$- \sum_{i=1}^N \sum_{j=1}^N M(i, j) \log M(i, j)$$

Variance Rozptyl

$$\sum_{i=1}^N \sum_{j=1}^N (i - \mu)^2 M(i, j)$$

Sum Average

$$\sum_{k=2}^{2N} k M_{x+y}(k)$$

, kde $M_{x+y}(k) = \sum_{i=1}^N \sum_{j=1, i+j=k}^N M(i, j)$, $k = 2, 3, \dots, 2N$.

Sum Variance

$$\sum_{k=2}^{2N} (k - SA)^2 M_{x+y}(k)$$

, kde SA je Sum Average.

Sum Entropy

$$-\sum_{k=2}^{2N} M_{x+y}(k) \log M_{x+y}(k)$$

Difference Entropy

$$-\sum_{k=0}^{N-1} M_{x-y}(k) \log M_{x-y}(k)$$

, kde $M_{x-y}(k) = \sum_{i=1}^N \sum_{j=1, |i-j|=k}^N M(i, j)$, $k = 0, 1, \dots, N - 1$.

Difference Variance

$$\text{var}(M_{x-y})$$

Information Measures of Correlation I, II

$$\frac{\text{Entropy} - H_{xy}^1}{\max\{H_x, H_y\}}$$
$$\sqrt{1 - \exp(-2(H_{xy}^2 - \text{Entropy}))}$$

, kde H_x a H_y jsou entropie hodnot vektoru, které vzniknou posčítáním řádků resp. sloupců.

$$H_{xy}^1 = -\sum_{i=1}^N \sum_{j=1}^N M(i, j) \log(M_x(i)M_y(j))$$

$$H_{xy}^2 = -\sum_{i=1}^N \sum_{j=1}^N M_x(i)M_y(j) \log(M_x(i)M_y(j)), \text{ kde } M_x(i) = \sum_{j=1}^N M(i, j) \text{ a } M_y(j) = \sum_{i=1}^N M(i, j).$$

Maximal Correlation Coefficient

druhá odmocnina druhého největšího vlastního čísla matice $Q(i, j) = \sum_{k=1}^N \frac{M(i, k)M(j, k)}{M_x(i)M_y(k)}$

Mezi nejpoužívanější příznaky s dobrou diskriminabilitou pak patří Energy (ASM), Entropy, Contrast a Inverse Difference Moment (např. [33]).

4.3 Image retrieval v Nephele

4.3.1 Zvolené příznaky

Tato kapitola je věnována výběru konkrétních příznaků a jejich testování na reálných datech (výsledky spolu s ukázkou reálných dat v databázi se nacházejí v kapitole 5). Základem obrazového vyhledávání v Nephele je porovnávání uživatelem zadané dvojice snímků v bílém a ultrafialovém světle proti obdobným dvojicím v databázi. Snímky v ultrafialovém světle obsahují informace, které v bílém světle nejsou zřejmé (např. fluorescence). Využití těchto informací by mohlo přispět k lepším výsledkům hledání. Příznaky by tedy měly být vybrány i s ohledem na tyto snímky.

Zvolení vhodných příznaků pro systém obrazového vyhledávání je motivováno dvěma základními cíli:

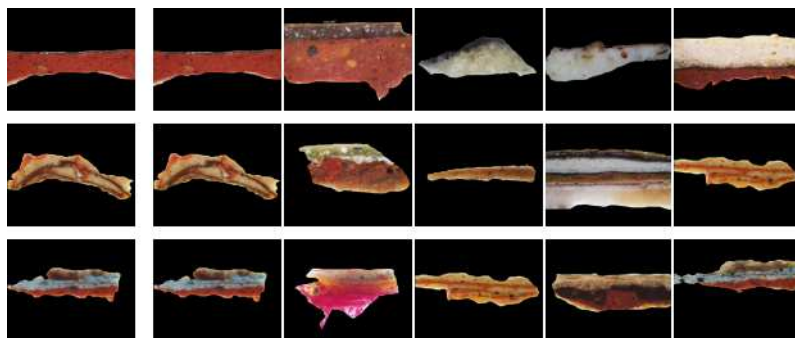
1. ve výsledcích vyhledávání by neměly chybět snímky, které jsou předloze podobné
2. výsledky vyhledávání by měly obsahovat co nejméně snímků, které neodpovídají předloze

Dosažení optimálního řešení nemusí být vždy snadné. V komplikovaném případě je možné přistoupit ke kompromisu, který by ale měl přinést vyhovující výsledky. Důležitým východním předpokladem je pak zajištění dostatečného množství testovacích dat.

Charakter mikroskopických řezů a jejich snímků vnáší do fáze výběru příznaků ještě jeden specifický problém. Tím je objektivní stanovení vizuální podobnosti jednotlivých snímků. Zůstává otázkou, zda má být jako výsledek hledání upřednostněn snímek s podobným složením vrstev před snímkem se shodným barevným zabarvením nebo zrnitostí. V případě snímků v ultrafialovém světle může být určení podobnosti ještě obtížnější. Problém by v budoucnu mohl trochu zjednodušit expertní systém, který by do rozhodování vnesl další poznatky.

Přistupme nyní k samotné volbě příznaků. Z důvodů, které jsou uvedeny v kapitole 4.1, byly k řešení vybrány kookurenční matice a Haralickovy příznaky. Kookurenční matice byly původně navrženy pro šedotónové obrázky. První experimenty byly tedy provedeny na snímcích řezů, které vznikly odstraněním barevné informace. Zvoleny byly matice s vektory posunutí $(1, 0)$, $(1, 1)$ a $(0, 1)$, z příznaků pak angular second moment, kontrast, entropie a korelace. Hned první výsledky však potvrdily, že odstranění barevné informace by nevedlo k úspěšnému řešení. Výsledky byly naprosto nevyhovující. Pochopitelně neodpovídaly barevně, neuspokojivá však byla i podobnost ve struktuře. Nadále se tedy budeme zabývat plnohodnotnými barevnými snímky. Příklady výsledků tohoto experimentu jsou přiloženy na obrázku 4.2.

