SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
MATEMATIČKI ODSJEK

Upotpunjenje slika pomoću difuzijskih preslikavanja i spektralne relaksacije

Ante Ćubela i Mirna Lovrić

Zagreb
22. svibanj 2023.

Sadržaj

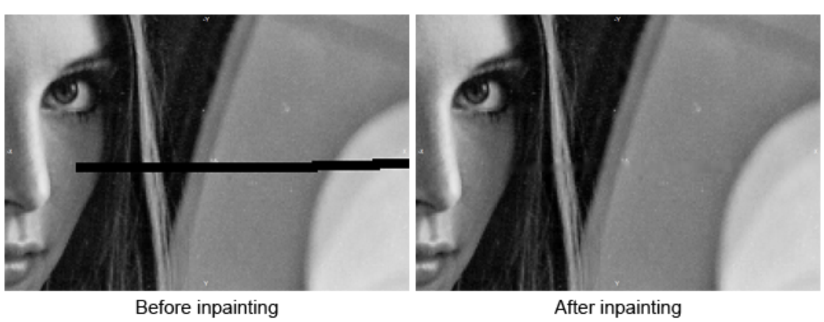
Uvod

Image inpainting je tehnika koja se koristi kako bismo rekonstruirali oštećenu sliku ili dio videozapisa, odnosno popunili odgovarajuće dijelove slike koji nedostaju na vizualno prihvatljiv način. Prvotno se ova tehnika koristila za rekonstrukciju oštećenih umjetničkih slika, no u današnjici nalazi veliku primjenu u uređivanju digitalnih slika. Na primjer, s ovom tehnikom možemo maknuti određene predmete sa slike na način da slika i dalje izgleda vizualno prihvatljivo.

Konkretno, ako imamo oštećenu sliku I, gdje je H nepoznati dio slike, želimo popuniti područje H tako da dobijemo vizualno prihvatljivu sliku $\hat{I}$. Postojeće metode koje rješavaju ovaj problem možemo podijeliti na dvije vrste. Prva skupina metoda bazira se na varijacijskom računu, odnosno problem *inpainting*-a se formulira kao minimizacija varijacijskog funkcionala koji prima određena prostorna ograničenja vezana uz glatkoću. Druga skupina metoda područje H popunjava koristeći poznati dio slike ili neku drugu sličnu bazu slika. Takve metode pokazale su se uspješnima, dok prva skupina metoda daje globalno optimalnije rješenje.

Definiramo referentni skup $\overline{H}$ kao $\overline{H} = I \setminus H$, odnosno to je poznati dio slike. U radu na koji smo se referirali, svakom pikselu $\mathbf{x}_{ij}$ slike I koji se nalazi u referentnom skupu, pridružujemo *patch* $\mathbf{p}_{ij} \in \mathbf{R}^D$, gdje je D uglavnom 5×5 ili 7×7. Na ovaj način dobili smo visoko dimenzionalni referentni skup $\{\mathbf{p}_{ij}\}$. Kako bismo ga lakše analizirali, koristimo difuzijska preslikavanja za smanjenje dimenzije. Radi toga, pretpostavljamo da su nepoznati *patchevi* i *patchevi* iz referentnog skupa na istoj nižedimenzionalnoj mnogostrukosti.

Ideja ovoga rada je piksele, odnosno patcheve iz nepoznatog dijela slike H preslikati u nižedimenzionalni prostor koristeći tzv. difuzijsko preslikavanje Ψ (koje prvotno nije definirano na H, ali koristeći interpolacijske tehnike definiramo preslikavanje s H) i zatim u takvom nižedimenzionalnom prostoru napraviti inpainting, pa dobivene vrijednosti *vratiti* u početnu domenu.



