

ACM MARIBOR

1. Delavnica sekcije ACM Maribor 2015

Zbornik povzetkov



Urednika: Iztok Fister Jr. in Niko Lukač

Maribor, 14. Maj 2015

ACM MARIBOR

1. Delavnica sekcije ACM Maribor 2015

Maribor, 14.05.2015

Urednika: Iztok Fister Jr. in Niko Lukač

Obdelava in oblikovanje: Iztok Fister Jr. in Niko Lukač

Izdajatelj: ACM Maribor

Spletna izdaja: <http://acm.um.si/delavnica2015>

Leto izdaje: maj 2015

Organizacijski odbor:

Iztok Fister Jr.

Niko Lukač

Denis Kolednik

Kazalo

Jani Dugonik - Sistem za strojno prevajanje na podlagi besednih zvez	4
Uroš Mlakar - Diferencialna evolucija in njena uporaba za večpragovno segmentacijo slik	5
Iztok Fister Jr. - Algoritmi po vzorih iz narave in njihova uporaba v športu	6
Denis Horvat - Klasifikacija vegetacije v podatkih LiDAR	7
Niko Lukač - Vrednotenje fotovoltaičnega potenciala nad podatki daljinskega zaznavanja	8
Štefan Kohek, Damjan Strnad - Paralelna simulacija rasti dreves	9
Denis Kolednik - Zaznavanje sprememb z uporabo matematične morfologije	10

Sistem za strojno prevajanje na podlagi besednih zvez

Jani Dugonik
Univerza v Mariboru

Povzetek

Strokovnjaki na področju strojnega prevajanja že dolgo poskušajo opisati jezik. Naravni jezik pa je živ in se nenehno spreminja, pravila so kompleksna in ne upoštevajo kreativnosti, zato jezika z vsemi izjemami ne morejo ukleščiti v svoje stroge okvire. Sistemi statističnega strojnega prevajanja so osnovani na statističnih modelih, ki so naučeni na poravnanih dvojezičnih besedil. Namesto, da bi prevajalnik razčlenjeval stavke po slovničnih pravilih, ta išče splošne vzorce, ki se pojavljajo v jezikovni rabi. Besedilo je prevedeno glede na verjetnostno porazdelitev. Prevod je tisto besedilo, ki ima najvišjo verjetnost, le-ta se običajno računa po posameznih povedih. Statistične metode so se prvotno ukvarjale s prevajanjem posameznih besed, v zadnjih letih pa so napredovale na raven besednih zvez oz. besednih nizov. Besedni nizi niso stavki kot slovnične strukture, temveč nizi besed, ki jih v korpusu prepoznajo sistemi, ustvarjeni po statistični metodi. Ustvarili bomo zbirko orodij za procesiranje naravnih jezikov, s pomočjo katere bomo nato zgradili sistem za prevajanje. Zbirka orodij bo napisana v programskem jeziku C++ in bo vsebovala različna orodja za obdelavo besedila, jezikovne in prevajalne modele, binarizacija modelov za hitrejša nalaganja, dodajanje lastnih modelov, ter iskalni algoritem (dekodirnik). S temi orodji bo nato zgrajen sistem za prevajanje, kateri bo tudi dosegljiv na spletu.

