

Iskanje podobnih enobarvnih fragmentov pri množičnem sestavljanju stenskih poslikav

Gregor Čepin^{1,2}, Tea Tušar², Bogdan Filipič²

¹Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Tržaška cesta 25, 1000 Ljubljana

²Institut "Jožef Stefan", Jamova cesta 39, 1000 Ljubljana

E-pošta: gregor.cepin@gmail.com, tea.tusar@ijs.si, bogdan.filipic@ijs.si

Searching for similar unicolored fragments in crowdsourced reassembly of wall paintings

Reassembly of a large number of wall painting fragments is a very time-demanding task involving manipulation of delicate fragments by professional restorators. In order to speed up this task, we recently introduced the mobile application e-Pedius in which digital images of fragments can be reassembled by crowds. To enhance the application's performance, unicolored fragments need to be combined into sprite sheets according to their similarity. This paper presents how this is performed—from the definition of similarity between fragment images to their grouping using hierarchical clustering. We demonstrate the results on the "Turška mačka" restoration project with 5162 unicolored fragments.

1 Uvod

Ohranjanje kulturne dediščine med drugim obsega tudi rekonstrukcijo umetnin, ki so bile uničene zaradi časovnega propadanja, nesreč ali katastrof (potresov in vojn). Na arheoloških najdiščih pogosto najdemo stenske poslikave, ki so pomembni elementi kulturne dediščine, a so večinoma fragmentirane in potrebujejo temeljito obnovo. Restavriranje stenskih poslikav iz fragmentov je zaradi njihovega velikega števila (tudi več tisoč), različnih oblik in velikosti ter poškodb in manjkajočih fragmentov zelo težavno in zahteva leta ročnega dela strokovnjakov.

Medtem ko je za večino faz restavriranja fragmentiranih stenskih poslikav (npr. odstranjevanje nečistoč, utrjevanje ometov, retuširanje) neizbežno rokovanje s fizičnimi fragmenti, se lahko faza sestavljanja fragmentov v prvotno poslikavo izvede najprej elektronsko, z digitalnimi slikami fragmentov. Šele ko na ta način najdemo zadovoljivo elektronsko postavitev, fizične fragmente sestavimo skupaj v peskovniku. Na Institutu "Jožef Stefan" smo v sodelovanju z Zavodom za varstvo kulturne dediščine Slovenije (ZVKDS) razvili dve računalniški aplikaciji, ki podpirata elektronsko sestavljanje fragmentov in tako bistveno pohitrita postopek restavriranja stenskih poslikav. Računalniška aplikacija *Pedius* restavratorjem omogoča digitalizacijo, evidentiranje in pomoč pri sestavljanju fragmentov v prvotno stensko poslikavo [2, 4], medtem ko je spletna in

mobilna aplikacija *e-Pedius* zasnovana kot igra, ki sestavljanje fragmentov prepusti množici [1].

V aplikaciji *e-Pedius* je sestavljanje stenskih poslikav s številnimi fragmenti razdeljeno na manjše naloge, ki vsebujejo tako vzorčaste kot enobarvne fragmente¹. Naloge restavratorji oblikujejo ročno in le na podlagi vzorčastih fragmentov (fragmenti s podobnim vzorcem so vključeni v isto nalogo), saj se za enobarvne fragmente ne ve zagotovo, h kateri nalogi sodijo, in jih zato lahko dodamo h katerikoli nalogi.

Zaradi zagotavljanja hitrega delovanja mobilne aplikacije fragmente predhodno združimo v *zlepke sličic* (angl. sprite sheets). Vzorčaste fragmente združujemo v zlepke na podlagi nalog, enobarvne pa na podlagi njihove medsebojne podobnosti. Ker je enobarvnih fragmentov lahko zelo veliko (tudi več tisoč), želimo iskanje podobnih enobarvnih fragmentov kar se da avtomatizirati. V prispevku opisujemo, kako poteka iskanje podobnih enobarvnih fragmentov in njihovo združevanje v skupine, iz katerih na koncu tvorimo zlepke.