Slika 1: Primjer upotpunjene slike

1 Difuzijska preslikavanja

Pretpostavimo da imamo skup podataka $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\},\ \mathbf{x}_i \in \mathbf{R}^D,\ i = 1, \dots, n$. Želimo naći skup $\mathbf{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_n\},\ \mathbf{y}_i \in \mathbf{R}^d,\ i = 1, \dots, n$, tako da y_i *dobro opisuje* $x_i, \forall\ i$. Ideja difuzijskog preslikavanja je podatke iz skupa $\mathbf{X}$ organizirati u usmjereni težinski graf $G = (V, E)$, gdje je skup vrhova $V = \mathbf{X}$, a brid između dva proizvoljna vrha $\mathbf{x}_i$ i $\mathbf{x}_j$ ima težinu w_{ij}, koja predstavlja bliskost ta dva vrha. Često se w_{ij} definira pomoću RBF jezgre:

$$w_{ij} = exp(-\|x_i - x_j\|^2/\sigma^2) = exp(-d_{ij}^2/\sigma^2), \tag{1}$$

gdje je $\sigma > 0$ parametar skaliranja. Želimo definirati slučajnu šetnju na $\mathbf{X}$. Ako definiramo matricu

$$\mathbf{A} = \mathbf{D}^{-1}\mathbf{W}, \tag{2}$$

gdje je $\mathbf{D}$ dijagonalna matrica definirana s $\mathbf{D}_{ii} = \sum_j w_{ij}$, vidimo da je $\mathbf{A}$ retčano stohastička, simetrična i da je pozitivna, tj. $a_{ij} > 0, \forall\ i, j$. Nadalje, na izraz a_{ij} možemo gledati kao na vjerojatnost prelaska iz vrha $\mathbf{x}_i$ u vrh $\mathbf{x}_j$ u jednom koraku, odnosno $\mathbf{A}$ predstavlja matricu prijelaza Markovljevog lanca $\mathbf{X}$. Spektralna dekompozicija matrice A daje nam

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}, \quad \mathbf{V} = [\psi_1 \dots \psi_n], \quad \mathbf{V}^{-1} = \begin{bmatrix} \phi_1^T \\ \vdots \\ \phi_n^T \end{bmatrix} \tag{3}$$

Također, lako se pokaže da je $\phi_0 = \frac{1}{\sum_i d_{ii}} d^T$ stacionarna distribucija matrice $\mathbf{A}$. Definiramo familiju difuzijskih udaljenosti $\mathbf{D}_t(\cdot, \cdot)$ sa

$$\mathbf{D}_t(\mathbf{x}_i, \mathbf{x}_j)^2 = \sum_{\mathbf{x} \in \mathbf{X}} \frac{(\mathbb{P}_t(\mathbf{x}_i, \mathbf{x}) - \mathbb{P}_t(\mathbf{x}_j, \mathbf{x}))^2}{\phi_0(x)}, \tag{4}$$

gdje je $\mathbb{P}_t(\mathbf{x}_j, \mathbf{x})$ vjerojatnost da se iz vrha $\mathbf{x}_j$ dođe u vrh $\mathbf{x}$ nakon t vremenskih koraka. Možemo razmišljati o ovoj udaljenosti kao Raspisivanjem dođemo do

$$\mathbf{D}_t(\mathbf{x}_i, \mathbf{x}_j)^2 = \sum_{l \geq 1} \lambda_l^{2t} (\psi_l(\mathbf{x}_i) - \psi_l(\mathbf{x}_j))^2, \tag{5}$$

dakle vidimo da ako definiramo difuzijsko preslikavanje

$$\mathbf{\Psi}_t(\mathbf{x}) = (\lambda_1^t \psi_1(\mathbf{x}), \dots, \lambda_{d(t)}^t \psi_{d(t)}(\mathbf{x})), \tag{6}$$

gdje je $d(t)$ broj svojstvenih vrijednosti čija veličina nije zanemariva, difuzijska udaljenost postaje

$$\mathbf{D}_t(\mathbf{x}_i, \mathbf{x}_j)^2 = \|\mathbf{\Psi}_t(\mathbf{x}_i) - \mathbf{\Psi}_t(\mathbf{x}_j)\|^2. \tag{7}$$

2 Algoritam

Mi smo prvotno pristupili problemu *inpainting*-a na sljedeći način... Prvo smo napravili dvije maske, za poznati dio slike $\overline{H}$ i nepoznati dio H. Koristili smo slike dimenzija 100×100, te izrezali 10×30 piksela za dio slike koji treba popuniti. Svaki piksel iz poznatog dijela slike $\overline{H}$ je izvorno reprezentiran kao vektor iz $\mathbb{R}^3$ gdje su sačuvane RGB vrijednosti piksela.