Analýzou jednotlivých barevných kanálů (RGB) bylo zjištěno, že každý kanál nese různé informace, které jsou pro kvalitní vyhledávání důležité. Kookurenční matice budou dále počítány na každém kanálu zvlášť a výpočty se nebudou kombinovat, aby se zabránilo ztrátě některých údajů. Prozkoumání obrázků též ukázalo, že se nedá jednoznačně určit, zda je některý barevný kanál globálně významný či nevýznamný. Některé snímky jsou velmi heterogenní v zeleném spektru, jiné v červeném. Proto je nutné pracovat se všemi barevnými složkami snímků.



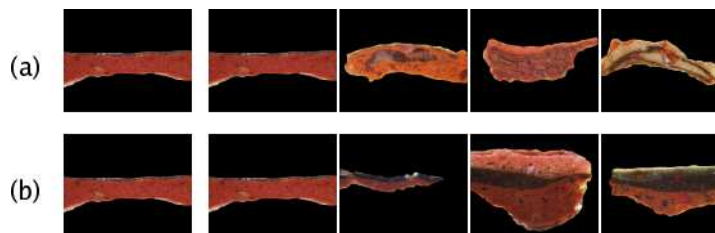
Obrázek 4.2: Příklady výsledků vyhledávání při použití kookurenčních matic, které byly vypočteny z šedotónového obrázku. Vlevo je vždy dotaz, vpravo v řádku jsou výsledky hledání. Pro přehlednost jsou uvedeny pouze snímky v bílém světle.

Výchozí pozice pro další experimenty byla následující. Bylo vybráno pět kookurenčních matic s vektory $(0, 1)$, $(1, 1)$, $(3, 0)$, $(3, 3)$ a $(0, 3)$. Matice se vzdáleností 1 jsou určeny k zachycení charakteru bezprostředního okolí jednotlivých pixelů. Vektor $(1, 0)$ byl oproti předchozímu experimentu vynechán, neboť vzorky se v horizontálním směru při takto krátké vzdálenosti téměř nemění. Matice se vzdáleností 3 nesou informaci s širším kontextem a jsou více než vhodným doplňkem. Dále byla stanovena množina šesti příznaků, s kterými budou jednotlivé experimenty prováděny. Jsou jimi kontrast, entropie, inverse difference moment (IDM), rozptyl, angular second moment (ASM) a korelace. Tato sada příznaků byla zvolena s ohledem na jejich vizuální význam. Také se jedná o příznaky, které jsou v jiných systémech používány a mají obecně největší diskriminabilitu (viz např. [20, 33]). Kookurenční matice a příznaky budou počítány na všech třech barevných kanálech jednotlivých snímků – R , G , B . Celkem máme k dispozici 180 příznaků pro každou dvojici snímků v bílém a ultrafialovém světle.

Následné experimenty se vyznačovaly postupným výběrem různých podmnožin příznaků a jejich testování na dostupných datech. Po provedení několika testů se však potvrdila hypotéza, že k úspěšné implementaci vyhledávání budou třeba i barevné příznaky. Kookurenční matice a Haralickovy příznaky sice relativně dobře odlišily řezy s jinou strukturou, problémem však zůstala barevná podobnost. Proto byly implementovány dva barevné příznaky: průměrná barva ve spektru a odpovídající směrodatná odchylka. V barevném prostoru RGB nám tak vznikne dvanáct nových příznaků.

Experiment, který použil pouze tyto barevné příznaky, prokázal dostatečnou schopnost barevně odlišit jednotlivé snímky. Podobný test byl proveden i s barevným prostorem HSV, konkrétně se složkou Hue. Výsledky však byly znatelně horší, a proto byly v sadě příznaků ponechány původní veličiny. Srovnání výsledků vyhledávání v jednotlivých barevných prostorech se nachází na obrázku 4.3.

Spojení s texturovými příznaky však nepřineslo očekávané zlepšení. Důvodem bylo zániknutí barevných příznaků v celkovém množství. Bylo nutné implementovat metriku, která by tuto nerovnováhu vyrovnala. Za tímto účelem byla zvolena **vážená euklidovská vzdálenost**:



Obrázek 4.3: Příklady výsledků vyhledávání podle barevných příznaků. (a) znázorňuje výsledky v barevném prostoru HSV (podle složky Hue). (b) znázorňuje výsledky v barevném prostoru RGB. Pro přehlednost jsou uvedeny pouze snímky v bílém světle.

$$d(x, y) = \sqrt{\sum_{i=1}^n w_i * |x_i - y_i|^2}$$

w_i je váha přiřazená příznaku i a vyjadřuje význam tohoto příznaku. Nastavením vhodných hodnot pak můžeme významně ovlivnit výsledky vyhledávání.

Pro sadu experimentů, jejichž cílem bude nastavení vhodných hodnot vah, byly vedle barevných vybrány následující Haralickovy příznaky (definované v 4.2): kontrast, IDM, entropie a rozptyl. Tyto příznaky se v předchozích experimentech ukázaly jako nejpřínosnější. Cílem stanovení jednotlivých parametrů je dosáhnout co nejlepšího vyvážení texturových a barevných příznaků a texturových příznaků mezi sebou. Tento úkol se ukázal jako velmi obtížný. Každá změna váhy totiž může ovlivnit výsledky v pozitivním i negativním smyslu. Dalším problémem je již zmíněné složité stanovení podobnosti jednotlivých dvojic snímků či určení významu ultrafialového snímku pro vyhledávání. Po nemalém počtu testů se parametry ustálily na hodnotách, při kterých vrací systém Nephele vizuálně vyhovující výsledky.