V nadaljevanju prispevka na kratko opišemo aplikaciji *Pedius* in *e-Pedius*. Tretji razdelek predstavlja zahteven projekt restavriranja stenskih poslikav imenovan *Turška mačka*. V četrtem razdelku so opisane uporabljene metode za iskanje podobnih fragmentov in njihovo združevanje v skupine. V petem razdelku so predstavljeni rezultati na projektu *Turška mačka*. Sklepni del vsebuje povzetek in zamisli za nadaljnje delo.

2 Aplikaciji *Pedius* in *e-Pedius*

2.1 Aplikacija *Pedius*

Aplikacija *Pedius* [2, 4] deluje na osebnih računalnikih in je namenjena restavratorjem ZVKDS. Digitalizacijo fragmentov omogoča z uporabo bodisi optičnih čitalnikov (to je primerno le za majhne in lahke fragmente) bodisi digitalnih fotoaparátov. Podpira tudi nadaljnjo obdelavo tako zajetih slik, pri kateri odstranjujemo črno ozadje in obrobe slik fragmentov, ki ustrezajo delcem ometa. Obdelane slike in podatki o fragmentih se shranjujejo v podatkovno bazo.

¹Fragmenti, ki jih manipulirajo računalniške aplikacije so seveda zgolj digitalne slike pravih (fizičnih) fragmentov.

Na osnovi te računalniške predstavitve fragmente v aplikaciji Pedius lahko elektronsko sestavljamo v grafičnem vmesniku, kjer ob strani izbiramo fragmente, na osrednjem delu pa jih premikamo, vrtimo in združujemo. Na voljo je tudi iskanje fragmentov po različnih kriterijih (npr. po velikosti, barvah fragmenta, vzorcih).

Aplikacija Pedius restavratorjem precej olajša delo, a ker je sestavljanje več tisoč fragmentov težka naloga, vseeno zahteva veliko časa. Za razbremenjevanje strokovnjakov smo v nadaljevanju razvili aplikacijo e-Pedius, ki sestavljanje omogoča tudi laičnim množicam.

2.2 Aplikacija e-Pedius

Aplikacija e-Pedius [1] je implementirana kot igra sestavljanja fragmentov in deluje kot domorodna mobilna aplikacija na tabličnih računalnikih z operacijskimi sistemi iOS, Android in Windows RT ter kot spletna aplikacija na ostalih platformah s tehnologijo HTML5. Temelji na konceptu *množičnega izvajanja* (angl. crowdsourcing), pri katerem večje število uporabnikov rešuje probleme, ki bi manjši skupini ljudi vzeli veliko časa. Uporabnike motivira s točkami in značkami, ki jih dobijo po uspešno opravljenem delu. Ker pravilna postavitev fragmentov ni znana vnaprej, tudi ocenjevanje postavitev prepustimo uporabnikom. Najbolje ocenjene postavitve bodo restavratorji lahko uporabili pri določanju pravih (dokončnih) postavitev.

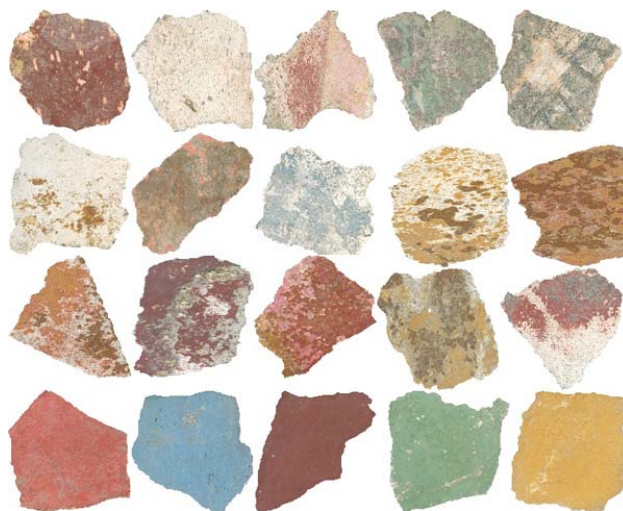
Sestavljanje fragmentov je v aplikaciji e-Pedius razdeljeno na manjše naloge, ki so primerne za sestavljanje na tabličnih računalnikih. Vsaka naloga vsebuje vnaprej določene vzorčaste fragmente in možnost izbire primernih enobarvnih fragmentov, ki so vnaprej neznani. Ko uporabnik s kapalko označi del danega fragmenta, se mu prikažejo enobarvni fragmenti, ki so izbranemu delu najbolj podobni. Ti se lahko uporabijo za dopolnjevanje postavitev, ki so bile sestavljene samo iz vzorčastih fragmentov. Na ta način lahko uporabnik reši nalogo z uporabo tako vzorčastih kot enobarvnih fragmentov.