Diferencialna evolucija in njena uporaba za večpragovno segmentacijo slik

Uroš Mlakar
Univerza v Mariboru

Povzetek

Danes živimo v svetu, kjer se vsakodnevno srečujemo z mnogimi signali, med katerimi so zagotovo najbolj pogoste digitalne fotografije. Ker je tega materiala ogromno, se mnogo ljudi ukvarja s proučevanjem in razvijanjem algoritmov, s katerimi bi enostavno in predvsem avtomatsko izluščili vsebino slike in jo na nedvoumen način predstavili računalniku za nadaljno obdelavo. Ena izmed bolj preprostih metod je zagotovo upragovljanje, ki je tudi veliko uporabljena za ločevanje zanimivih objektov od ozadja v sliki. Iskanje optimalnega pragu lahko smatramo kot optimizacijski proces z enim globalnim optimumom, zato se tudi veliko raziskovalcev ukvarja z razvojem algoritmov, ki bi poenostavili in tudi pohitrili ta proces. Za iskanje optimalnih pragov je bilo razvitih že več stohastičnih algoritmov, katere bomo v tem delu tudi podrobneje pregledali, predvsem pa se bomo osredotočili na metodo Otsu-ja, na kateri temelji tudi naš algoritem. Na področju upragovljanja so veliko zanimanja dobile hevtistične metode, saj so stohastični algoritmi velikokrat neučinkoviti oz. preprosto računsko prezahtevni. Predstavili bomo algoritem diferencialne evolucije in njegovo uporabo za upragovljanje slik. Uporabili bomo različne mutacijske strategije in opazovali njihov vpliv na kvaliteto rešitev, ter prav tako na hitrost algoritma. Naš algoritem bomo primerjali s 3 algoritmi iz literature (kukavičje iskanje, optimizacija z roji delcev ter samo-prilagodljiva diferencialna evolucija). Eksperimente smo izvedli na 11 testnih slikah iz standardne testne baze, ki je javno dostopna. Na koncu se bomo še dotaknili analize učinkovitosti hevtističnih algoritmov in težav, ki smo jih imeli s primerjavo algoritmov pri iskanju optimalnih pragov.

Algoritmi po vzorih iz narave in njihova uporaba v športu

Iztok Fister Jr.
Univerza v Mariboru

Povzetek

Algoritmi po vzorih iz narave poskušajo uporabiti naravne značilnosti bioloških organizmov (npr. obnašanje čebel, kukavic, mravelj) v računalništvu. Ti algoritmi so se pokazali kot zelo učinkoviti pri reševanju težkih problemov na različnih področjih, kot na primer v matematiki in ekonomiji. Nedavne raziskave so razkrile raznovrstne možnosti uporabe teh algoritmov tudi na področju športa, še posebej planiranja športnega treninga. V tej predstavitvi predstavimo univerzalno rešitev poimenovano umetni športni trener s sposobnostmi primerljivimi klasičnemu športnemu trenerju. Umetni športni trener je sestavljen iz algoritmov podatkovnega rudarjenja, ki poskušajo poiskati glavne navade treniranca in algoritmov računske inteligence, katerih naloga je generiranje dolgoročnega in kratkoročnega načrta treningov. Predstavitev bomo končali s pogledom na prihodnje izzive in razširitve umetnega športnega trenerja.

Klasifikacija vegetacije v podatkih LiDAR

Denis Horvat

Univerza v Mariboru

Povzetek

Predstavljamo novo metodo klasifikacije vegetacije v podatkih LiDAR. Metoda za svoje delovanje uporablja samo položaje točk ter njihovo izpeljano lokalno geometrijo kar pomeni, da dodatnih informacij točk (npr. število odbojev, intenziteta in ortofoto), ki niso vedno na voljo ne potrebujemo. Predlagan pristop vzpostavi topologijo med točkami z enakomerno delitvijo prostora v celice. Nad celicami nato izvedemo transformacijo lokalno prilegajočih ploskev kjer za vsako celico pridobimo ravnino katere izračun je odvisen od lokalne geometrije površja podatkov. Prileganje vsake ravnine ocenimo z napako prileganja, ki jo določajo ortogonalne razdalje točk do ravnine v okolici. Tako bodo ravninski predeli površja podatkov proizvedli majhne napake prileganja, pri predelih s statistično veliko višinsko razpršenostjo točk, med katerimi smatramo tudi vegetacijo, pa bo ta napaka občutno večja. Ker večjih napak prileganja ne proizvede samo vegetacija, te predele dodatno filtriramo z uporabo kontekstne analize, kjer kontekst predstavlja sosednost ravninskih ter neravninskih področij (področij z nizko in visoko napako prileganja). Napake, ki jih obravnavamo kategoriziramo na predele, ki se nahajajo med ravninami, predele, ki so pritrjeni k ravninam in predele, ki se nahajajo nad ravninami. Učinkovitost metode ocenimo z metriko popolnosti (angl. completeness), pravilnosti (angl. correctness) in kvalitete (angl. quality). Dosežena kvaliteta klasifikacije vegetacije za urbana področja sega od 84,2% do 89,7% za podeželska 94,7%, za gozdna področja pa 99,5.