Neka je n broj piksela u $\overline{H}$. Koristeći RBF (*radial basis function*) jezgru smo napravili matricu bliskosti $\mathbf{W}$, te pomoću nje dobili retčano stohastičku matricu $\mathbf{A} \in \mathbb{R}^{n \times n}$. Napravili smo spektranu dekompoziciju matrice $\mathbf{A}$, kako je spomenuto ranije, te iz spektralne dekompozicije smo dobili matricu $\mathbf{\Psi}$. Zbog opadanja svojstvenih vrijednosti u spektralnoj dekompoziciji, za dovoljno dobru aproksimaciju difuzijskih udaljenosti nam je dovoljan samo mali broj sumanada u izrazu (5). To jest, piksele smo prvo reprezentirali sa svojstvenim vektorima dobivene matrice, a zatim samo sa vodećim vrijednostima tih svojstvenih vektora, promatravši samo prvih k redaka matrice $\mathbf{\Psi}$. "Cutoff" je bio taj što smo uzeli u obzir broj svojstvenih vrijednosti većih od 0.1. Novu matricu smo nazvali $\mathbf{\Psi}_k$.

Kada smo piksele iz originalne domene naše slike prikazali kao stupce matrice $\mathbf{\Psi}$ u višoj dimenziji, cilj nam je bio popuniti tu sliku novom difuzijskom prostoru, i to na sljedeći način; Kako je područje H koje treba upotpuniti pravokutnik, odlučili smo se upotpunjavati vrijednosti od H od vrha prema dolje idući po retcima slijeva na desno, umjesto da idemo od rubova prema unutrašnjosti. Za neki piksel $\mathbf{P} \in H$, smo uzeli u obzir sve njegove neposredno susjedne piksele (horizontalno, vertikalno i dijagonalno), reprezentirane kao stupci matrice $\mathbf{\Psi}_k$. Zatim smo uzeli prosjek tih "susjednih" stupaca od $\mathbf{\Psi}$ (četiri do šest, ovisno o mjestu), te koristili taj prosjek kao naš *inpainting.* Koristili smo i drugu varijaciju popunjavanja H u difuzijskoj domeni, time što smo umjesto od vrha prema dnu, područje H popunjavali od rubova prema unutra, koristeći opet prosjeke vrijednosti susjednih piksela u difuzijskoj domeni.

Zatim smo za nove *inpaint*-ane piksele, izračunali udaljenosti u difuzijskoj domeni od svih drugih piksela, te uzeli nekoliko najbližih susjeda ($k-$ *Nearest Neighbours* algoritam). Za te najbliže susjede smo odredili indekse odgovarajućih piksela u originalnoj domeni slike, te time dobili neku vrstu "najbližih susjeda" za nepopunjene indekse u domeni slike. Zatim da popunimo vrijednosti tih piksela iz H, jednostavno uzmemo prosjeke RGB vrijednosti piksela koji su sada označeni kao najbliži susjedi.

Cijeli algoritam je sažet ispod:

Algorithm 1 Upotpunjenje slika

Ulazni podaci: Slika koju je potrebno upotpuniti, maske za poznati i za nepoznati dio slike
Izlazni podaci: Upotpunjena slika

1: Od poznatog dijela slike ćemo napraviti matricu bliskosti $\mathbf{W}$ koristeći RBF jezgru:

$$w_{i,j} = exp(-||x_i - x_j||^2/\sigma^2),$$

gdje je $x_i \in \mathbb{R}^3$ RGB vektor piksela na poziciji i.

2: $\mathbf{A} = \mathbf{D}^{-1}\mathbf{W}$, gdje je $\mathbf{D}_{ii} = \sum_j w_{i,j}$, te $\mathbf{D}$ dijagonalna matrica.

3: Napraviti spektralnu dekompoziciju od $\mathbf{A}$ ($\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$).