Da je delovanje aplikacije dovolj hitro, slike fragmentov vnaprej združimo v zlepke sličic. Slike vzorčastih fragmentov združujemo na podlagi nalog, enobarvnih pa na podlagi njihove medsebojne podobnosti, kar zmanjšuje število zlepkov sličic, ki jih mora aplikacija prenesti na tablični računalnik ob izboru s kapalko. To je pomemben korak, saj imamo opraviti z obsežnimi poslikavami z več tisoč fragmenti, kot je podrobneje predstavljeno v nadaljevanju.

3 Restavratorski projekt Turška mačka

Čeprav sta bili obe aplikaciji zasnovani kar se da neodvisno od posameznih nalog sestavljanja, je bila glavna motivacija zanj projekt Turška mačka.

Eno mnogih izkopavanj v rimski Celeii (današnjem Celju) je razkrilo ostanke rimskih bivališč iz prvega stoletja našega štetja. V najdišču imenovanem Turška mačka so leta 1978 našli več kot 9000 fragmentov stenskih poslikav. Zaradi velike koncentracije fragmentov na enem samem mestu še zdaj ni jasno, ali so bile vse poslikave v isti



Slika 1: Digitalne slike enobarvnih fragmentov

zgradbi ali pa so arheologi našli odpadno jamo, v katero so Rimljani odlagali material pri obnovah. Fragmenti so različnih barv, nekateri vsebujejo motive, kot so živali in rože, razločimo lahko številne okrasne vzorce.

Izkopani fragmenti se zdaj nahajajo v Restavratorskem centru ZVKDS, kjer so že pred leti izvedli začetne faze restavriranja fragmentov. Prvi poskusi ročnega sestavljanja fragmentov projekta Turška mačka segajo že 30 let nazaj, vendar so sestavljanje opustili, saj je bilo preveč zamudno in praktično neizvedljivo. Velikost peskovnika je namreč omejena in težko je naenkrat pregledati in sestavljati zadostno število fragmentov. Šele aplikacija Pedius je omogočila nadaljevanje tega projekta. ZVKDS je z njo zagotovil digitalizirane predstavitve vseh 9521 fragmentov, ki so zdaj na voljo za sestavljanje tudi preko aplikacije e-Pedius. Iz njih so restavratorji določili več kot 100 nalog sestavljanja.

Fragmente imenujemo vzorčasti, če vsebujejo vzorce, črte ali kombinacije barv. Preostalim fragmentom pravimo enobarvni fragmenti. V njih prevladuje en barvni odtenek, ta barva pa ponavadi variira zaradi obdelosti, vsebuje barvo ometa zaradi obrabljenosti ali vsebuje tudi manjše količine drugih barv, ki pa niso dovolj izrazite, da bi fragmente lahko prepoznali kot vzorčaste. Slika 1 prikazuje nekaj enobarvnih fragmentov projekta Turška mačka.

4 Iskanje podobnih enobarvnih fragmentov

Iskanje podobnih enobarvnih fragmentov v aplikaciji e-Pedius obsega dva koraka. Najprej definiramo, kdaj sta si dva fragmenta podobna, nato pa vse enobarvne fragmente združimo v skupine, ki vsebujejo med seboj čim bolj podobne fragmente.

4.1 Podobnost fragmentov

Slike fragmentov lahko predstavimo z več barvnimi modeli, ki se različno dobro obnesejo pri računalniški primerjavi slik. Uporabili smo tri barvne modele: RGB, HSV in

LAB. Vsi predstavijo točko na sliki s tremi vrednostmi oziroma barvnimi kanali. Vrednost vsakega barvnega kanala v neki točki je v pomnilniku računalnika zapisana z osmimi biti, tako da ima vsak kanal 256 možnih vrednosti.

