

Raport stiintific

privind implementarea proiectului in perioada noiembrie 2013 – noiembrie 2014

Proiect PN-II-ID-PCE-2011-3-0400, "Reprezentari rare in prelucrarea semnalelor"

Rezultate obtinute si publicatii

Pentru lectura rapida a acestui raport, voi descrie la inceput pe scurt rezultatele obtinute si apoi le voi prezenta cu unele detalii tehnice.

In anul 2014 am reusit publicarea a trei articole in conferinte ISI, doua la ICASSP (cea mai importanta conferinta in prelucrarea semnalelor) si unul la EUSIPCO (cea buna conferinta europeana in prelucrarea semnalelor). Un manuscris este aproape gata pentru trimiterea la o revista, un altul in pregatire pentru o conferinta.

Toate rezultatele obtinute sunt de natura algoritmica. Articolele se sprijina pe programe scrise in Matlab sau C. Unele din aceste programe sunt disponibile pe site-ul pe care se afla si pagina proiectului. Pe celelalte le distribuim la cerere prin email.

Directii de cercetare

In planul de realizare a proiectului pe 2014 au fost prevazute doua obiective:

- O1.** Proiectarea dictionarelor pentru reprezentari rare si clasificare
- O2.** Filtre si polinoame rare. Utilizare in aproximari de functii pozitive

Activitatea a avut loc cu precadere in cadrul **O1**.

Componenta echipei

Drd. Paul Irofti a fost angajat de la inceputul anului, el colaborand pe teme legate de O1 inca de anul trecut. Pe parcursul lui 2014, dr. Cristian Rusu a reinceput colaborarea, dupa un stagiu post-doc in strainatate. In schimb, colaborarea cu dr. Bogdan Sicleru a incetat, din motive de sanatate. Membrii echipei vor fi numiti prin initialele lor.

Descriere tehnica a rezultatelor

Voi detalia in continuare avansurile si rezultatele efective obtinute.

O1. La aceasta tema au lucrat PI, CR si BD. Am abordat doua directii de cercetare.

1. Problema de optimizare asociata proiectarii dictionarelor pentru reprezentari rare este dificila, iar algoritmi pentru rezolvarea ei sunt iterativi, vezi de exemplu K-SVD [AEBO6,RBE10]. Acesti algoritmi sunt

de obicei initializati cu un dictionar aleator sau cu o submultime a vectorilor de antrenare. Pe de alta parte, algoritmi de antrenare au complexitate mare, timpii de executie fiind de ordinul minutelor pentru dimensiuni curente ale dictionarului si crescand semnificativ odata cu acestea. Aceste constatari ne-au facut sa ne indreptam atentia catre initializare, pentru care merita alocat mai mult efort, cu conditia ca rezultatele antrenarii sa fie mai bune sau antrenarea sa decurga mai rapid. De aceea, am propus un algoritm rapid de initializare pentru K-SVD. Procedura porneste de la un dictionar incoerent obtinut cu [Rusu13] (el poate fi inlocuit, in cazul cel mai simplu, cu o matrice ortogonala generata aleator) si elimina o parte din atomii lui. In acest fel se obtine o multime de atomi cu proprietati bune de coerenta (adica foarte bine separati). Peste care se adauga atomi pornind de la vectorii de antrenare cel mai prost reprezentati. Complexitatea mica este obtinuta din faptul ca in pasul de eliminare, antrenarea se face doar prin rotatii (inmultiri cu matrice ortogonale), deci atomii initiali nu sunt antrenati individual, iar in pasul de adaugare nu se reantreneaza intreg dictionarul, ci doar atomii nou adaugati. Rezultatele, publicate in [2], sunt cele scontate, adica initializarea propusa scurteaza timpul total de antrenare pentru erori similare de reprezentare cu K-SVD, sau, la timpi de executie egali (reglati din numarul de iteratii K-SVD), produce eroare mai mica.