Vrednotenje fotovoltaičnega potenciala nad podatki daljinskega zaznavanja

Niko Lukač
Univerza v Mariboru

Povzetek

Sončna energija je dandanes ena iz med najbolj pomembnih virov energij za doseganje trajnostnega razvoja mest. Kljub veliki količini prejete sončne energije, je globalni izkoristek nizek. Na učinkovitost izkoriščanja vplivajo različni faktorji, kot npr. senčenje iz okolice, lokalna mikroklima ter fotovoltaična tehnologija. V zadnjih letih je tehnologija daljinskega zaznavanja pripomogla, da zelo natančno zajamemo karakteristike površja, kar nam omogoča nova aplikacije oz. simulacije nad realnimi podatki. V tem prispevku predstavimo učinkovito metodo za izračun fotovoltaičnega potenciala nad stavbami in terenom zajetih iz podatkov daljinskega zaznavanja. Osredotočimo se na zračno snemanje na danem področju, pri čem uporabimo stereoskopske optične posnetke ter podatke LiDAR (Light Detection And Ranging). V tem delu predpostavimo, da so dani podatki že georeferencirani. Stereo posnetki nam omogočajo pokritost širšega geografskega področja, iz katerih generiramo digitalni model terena (DMT) nižje ločljivosti. Laserski posnetek LiDAR pa nam dajejo visokoločljivostni topološko nestrukturiran oblak točk manjšega področja nad katerim izvedemo izračun fotovoltaičnega potenciala. Oblak točk umestimo v 2.5D mrežo za vzpostavitev topologije, nato točke klasificiramo v razrede stavb, terena in vegetacije. Obdelani stereo in LiDAR podatki nam nato predstavljajo vhod v izračun, saj nam omogočajo zelo natančno topografsko sliko realnih površij. Senčenje igra pomembno vlogo pri modeliranju sončnega obsevanja, saj znatno zniža direktno obsevanje. Simulacijo senčenja izvedemo na 2.5D mreži, kjer za izračun pozicije Sonca za dano geografsko lokacijo uporabimo metodo SPA (Solar Positioning Algorithm). V izračunu upoštevamo tudi lokalno mikroklimo, s pomočjo lokalnih meteoroloških meritev globalnega in difuznega obsevanja ter satelitskih meritev kratkovalovnega sevanja širšega geografskega področja. Pri vsakem koraku časovne integracije upoštevanjem nelinearnih karakteristik fotovoltaičnih modulov. Izračun smo izvedli nad tremi geografskimi področji z različnim lokalnim podnebjem in topografijo površja.