V modelu RGB je v vsaki točki podana vrednost rdeče (R), zelene (G) in modre (B) barve. V modelu HSV so v vsaki točki zapisane naslednje vrednosti: barvni odtенок (H), intenzivnost barve (S), ki pove, kje med popolnoma belo in popolnoma čisto barvo se nahaja izbrana barva, in svetlost (V), ki določa zatemnjenost med popolnoma črno in popolnoma čisto barvo. V modelu LAB vrednost L predstavlja svetlost točke, vrednosti A in B pa položaj v barvni ravnini. Ustvarjen je bil z namenom, da bi posnel človeško dojemanje barv. Tako imata dve barvi, ki se ljudem zdita podobni, majhno razliko vrednosti v modelu LAB.

Sliki posameznega fragmenta smo opisali s histogrami, ki povedo, koliko točk na sliki ima določeno vrednost barvnega kanala. Ker imamo v vseh barvnih modelih po tri barvne kanale, ustrezajo vsaki sliki fragmenta trije histogrami. Ker ima vsak barvni kanal 256 možnih vrednosti, je največja velikost histograma 256 polj, z združevanjem sosednjih polj pa jih dobimo manj. Preizkusili smo histograme z 256, 128 in 64 polji.

S histogrami opisane slike fragmentov lahko primerjamo med seboj z različnimi metrikami. Preizkusili smo naslednje [3]: evklidsko razdaljo, manhattansko razdaljo, Hellingerjevo razdaljo, razdaljo χ^2 , korelacijo med histogramoma ter manhattansko razdaljo na kumulativnih histogramih. Zaradi omejitve prostora predstavljamo le formulo za Hellingerjevo razdaljo:

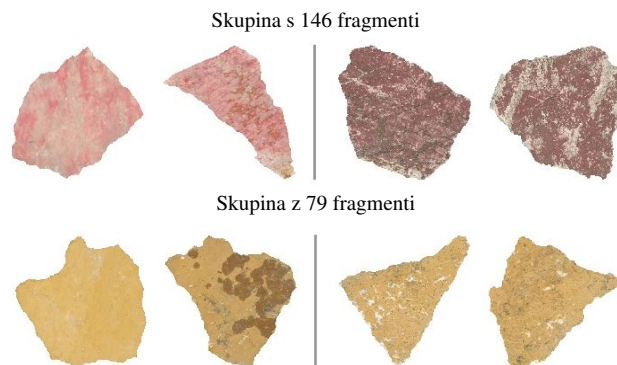
$$d(a, b) = \sqrt{1 - \frac{1}{\sqrt{a_{\text{avg}} b_{\text{avg}} N^2}} \sum_{i=1}^N \sqrt{a_i b_i}},$$

kjer sta a in b histograma, ki ju primerjamo, a_i in b_i i -ti polji histogramov, a_{avg} in b_{avg} povprečni vrednosti histogramov ter N število polj v vsakem histogramu.

Ker smo izbrali tri barvne modele, tri velikosti histogramov in šest metrik, lahko dve sliki fragmentov primerjamo na $3 \times 3 \times 6 = 54$ načinov. Zanima nas, kateri način primerjave je najustrežnejši, a tega ne moremo enostavno izmeriti, saj nimamo na voljo podatkov o tem, kateri fragmenti so si dejansko najbolj podobni. Zato smo uspešnost načinov primerjave ugotavljali tako, da smo vsako sliko fragmenta razdelili na štiri enako velike dele in uporabili naslednje točkovanje. Način primerjave fragmentov je prejel točke vsakič, ko je za izbrano četrtino slike fragmenta katero od preostalih treh četrtin uvrstil med prvih deset najbolj podobnih slik fragmentov. Za uvrstitev slike na k -to mesto, kjer je $k = 1, \dots, 10$, je prejel $11 - k$ točk. Postopek smo izvedli za vse štiri četrtine slike vseh fragmentov in točke vsakič sešteli preko uvrščenih ostalih četrtin. Način primerjave, ki je zbral največ točk, smo privzeli za najustrežnejšega.