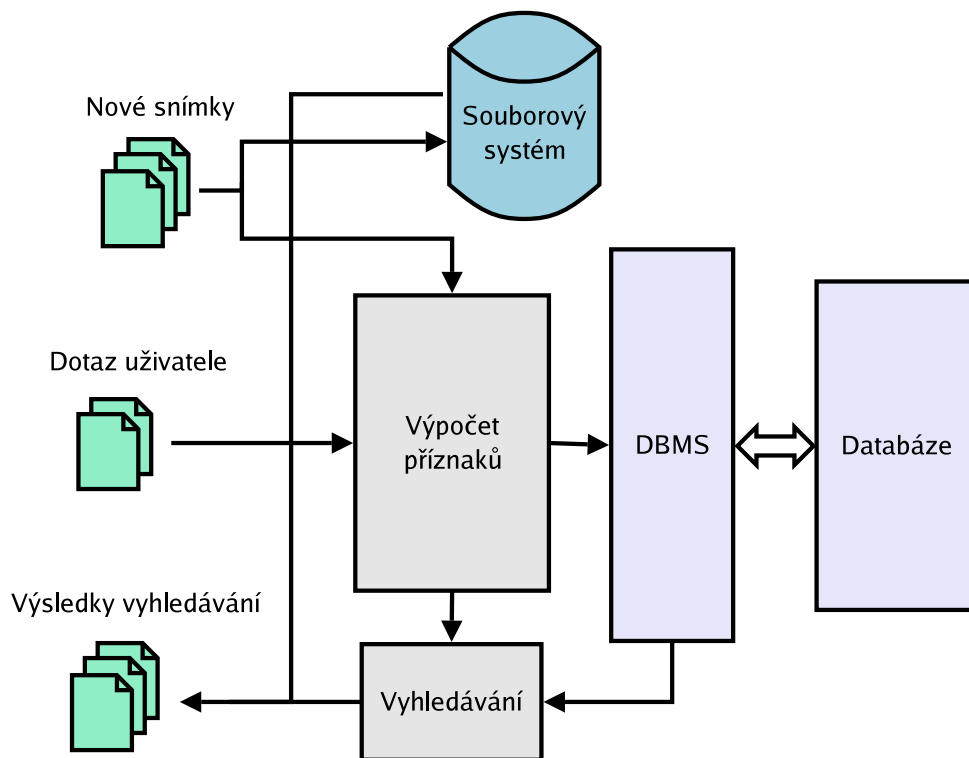
Výsledky na reálných datech jsou uvedeny v kapitole 5.

4.3.2 Vyhledávání v Nephele

Po stanovení sady příznaků si podrobněji popíšeme princip obrazového vyhledávání a indexace snímků v systému Nephele. Image retrieval lze v Nephele rozdělit na dvě části. První je výpočet příznaků na obrázcích v databázi, druhou je pak samotné vyhledávání. Situace je ilustrována na obrázku 4.4.

Příznaky nových snímků se počítají při přidání nové restaurátorské zprávy. Vyberou se všechny obrázky v bílém a ultrafialovém světle a spočítá se sada příznaků, která byla popsána v předchozí kapitole 4.3.1. Samotné snímky se uloží na vyhrazené místo na disku a vektor příznaků, který vznikne spojením spočítaných hodnot ze snímků v bílém a UV světle, se vloží do databáze. Tabulka, která je pro ně určena, obsahuje kromě samotných příznaků ještě identifikaci zprávy, k níž snímky patří (číslo zprávy a číslo vzorku).

Vyhledávání má následující průběh. Dotaz tvoří dva snímky, jeden v bílém světle a odpovídající snímek v ultrafialovém světle. Na nich se provede výpočet stejné sady pří-



Obrázek 4.4: Obrazové vyhledávání a indexace v Nephele

znaků jako při vkládání nových zpráv do databáze. Vektor příznaků se pak použije pro vyhledávání. To probíhá na základě podobnosti s vektory příznaků, které jsou uloženy v databázi. K porovnávání se používá vážená euklidovská vzdálenost. Jelikož by prohledávání celé tabulky bylo při mnoha záznamech pomalé a neefektivní, je k vyhledávání použita indexační struktura. Pro tento účel byl vybrán R^* -strom (popsán v [7]). Ve struktuře se hledá k nejbližších (nejpodobnějších) snímků, kde k je specifikováno uživatelem při zadávání dotazu.

R^* -strom je jedna z variant R -stromu (více v [18]). Jedná se o dynamickou strukturu určenou pro ukládání a vyhledávání dat ve vícerozměrném prostoru. Proto se hodí k řešení urychlení obrazového vyhledávání na základě podobnosti (vyhledání podobných obrázků se redukuje na hledání objektů ve vícerozměrném prostoru). Strom se staví vždy při startu serveru z dostupných dat. Nadále se strom již nemění.