Paralelna simulacija rasti dreves

Štefan Kohek in Damjan Strnad

Univerza v Mariboru

Povzetek

Splošna razširjenost več-jedrnih centralnih in grafičnih procesnih enot, ki omogočajo paralelno obdelavo podatkov, zahteva ustrezno prilagoditev algoritmov iz eno-nitnega izvajalnega modela. Simulacija rasti dreves je eden izmed problemov, ki jih je težje paralelizirati predvsem zaradi kompleksnejše medsebojne odvisnosti podatkov. Simulacija rasti dreves namreč sestoji iz več zaporednih korakov, kjer rezultat trenutnega koraka v mnogo primerih predstavlja vhod v naslednje korake. Dodaten problem predstavlja rast podatkovnih struktur skupaj s samo rastjo dreves. Uporabljen pristop simulacije rasti upošteva medsebojno tekmovanje delov rastlin za vire iz okolja, kot je v našem primeru svetloba. Računsko najintenzivnejša koraka, katerima je smiselno nameniti največ pozornosti, sta tako redistribucija virov znotraj drevesa kot njegova interakcija z okoljem. V tem delu predstavimo aplikacijo, ki omogoča tvorbo dreves s pomočjo paralelnih procesnih enot. Aplikacija omogoča interaktiven prikaz dreves in spreminjanje njihovih parametrov. Za večjo kontrolo rasti dreves predstavljamo tudi razširitve in poenostavitve nad obstoječim modelom rasti, kjer razširitve ne predstavljajo bistvene upočasnitve pri paralelni implementaciji, poenostavitve pa omogoča občutno pospešitev redistribucije virov znotraj dreves. Za samo implementacijo je uporabljen OpenCL, ki omogoča poganjanje algoritmov tako na grafičnih procesnih enotah kot tudi na ostalih procesnih enotah. Vsi koraki simulacije rasti so implementirani v OpenCL, sama kontrola zagonov posameznih korakov pa ostaja na centralni procesni enoti. V nadaljevanju pokažemo, da paralelizacija prinese občutne pospešitve, ki že omogočajo interaktivno oblikovanje oziroma spremembe parametrov dreves. Predstavljene razširitve in poenostavitve modela rasti pa omogočajo pridobitev raznolikih oblik dreves. Skupen rezultat je tvorba raznolikih gozdov, ki sestojijo iz več kot 1000 preprostih dreves, v manj kot sekundi.

Zaznavanje sprememb z uporabo matematične morfologije

Denis Kolednik
Univerza v Mariboru

Povzetek

Splošna razširjenost več-jedrnih centralnih in grafičnih procesnih enot, ki omogočajo paralelno obdelavo podatkov, zahteva ustrezno prilagoditev algoritmov iz eno-nitnega izvajalnega modela. Simulacija rasti dreves je eden izmed problemov, ki jih je težje paralelizirati predvsem zaradi kompleksnejše medsebojne odvisnosti podatkov. Simulacija rasti dreves namreč sestoji iz več zaporednih korakov, kjer rezultat trenutnega koraka v mnogo primerih predstavlja vhod v naslednje korake. Dodaten problem predstavlja rast podatkovnih struktur skupaj s samo rastjo dreves. Uporabljen pristop simulacije rasti upošteva medsebojno tekmovanje delov rastlin za vire iz okolja, kot je v našem primeru svetloba. Računsko najintenzivnejša koraka, katerima je smiselno nameniti največ pozornosti, sta tako redistribucija virov znotraj drevesa kot njegova interakcija z okoljem. V tem delu predstavimo aplikacijo, ki omogoča tvorbo dreves s pomočjo paralelnih procesnih enot. Aplikacija omogoča interaktiven prikaz dreves in spreminjanje njihovih parametrov. Za večjo kontrolo rasti dreves predstavljamo tudi razširitve in poenostavitve nad obstoječim modelom rasti, kjer razširitve ne predstavljajo bistvene upočasnitve pri paralelni implementaciji, poenostavitve pa omogoča občutno pospešitev redistribucije virov znotraj dreves. Za samo implementacijo je uporabljen OpenCL, ki omogoča poganjanje algoritmov tako na grafičnih procesnih enotah kot tudi na ostalih procesnih enotah. Vsi koraki simulacije rasti so implementirani v OpenCL, sama kontrola zagonov posameznih korakov pa ostaja na centralni procesni enoti. V nadaljevanju pokažemo, da paralelizacija prinese občutne pospešitve, ki že omogočajo interaktivno oblikovanje oziroma spremembe parametrov dreves. Predstavljene razširitve in poenostavitve modela rasti pa omogočajo pridobitev raznolikih oblik dreves. Skupen rezultat je tvorba raznolikih gozdov, ki sestojijo iz več kot 1000 preprostih dreves, v manj kot sekundi.