4.2 Združevanje v skupine

Ko izberemo najustrežnejši način ugotavljanja podobnosti in z njim izračunamo razdalje med vsemi slikami fra-



Slika 2: Na grafičnem vmesniku za ogled skupin je vsaka skupina predstavljena s štirimi fragmenti: leva fragmenta imata največjo oddaljenost od ostalih fragmentov v skupini, desna fragmenta pa sta od ostalih fragmentov v skupini najmanj oddaljena

gmentov, želimo najbolj podobne slike fragmentov združiti v skupine. Za to uporabimo hierarhično združevanje v skupine (angl. hierarchical clustering). To deluje tako, da začne v stanju, kjer je vsak objekt v svoji skupini in je edini član te skupine. Nato ponavlja postopek, v katerem združi dve najbolj podobni skupini. To ponavlja, dokler ne dobi le ene skupine, v kateri so vsi objekti. Tako dobimo na koncu drevo, imenovano dendrogram, ki prikazuje, kako so se skupine postopoma združevale med seboj. Če ta postopek združevanja na neki točki prekinemo, dobimo določeno število skupin.

Vprašanje je, kdaj končati postopek združevanja. V našem primeru je to težko formalno definirati. Želimo predvsem, da so slike podobnih fragmentov v isti skupini, medtem ko je število skupin ali slik fragmentov v skupini manj pomembno. Ker ustreznost združevanja v skupine najlažje ocenimo na pogled, smo za določanje skupin razvili grafični uporabniški vmesnik, ki uporabniku prikaže vse skupine na neki stopnji združevanja in mu v primeru, da z videnim ni zadovoljen, omogoča, da stopnjo združevanja poveča oz. zmanjša, tako da dobi manj oz. več skupin (slika 2).

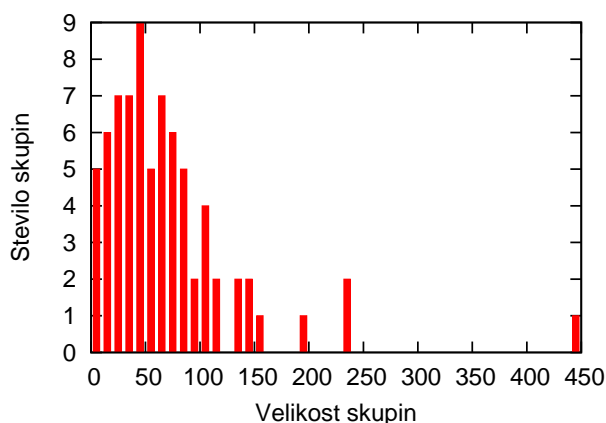
5 Rezultati

Opisani postopek iskanja podobnih fragmentov smo uporabili na projektu Turška mačka. Najprej smo 1000 naključno izbranih fragmentov² primerjali med seboj na vseh 54 načinov (z uporabo različnih barvnih modelov, velikosti histogramov in metrik). Uspešnost načinov primerjave prikazuje tabela 1, v kateri so zbrane točke normirane glede na $(10 + 9 + 8) \times 1000 \times 4 = 108.000$ točk, ki predstavljajo idealno uspešnost načina primerjave. Kot vidimo, se za najustrežnejšega izkaže barvni model HSV s histogramom velikosti 256 polj in Hellingerjevo razdaljo. Zaradi majhne razlike (le 4 %) med rezultatom na histogramih velikosti 256 in 128 polj, smo se zaradi večje učinkovitosti odločili

²Zaradi velike časovne zahtevnosti primerjav smo število uporabljenih fragmentov v tem koraku omejili na 1000.

Tabela 1: Rezultati za različne metrike, barvne modele in velikosti histogramov, razvrščeni po uspešnosti (v %)

Razdalja	HSV 256	HSV 128	LAB 256	HSV 64	LAB 128	LAB 64	RGB 128	RGB 256	RGB 64
Hellingerjeva	60,1	59,7	58,4	58,6	58,1	57,2	48,5	49,1	49,1
χ^2	56,8	56,5	55,8	55,6	55,5	54,8	46,7	47,2	47,7
mahnattanska	55,9	55,2	54,2	53,6	53,6	52,1	44,5	44,5	44,5
korelacija	53,0	51,9	51,5	49,9	50,8	48,9	42,9	42,9	42,7
evklidska	48,9	48,7	48,3	48,2	48,0	47,2	42,9	42,9	42,8
manh. na kum. hist.	43,0	43,0	42,4	43,0	42,4	42,4	36,8	36,8	36,7



Slika 3: Število skupin v odvisnosti od njihove velikosti (0–10, 10–20, ..., 440–450 fragmentov) v primeru združitve enobarvnih fragmentov v 73 skupin

v praksi uporabljati slednjega.