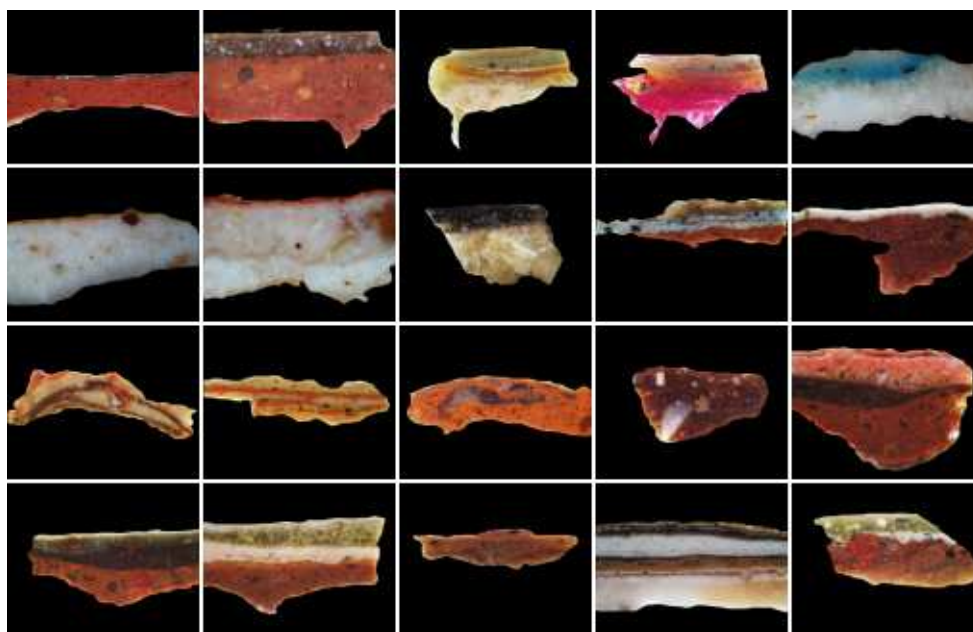
Průchodem R^* -stromu získáme identifikaci vzorků a zpráv, které představují výsledky hledání. Z databáze se vyzvednou podrobnější informace, z disku náhledy snímků a výsledek se odešle uživateli.

Kapitola 5

Výsledky na reálných datech

Tato kapitola je věnována výsledkům obrazového vyhledávání v systému Nephele. Metody navržené v kapitole 4.3.1 byly testovány na reálných datech, které byly poskytnuty Akademií výtvarného umění v Praze. Na obrázku 5.1 je přehled některých vzorků, které se v databázi nacházejí. Pro testy bylo k dispozici celkem 44 dvojic snímků v bílém a ultrafialovém světle. Ze snímků bylo ručně vysegmentováno pozadí, které by testování zvolených metod mohlo ovlivnit.

Na obrázku 5.2 jsou uvedeny některé vypočtené příznaky ke konkrétnímu snímku. Výsledky vyhledávání v systému Nephele jsou zobrazeny na obrázku 5.3. Zvolené příznaky popsané v kapitole 4.3.1 ukázaly dobré vlastnosti pro odlišení snímků jednotlivých mikroskopických řezů. Systém vyhledávání je schopný nalézt v databázi snímky, které odpovídají dotazu zadanému uživatelem. Pokud se v databázi žádný podobný snímek nevykazuje, je metoda pochopitelně méně úspěšná.



Obrázek 5.1: Výběr vzorků z databáze. Pro přehlednost jsou uvedeny pouze snímky v bílém světle.



Kontrast (M (0,1)):	394.205962
Kontrast (M (3,0)):	502.931331
Entropie (M (3,0)):	12.722056
Průměrná barva:	101.672414
Směrodatná odchylka:	45.014762



Kontrast (M (0,1)):	54.248786
Kontrast (M (3,0)):	74.986013
Entropie (M (3,0)):	10.541202
Průměrná barva:	136.455075
Směrodatná odchylka:	22.613052

Obrázek 5.2: Příklad vypočtených příznaků ze dvou rozdílných obrázků (přesněji z jejich zeleného spektra). V závorce u kontrastu a entropie je uveden displacement vektor použité kookurenční matice.



Obrázek 5.3: Příklady výsledků obrazového vyhledávání v Nephele. V levém sloupci jsou dvojice snímků v bílém a UV světle, které byly systému kladeny jako dotaz. V odpovídajících řádcích jsou seřazeny výsledky hledání podle podobnosti. První sloupec ve výsledcích je identický dotazu (jsou to nejpodobnější snímky).

Kapitola 6

Závěry

Cílem této práce byl návrh databáze restaurátorských zpráv a systému, který by umožnil v záznamech vyhledávat jak podle textové, tak podle obrazové informace. Databáze by měla poskytnout restaurátorům snadný přístup k informacím o již restaurovaných dílech. Akademie výtvarného umění v Praze zatím žádný takový software nevlastní.

Stěžejním úkolem návrhu systému obrazového vyhledávání byl průzkum dostupných metod a přístupů k obrazovému vyhledávání. Po studiu literatury bylo zvoleno řešení na bázi příznaků z kookurenčních matic. Metoda byla prostudována a vybrané příznaky implementovány. Zvoleny byly Haralickovy příznaky kontrast, entropie, korelace a rozptyl. K vyhledávání byl použit R*-strom. Dále byla navržena aplikace, která vedle obrazového vyhledávání, poskytuje správu restaurátorských zpráv v databázi. Bylo implementováno přidávání nových zpráv, textové vyhledávání apod.

Systém Nephele představuje dobrý základ softwaru, který bude předán laboratoři AVU v Praze. Po rozšíření o další potřebné funkce by měl být nasazen do reálného prostředí. Výsledky a metody obrazového vyhledávání také byly prezentovány na konferenci ICIP 2005 v Janově.

Námětem do budoucnosti je další vylepšení aplikace. Jedná se zejména o rozšíření funkčnosti celého systému a začlenění dalších metod digitálního zpracování obrazu, které by mohly pomoci restaurátorům v jejich práci. Samotné restaurátorské zprávy rovněž procházejí vývojem a systém by měl na tyto změny reagovat. Počítá se i s rozšířením počtu prováděných analýz. Dlouhodobějším cílem je vytvoření expertního systému, který by některé úkony spojené s přípravou restaurace a rozboru uměleckého díla prováděl bez asistence odborníka.

Příloha A

Implementační detaily

Příloha je věnována implementačním detailům systému Nephele. V žádném případě se nejedná o plnohodnotnou programátorskou dokumentaci. Jsou tu zdůrazněny zajímavé aspekty návrhu a implementace.