4: Napravimo matricu $\mathbf{\Psi} = \mathbf{\Lambda}\mathbf{V}^T$, the uzmemo njenih prvih k redaka: $\mathbf{\Psi}_k = \mathbf{\Psi}(1 : k, ;)$.

5: Koristeći stupce od $\mathbf{\Psi}_k$ kao reprezentaciju piksela u difuzijskoj domeni, napravimo upotpunjenje od H u difuzijskoj domeni koristeći prosjeke "susjednih" stupaca od $\mathbf{\Psi}_k$.

6: Za svaki $y_i \in H$ dobiven u difuzijskoj domeni, nađemo nekoliko udaljenošću najbližih stupaca od $\mathbf{\Psi}_k$,te pripadajuće piksele u domeni slike proglasimo najbližim susjedima od piksela kojemu odgovara y_i u difuzijskoj domeni ($k -$ *nearest neighbours*).

7: Svakomu pikselu $x_i \in H$ dodijelimo prosjeke RGB vrijednosti od njegovih najbližih susjeda dodijeljenih u koraku (6) algoritma, te smatramo to upotpunjenjem u originalnoj domeni.

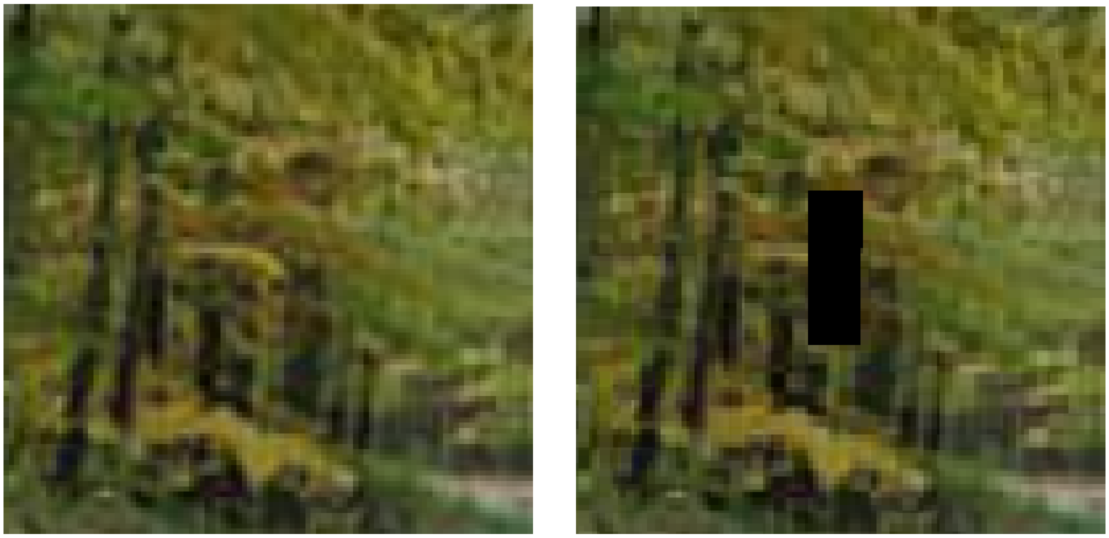
---

3 Rezultati

Isprobali smo naš algoritam na dvije slike dimenzija 100×100 piksela. Nismo mogli uzimati puno veće dimenzije zbog memorijskih ograničenja. Na slikama ispod su prikazani rezultati našeg algoritma, gdje su korišteni parametri $\sigma = 0.2$ (u jezgrenoj funkciji), $k = 10$ (u $k -$ *nearest neighbours*), te su prikazani rezultati za oba različita načina inpaintinga.