Tako definirano podobnost med fragmenti smo nato uporabili hierarhično združevanje v skupine (tokrat upoštevajoč vseh 5162 enobarvnih fragmentov projekta Turška mačka). V primeru, ko smo fragmente združili v 73 skupin, smo dobili precej raznolike skupine (histogram na sliki 3 prikazuje število skupin glede na njihovo velikost). Medtem ko polovica skupin vsebuje manj kot 50 fragmentov, so nekatere druge skupine zelo velike, največja ima skoraj 450 fragmentov. Dve tako dobljeni skupini sta predstavljeni na sliki 2.

6 Sklep

V prispevku smo predstavili problem iskanja podobnih enobarvnih fragmentov, s katerim se soočamo med pripravo podatkov za mobilno in spletno aplikacijo e-Pedius za množično sestavljanje fragmentov stenskih poslikav. Izmed 54 načinov ugotavljanja podobnosti je na fragmentih restavratorskega projekta Turška mačka najboljše rezultate dala Hellingerjeva razdalja na histogramu HSV z 256 polji. Ker pa so bili rezultati na histogramu HSV s 128 polji primerljivi, zaradi njegove polovične velikosti (in posledično večje učinkovitosti) v praksi raje uporabljamo ta način.

Postopka združevanja fragmentov v skupine nismo avtomatizirali v celoti, ampak smo končno odločitev prepustili strokovnjakom, ki lahko na podlagi prikazanih skupin

fragmentov interaktivno določijo končne skupine.

V prihodnosti bi lahko iskanje podobnih fragmentov dodatno izboljšali z uporabo informacij, pridobljenih od uporabnikov aplikacije e-Pedius. Medtem ko vnaprej ne vemo, kateri fragmenti so si v resnici najbolj podobni, lahko to razberemo iz skupnih uvrstitev v postavitve, ki jih posredujejo uporabniki. Podobno lahko zaznamo tudi morebitno neustrezno razvrščanje fragmentov po podobnosti. Če na primer uporabniki izbirajo fragmente globlje iz ponujene seznama podobnih, to pomeni, da niso zadovoljni s predlaganim razvrščanjem.

Bralce vabimo k sodelovanju pri množičnem sestavljanju fragmentov stenskih poslikav – aplikacija e-Pedius je na voljo na <http://e-pedius.si/>.

Zahvala

Opisano delo je bilo opravljeno v okviru projekta e-Pedius, izbranega na javnem razpisu za sofinanciranje projektov razvoja e-storitev in mobilnih aplikacij za javne in zasebne neprofitne organizacije 2012–2013. Projekt so sofinancirali Republika Slovenija, Ministrstvo za izobraževanje, znanost in šport, Evropska unija, Evropski sklad za regionalni razvoj, Institut “Jožef Stefan” in Zavod za varstvo kulturne dediščine Slovenije.

Literatura

- [1] e-Pedius, Podpora množičnemu sestavljanju fragmentov stenskih poslikav, 2013. URL <http://e-pedius.si/>. Pridobljeno 19. 7. 2013.
- [2] B. Filipič, M. Mlakar, E. Dovgan in T. Tušar. Razvoj sistema za računalniško podprto evidentiranje in sestavljanje fragmentov stenskih poslikav. *Zbornik 14. mednarodne multikonference Informacijska družba – IS 2011*, zvezek A, strani 45–48. Institut “Jožef Stefan”, Ljubljana, 2011.
- [3] Histograms, OpenCV API Reference, OpenCV 2.4.6.0 Documentation, 2011. URL <http://docs.opencv.org/modules/imgproc/doc/histograms.html>. Pridobljeno 19. 7. 2013.
- [4] Pedius, Sistem za evidentiranje in sestavljanje fragmentov stenskih poslikav, 2011. URL <http://dis.ijs.si/ci/pedius/index.html>. Pridobljeno 19. 7. 2013.