A.1 Prostředí

Aplikace Nephele se skládá ze dvou relativně nezávislých částí: klienta a serveru. Klient především zprostředkovává komunikaci s uživatelem pomocí grafického rozhraní. Role serveru je ve zpracování dotazů a komunikaci s databází.

Pro implementaci klienta bylo zvoleno prostředí Visual Studio .NET a jazyk Managed C++. Platforma operačního systému Windows byla předem dána vybavením počítačů, na kterých systém v budoucnu poběží. Hlavním důvodem pro volbu VS .NET byl relativně snadný a rychlý návrh GUI.

Server je implementován v jazyce C/C++ a běží v prostředí operačního systému Linux. Ten byl zvolen zejména pro svou stabilitu a bezpečnost, což jsou nejdůležitější požadavky na finální produkt. Server komunikuje s databází MySQL přes rozhraní ODBC. ODBC umožňuje snadno vyměnit databázi MySQL za jinou relační databázi.

Klient a server spolu komunikují přes síťové sockety. Posílaná data jsou ve formátu XML. Výhodou je možnost implementace zcela nového klienta v jiném prostředí. Příkladem může být potenciální webový klient v jazyce PHP. Celá komunikace je zabezpečena protokolem SSL.

A.2 Klient a server

Následující kapitoly jsou věnovány podrobnějšímu rozboru jednotlivých částí.

A.2.1 Komunikační rozhraní

Klient a server mezi sebou posílají data. Klient typicky pošle dotaz (případně s dalšími údaji) a server odešle odpovídající odpověď. Tato data jsou zapsána v jazyce XML. Nevýhodou XML je jeho zaměření výhradně na textová data. Přenos binárních dat (obrázky, soubory apod.) či souvislejších textů (např. zdrojové kódy, které obsahují pro XML nebezpečné znaky) se tedy může jevit jako obtížný problém. Řešením je kódování Base64. Binární data se překódují do Base64, a tak se mohou přenášet jako pouhý text.

Struktura XML je pevně daná a vypadá následovně:

```
<nephele>
  <head>
    <xmltype></xmltype>
    <errmsg></errmsg>
  </head>
  <body>
    <block tag='''>
      <row>
        <item>
          <name></name>
          <type></type>
          <data></data>
        </item>
        ...
      </row>
      ...
    </block>
    ...
  </body>
</nephele>
```

Vysvětlení jednotlivých položek:

- **head** – neboli hlavička. Nese informace o celém XML.
 - *xmltype* – typ xml, resp. identifikace dat v XML. Liší se v závislosti na směru, kterým data cestují. Klient do *xmltype* vyplňuje identifikaci dotazu. Podle této informace pak server pozná, jak má s daty naložit. Server nastaví *xmltype* na jednu ze tří možností: *error* (při zpracování dotazu došlo k chybě), *reply* (odpověď s daty), *replyok* (potvrzovací odpověď, zpravidla bez dalších dat).
 - *errmsg* – v případě chyby serveru nese tato položka chybový řetězec
- **body** – tělo XML. Obsahuje veškerá data.
 - *block* – blok dat jednoho typu. Hierarchicky odpovídá tabulce v databázi. *Tag* pak říká, jaká data blok obsahuje.
 - *row* – sada položek v rámci jednoho bloku. Odpovídá řádku v tabulce.
 - *item* – jednotlivé položky. *Name* je název položky, *type* typ položky a *data* obsahuje data.

A.2.2 Klient

Cílem klienta je zpřístupnění funkcí celého systému uživateli. Děje se tak skrze grafické prostředí. Uživatel se do systému nejprve přihlásí svým uživatelským jménem a heslem. Administrátorem systému má přidělena práva, která umožní provádět povolené akce. Patří sem vyhledávání záznamů, přidávání záznamů a další. Pokud je uživatel administrátorem, má k dispozici další funkce (zakládání a správa uživatelských účtů, správa přístupových práv).

Uživatel se pohybuje mezi jednotlivými formuláři systému. Pokud jsou k další činnosti potřebná data od serveru, sestaví se příslušné XML a dotaz se odešle. Po příchodu odpovědi se data zpracují a zobrazí uživateli. Činnost končí odhlášením ze systému. Celý cyklus je ilustrován na obrázku A.1.



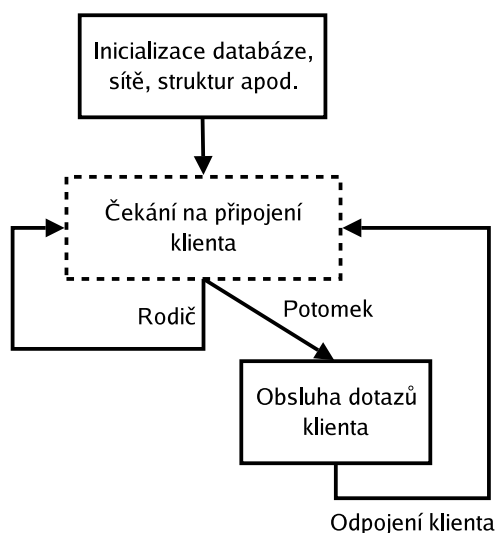
Obrázek A.1: Běh klienta

A.2.3 Server

Server obsluhuje veškeré dotazy, které přichází od klientů. Při svém startu zavolá funkci `fork()` a rodičovský proces se ukončí. Tím se server odsune do pozadí, kde dále běží a čeká na připojení klientů. Poté server načte konfigurační soubor, který obsahuje různá nastavení systému. Následují inicializace jednotlivých částí systému: otevření logu, ošetření signálů, inicializace a připojení k databázi (MySQL) a inicializace sítě. K té patří vedle vytvoření socketu a začátku naslouchání na určeném portu také inicializace protokolu SSL, kterým je přenos dat po síti zabezpečen.