Slika 2: Originalna slika kamenčića na plaži i slika s maskom od $\overline{H}$

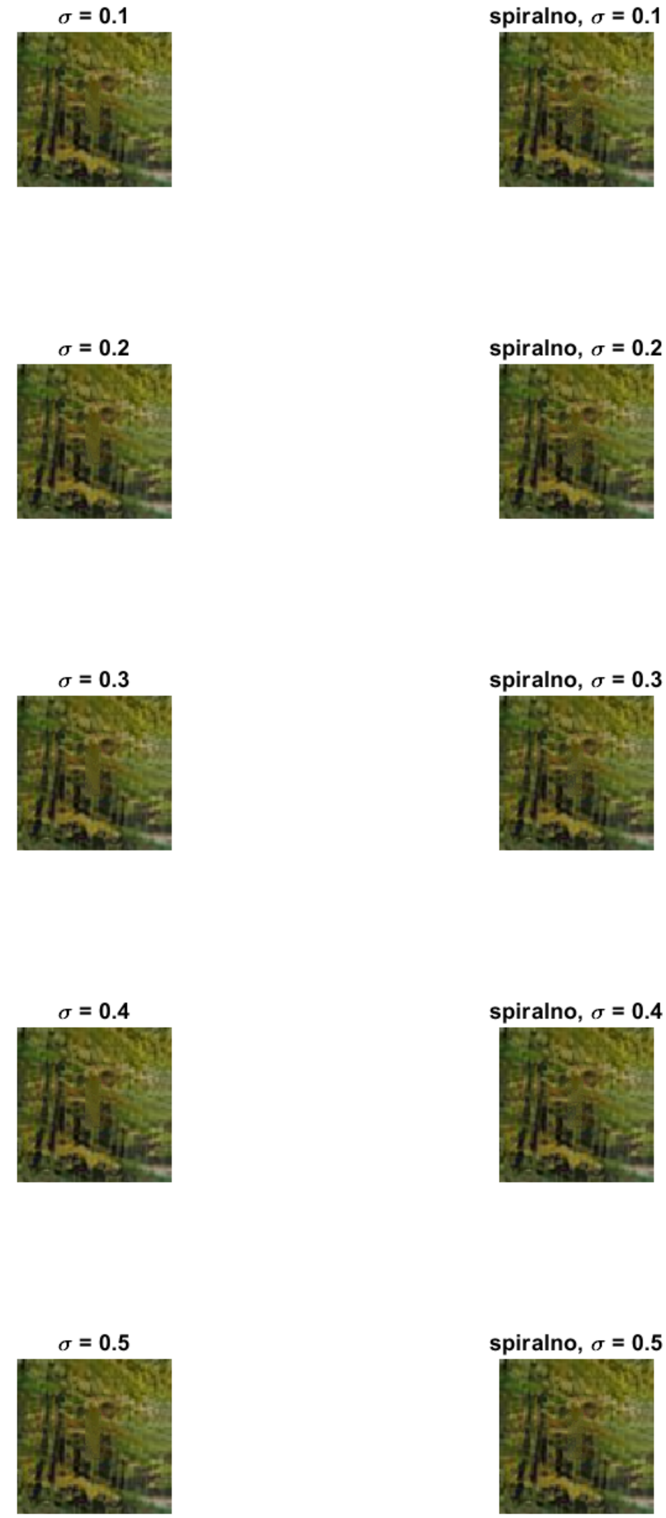
Slika 3: Popunjenje retčanim redoslijedom (lijevo) i spiralnim redoslijedom (desno)