Inicializace pokračuje nahráním dostupných dynamických knihoven (viz A.3) a postavením R*-stromu, který bude potřeba pro obrazové vyhledávání. Stavba spočívá v získání všech záznamů z tabulky určené pro ukládání příznakových vektorů a jejich uložení do stromové struktury.

Poté server čeká na připojení klientů. Pokud se na socket nějaký klient připojí, server opět zavolá funkci `fork()` a komunikaci s klientem předá svému potomkovi. Ten zdědí prostředí svého rodiče, a tedy i vyhledávací strom. Rodič pak čeká na připojení dalšího klienta. Počet potomků, které server může mít, je omezen konfigurací. Pokud je tento počet na svém maximu, server čeká, dokud se nějaký klient neodpojí. Běh serveru a rodiče ilustruje schéma A.2.



Obrázek A.2: Běh serveru

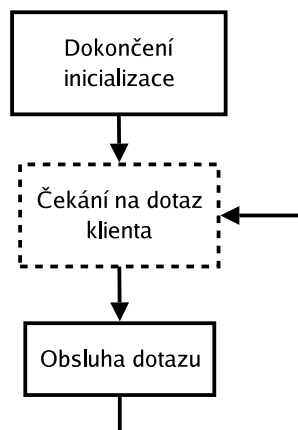
Úkolem potomka je obsloužit všechny dotazy klienta. Nejprve dokončí inicializaci některých hodnot a čeká na dotazy od klienta. Po příchodu dotazu naparsuje část XML a zjistí typ požadavku (z hodnoty *xmltype*). Potomek se podívá do databáze a zjistí, která funkce z jaké knihovny je pro obsluhu dotazu určena. Pokud byla knihovna zavedena do paměti při startu serveru a funkce se v ní nachází, je tato funkce zavolána. Pokud knihovna zavedena nebyla, nebo pokud se funkce v již zavedené knihovně nenachází, pokusí se server danou knihovnu zavést do paměti, resp. načtení opakovat. Tento postup umožňuje přidávání obslužných funkcí za běhu bez nutného restartu serveru.

Funkce zpracuje dotaz klienta. Naparsuje XML, provede potřebné úkony a vrátí výsledné XML v odpověď. Potomek toto XML odešle a čeká na další dotazy. Pokud je komunikace přerušena (chyba, ukončení klienta), provede se odhlášení uživatele a proces se ukončí. Schéma A.3 ukazuje činnost potomka.

A.3 Podpůrné knihovny

Tato kapitola se věnuje popisu použitých knihoven.

- **MySQL 4.1** – v tomto případě se nejedná o použitou knihovnu. MySQL je relační databáze, kterou systém Nephele v současné době využívá. Díky rozhraní ODBC je ale přechod na jinou databázi stejného typu bezproblémový.
- **unixODBC 2.2, MyODBC 3.51** – unixODBC poskytuje API rozhraní ODBC. MyODBC je ODBC driver pro MySQL.

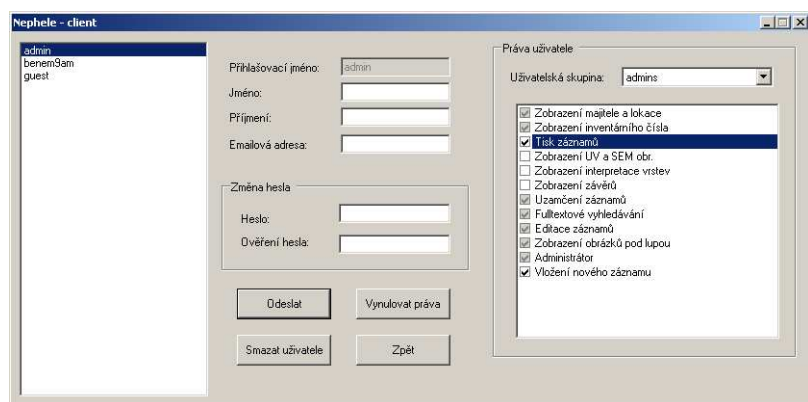


Obrázek A.3: Činnost potomka

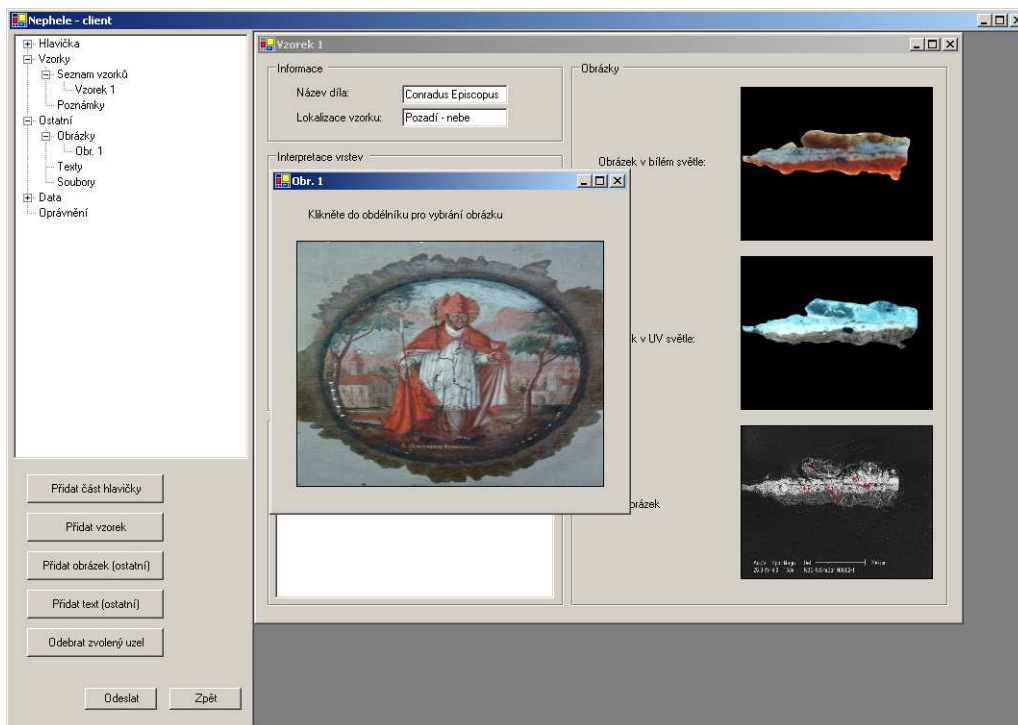
- **OpenSSL 0.7.9g** – knihovna implementující protokoly SSL (secure sockets layer) verze 2 a 3 a TLS (transport layer security) verze 1. Zároveň obsahuje silnou kryptografickou část. Nephele tuto knihovnu používá pro zabezpečení komunikace mezi klientem a serverem.
- **Xerces 2.6.0, ICU 3.2** – Xerces je rozsáhlá knihovna pro parsování a validaci XML dokumentů. Poskytuje rozhraní DOM i SAX. Je zkompileován s podporou knihovny ICU (International Component for Unicode) pro zajištění správného zacházení s kódováním UTF-8.
- **ImageMagick 6.2.3** – knihovna pro práci s obrázky.
- **libVIPS 7.10.10** – knihovna, která implementuje vybrané algoritmy digitálního zpracování obrazu. Některé části kódu byly upraveny a použity v Nephele.
- **tools 0.43b, spatialindex 0.83b** – knihovny, které mj. implementují R*-stromy.

A.4 Uživatelské rozhraní

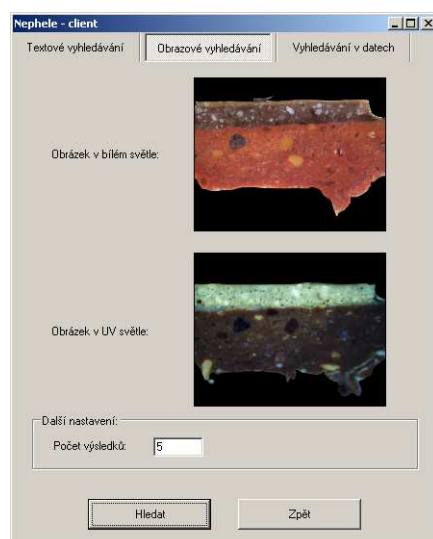
Tato kapitola obsahuje několik snímků klienta a uživatelského prostředí.



Obrázek A.4: Editace uživatelských údajů a práv



Obrázek A.5: Přidávání nového záznamu



Obrázek A.6: Formulář pro obrazové vyhledávání

Příloha B

Rejstřík obrázků a tabulek

Na dalších stránkách se nachází seznam obrázků a tabulek, které jsou v textu použity.

Seznam obrázků

1.1	Snímky mikroskopických řezů	5
2.1	Umělecké dílo s vyznačenými místy odběru vzorků	6
2.2	Snímek z elektronického mikroskopu (SEM snímek)	7
2.3	Zjednodušené ER-schéma pro restaurátorskou zprávu	11
2.4	Architektura Nephele	13
3.1	Příklad vyhledávání podle podobnosti (Blobworld)	17
3.2	Schéma obrazového vyhledávání (QBE, QBS)	18
3.3	Příklady textur	20
3.4	Histogram obrázku v grafickém programu Gimp (www.gimp.org)	21
3.5	Textury ilustrující Tamurovy příznaky	22
4.1	Ukázka výpočtu kookurenční matice	30
4.2	Příklady výsledků vyhledávání při použití šedotónových kookurenčních matic	34
4.3	Příklady výsledků vyhledávání podle barevných příznaků	35
4.4	Obrazové vyhledávání a indexace v Nephele	36
5.1	Výběr vzorků z databáze	37
5.2	Příklad vypočtených příznaků z konkrétního obrázku	38
5.3	Příklady výsledků obrazového vyhledávání v Nephele	39
A.1	Běh klienta	43
A.2	Běh serveru	44
A.3	Činnost potomka	45
A.4	Editace uživatelských údajů a práv	46
A.5	Přidávání nového záznamu	46
A.6	Formulář pro obrazové vyhledávání	47

Seznam tabulek

2.2	Tabulky navržené pro ukládání restaurátorských zpráv	10
2.4	Přístupová oprávnění	15
2.6	Tabulky pro zajištění chodu serveru	15

Literatura

- [1] Akademická laboratoř materiálového průzkumu malířských děl. <http://www.alma-lab.cz>.
- [2] National Gallery of Art, Washington D.C. <http://www.nga.gov>.
- [3] Národní galerie v Praze. <http://www.ngprague.cz>.
- [4] Referenční manuál k databázovému serveru MySQL. <http://dev.mysql.com/doc/>.
- [5] The National Gallery, London. <http://www.nationalgallery.org.uk>.
- [6] Vision Texture Database. <http://www-white.media.mit.edu/vismod/imagery/VisionTexture/vistex.html>.
- [7] N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger. The R*-tree: An efficient and robust access method for points and rectangles. *Proceedings of the 1990 ACM SIGMOD International Conference on Management of Data*, 19:322–331, 1990.
- [8] J. Bentley and J. Friedman. Data structures for range searching. *ACM Computing Surveys*, 11(4):397–409, 1979.
- [9] A. D. Bimbo and P. Pala. Visual image retrieval by elastic matching of user sketches. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(2):121–132, 1997.
- [10] C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, and J. Malik. Blobworld: A system for region-based image indexing and retrieval. In *Proc. 3rd International Conference on Visual Information and Information Systems, Lecture Notes in Computer Science*, 1614:509–516, 1999. <http://elib.cs.berkeley.edu/photos/blobworld>.
- [11] C.-C. Chang and C.-F. Lee. Relative coordinates oriented symbolic string for spatial relationship retrieval. *Pattern Recognition*, 28(4):563–570, 1995.
- [12] G. Ciocca, I. Gagliardi, and R. Schettini. Quicklook2: An integrated multimedia system. *Journal of Visual Languages and Computing*, 12:81–103, 2001.
- [13] T. Deselaers, H. Ney, T. Seidl, and D. Keysers. Features for image retrieval. <http://www-i6.informatik.rwth-aachen.de/~deselaers/files/deselaers-features-for-ir.pdf>.
- [14] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification, 2nd Edition*. Wiley Interscience, 2000.
- [15] F. Ferri, P. Pudil, M. Hatef, and J. Kittler. Comparative study of techniques for large-scale feature selection. *Pattern Recognition in Practice*, 4:403–413, 1994.