Slika 4: Originalna slika šumice i slika s maskom od $\overline{H}$

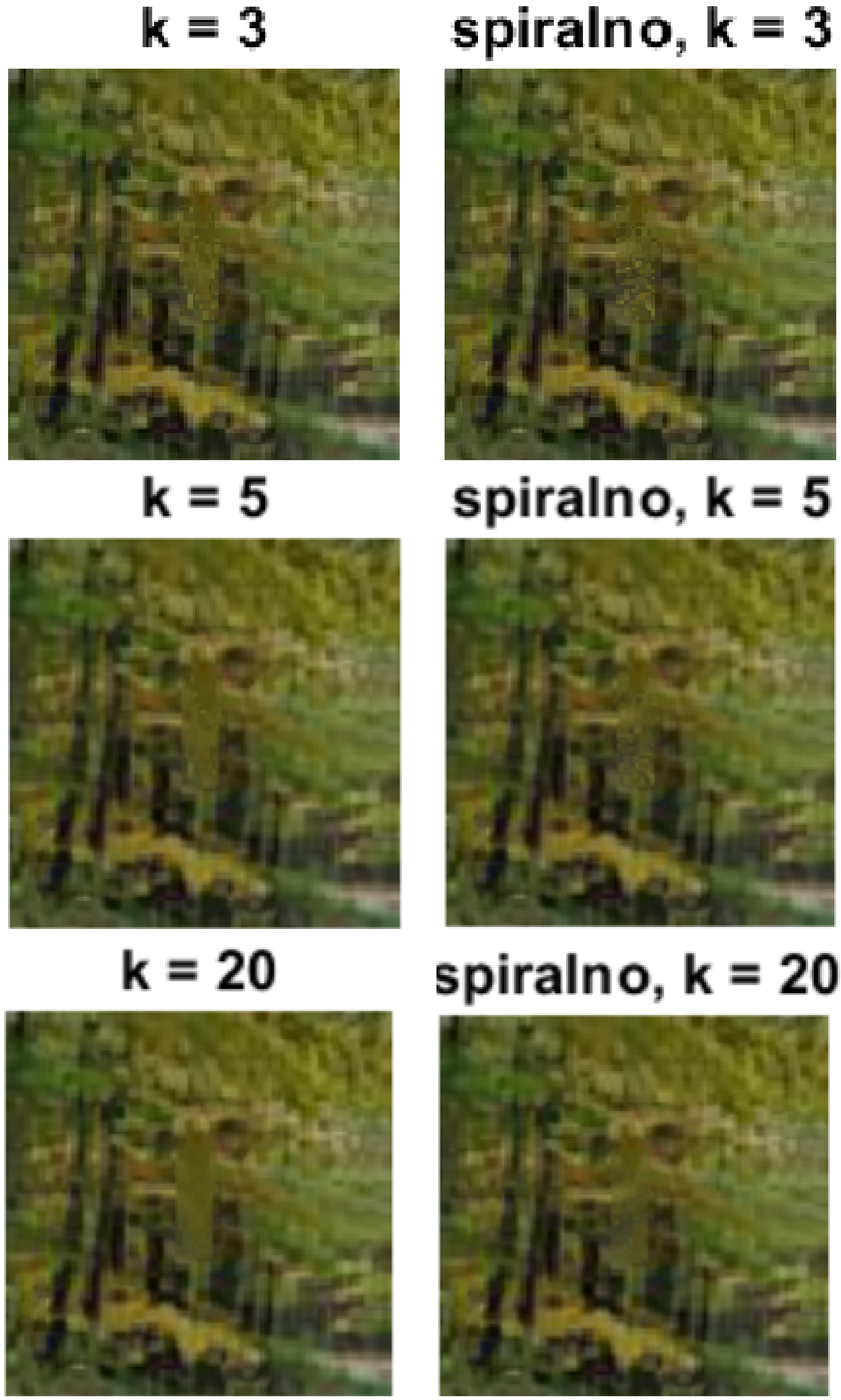
Slika 5: Popunjenje retčanim redoslijedom (lijevo) i spiralnim redoslijedom (desno)

Također smo isprobali kako parametar σ utječe na upotpunjenje slike. Upotpunjenja za različite vrijednosti parametra σ su pokazana na slikama ispod.



Slika 6: Slijeva je napravljeno upotpunjenje po recima, a nadesno upotpunjenje spiralno, za različite vrijednosti parametra σ

Isprobali smo i različite vrijednosti parametra k, korištenog kod uzimanja k najbližih susjeda. Neki rezultati su pokazani na slici ispod.



Slika 7: Upotpunjenje po recima (lijevo) i upotpunjenje spiralno (desno) različite vrijednosti k u kNN metodi

Literatura

[1] S. Gephstein, Y. Keller: *Image Completion by Diffusion Maps and Spectral Relaxation*, Faculty of Engineering, Bar Ilan University, Israel 2013.

[2] Z. Drmač: *Napredne Linearne i Nelinearne Numeričke Metode u Analizi Podataka*. PMF-MO Zagreb, 2023.