- [16] J. Flusser and T. Suk. Pattern recognition by affine moment invariants. *Pattern Recognition*, 26(1):167–174, 1993.
- [17] J. Flusser and B. Zitová. Přednáška Digitální zpracování obrazu. MFF UK, Praha.
- [18] A. Guttman. R-tree: A dynamic index structure for spatial searching. *Proceedings of ACM SIGMOD International Conference on Management of Data*, 14:47–57, 1984.
- [19] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(6):610–621, 1973.
- [20] P. Howarth and S. Ruger. Evaluation of texture features for content-based image retrieval. *Proceedings of International Conference on Image and Video Retrieval*, pages 326–334, 2004.
- [21] D. Hradil and J. Hradilová. Minerály starých obrazů. *Vesmír*, 7:383–385, 2002.
- [22] M.-K. Hu. Visual pattern recognition by moment invariants. *IEEE Transactions on Information Theory*, 8(2):179–187, 1962.
- [23] C.-L. Huang and D.-H. Huang. A content-based image retrieval system. *Image and Vision Computing*, 16:149–163, 1998.
- [24] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih. Image indexing using color correlograms. *Proc. of 1997 Conf. on Computer Vision and Pattern Recognition*, pages 762–768, 1997.
- [25] R. Kašpar, L. Petrů, B. Zitová, J. Flusser, and J. Hradilová. Microscopic cross-sections of old artworks. *Přijato na ICIP 2005, v září v Janově, Itálie*, 2005.
- [26] V. Kovalev and S. Volmer. Color co-occurrence descriptor for querying-by-example. *Multimedia Modeling*, pages 32–38, 1998.
- [27] A. S. Kurani, D.-H. Xu, and J. Furst. Co-occurrence matrices for volumetric data. <http://shrike.depaul.edu/~dxu/Papers/iasted-06-01.pdf>.
- [28] S.-Y. Lee and F.-J. Hsu. Spatial reasoning and similarity retrieval of images using 2D C-String knowledge representation. *Pattern Recognition*, 25(3):305–318, 1992.
- [29] W.-Y. Ma and B. Manjunath. Netra: A toolbox for navigating large image databases. *Multimedia Systems*, 7:184–198, 1999. <http://vision.ece.ucsb.edu/netra>.
- [30] B. Manjunath and W. Ma. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:837–842, 1996.
- [31] W. Niblack, R. Barber, and et al. The QBIC project: Querying images by content using color, texture and shape. *Proceedings of SPIE Storage and Retrieval for Image and Video Databases*, pages 173–187, 1993. <http://wwwqbic.almaden.ibm.com>.
- [32] V. E. Ogle and M. Stonebraker. Chabot: Retrieval from a relational database of images. *IEEE Computer*, 28(9):40–48, 1995. <http://http.cs.berkeley.edu/~ginger/chabot.html>.

- [33] M. Partio, B. Cramariuc, M. Gabbouj, and A. Visa. Rock texture retrieval using gray level co-occurrence matrix. *Proc. of 5th Nordic Signal Processing Symposium*, 2002.
- [34] A. Pentland, R. W. Picard, and S. Sclaroff. Photobook: Content-based manipulation of image databases. *International Journal of Computer Vision*, 18(3):233–254, 1996. <http://vismod.media.mit.edu/vismod/demos/photobook/index.html>.
- [35] J. Peterka. Přednášky Počítačové sítě a Rodina protokolů TCP/IP. <http://www.earchiv.cz>.
- [36] L. Petru. Použití metod digitálního zpracování obrazu pro restaurování uměleckých děl. Master's thesis, MFF UK, Praha, 2003.
- [37] J. Pokorný. *Základy implementace souborů a databází*. Vydavatelství Karolinum, Univerzita Karlova, 1997.
- [38] Y. Rui, T. S. Huang, and S.-F. Chang. Image retrieval: Current techniques, promising directions and open issues. *Journal of Visual Communication and Image Representation*, 10:123–135, 1999.
- [39] H. V. Ríos, A. Santamaría, C. G. Maldonado, and J. Pena. Visual information retrieval for images in digital libraries plants of Mexico. Zdroj - osobní kontakt.
- [40] S.-O. Shim and T.-S. Choi. Image indexing by modified color co-occurrence matrix. *ICIP*, 3:493–496, 2003.
- [41] H. Tamura, S. Mori, and T. Yamawaki. Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man and Cybernetics*, 8:460–473, 1978.
- [42] M. Turner. Texture discrimination by Gabor functions. *Biological Cybernetics*, 55:71–82, 1986.
- [43] R. C. Veltkamp and M. Tanase. Content-based image retrieval: A survey. <http://www.aa-lab.cs.uu.nl/cbirsurvey/>.
- [44] D. White and R. Jain. Similarity indexing with SS-tree. *Proceedings of 12th ICDE*, pages 516–523, 1996.
- [45] D. Zhang and G. Lu. A comparative study of fourier descriptors for shape representation and retrieval. *Proc. of 5th Asian Conference on Computer Vision*, 2002.