2. Deoarece optimizarea dictionarelor pentru dictionare rare are complexitate mare, o idee naturala este de a incerca obtinerea unor algoritmi paraleli. PI s-a ocupat de implementarea pe GPU, in OpenCL, a unor versiuni paralele ale algoritmilor consacrați. Dupa depasirea unor dificultati tehnice legate de implementarea OpenCL si BLAS pe tipul de placa AMD de care dispunem (OpenCL e un standard relativ nou si implementarile unor functii de baza nu sunt 100% stabile, dupa cum PI a constatat direct), primele rezultate au fost obtinute prin modificarea algoritmului K-SVD. In afara paralelizarii unor blocuri ale algoritmului, am testat si varianta in care mai multi atomi sunt actualizati simultan, spre deosebire de functionarea pur secventiala presupusa optima pentru K-SVD. Surprinzator, aceasta asigura nu doar un timp de executie mai scurt, ci si o eroare de reprezentare finala mai mica, cu pretul minor al unei convergente neuniforme. Noul algoritm si prezentarea rezultatelor obtinute pe placa GPU de care dispunem au produs lucrarea [3].

Un alt algoritm de antrenare pe care PI l-a paralelizat este dedicat dictionarelor formate din matrice ortogonale, in care reprezentarea unui vector de antrenare se face intr-o singura astfel de matrice, continuand munca din [Rusu13]. Aceasta structura de dictionar permite o reprezentare extrem de rapida, iar paralelizarea aduce un salt suplimentar de viteza, obtinandu-se timpi de executie mult mai mici decat ai unor algoritmi standard ca Orthogonal Matching Pursuit. De asemenea, timpul de antrenare este net inferior celui al K-SVD (inclusiv in varianta paralela). Am investigat de asemenea situatiile in care eroarea de reprezentare este comparabila cu cea a K-SVD. O lucrare asupra acestui subiect este aproape finalizata.

02. La aceasta tema au lucrat BS si BD. Subiectul principal a fost aproximarea sumelor de exponentiale (SdE) reale pozitive cu polinoame. SdE apar in mod natural in modelarea densitatilor de probabilitate, cu aplicatii in statistica [Duf07], finante, asigurari, dar si in raspunsurile unor clase de sisteme. Am utilizat doua tehnici pentru rezolvarea problemelor de optimizare cu SdE printre variabile. In ambele cazuri, o aproximatie initiala a exponentilor se obtine adaugand functiei obiectiv un termen care forteaza raritate SdE. Apoi se cauta valori mai bune ale exponentilor cu simpli pasi de gradient. De asemenea, se profita de faptul ca pentru exponenti ficși, problema de optimizare este convexa, deci coeficientii SdE pot fi optimizati exact. Am aplicat metodele obtinute pentru modelare densitatilor de probabilitate empirice, dar si pentru proiectare filtrelor IIR cu raspuns pozitiv.

1. Daca exponentialele au exponenti cu valori parte dintr-o progresie aritmetica, atunci SdE poate fi vazuta ca un polinom rar, pentru care restrictia de pozitivitate poate fi exprimata printr-o inegalitate liniar matriceala (LMI). Pasii de gradient produc in general exponenti care nu apartin unei progresii aritmetice; totusi, aproximand exponentii cu o progresie aritmetica cu pas suficient de mare astfel incat gradul polinomului asociat sa nu fie prea mare, si astfel complexitatea problemei de optimizare sa ramana acceptabila, am obtinut o procedura de optimizare care da rezultate foarte bune pentru SdE cu un numar rezonabil de termeni (<10). Rezultatele au fost publicate in [1].

2. Pentru cazul general, cand exponentii au valori arbitrare, am utilizat doua aproximari polinomiale unilaterale ale exponentialei, una superioara, alta inferioara. Aproximarile se fac pe intervale a caror lungime poate fi calculata automat, pentru o precizie dorita de aproximare; am constatat ca este suficient un numar mic de intervale pentru a obtine precizii practic bune. In acest fel, probleme de optimizare cu SdE pot fi approximate, cu precizie arbitrara, prin probleme cu polinoame, a caror rezolvare este posibila cu rutine standard de genul celor din biblioteca POS3POLY [SiDu13]. Un articol a fost finalizat pe aceasta tema.

Greutati intampinate si alte informatii

Principala dificultate continua sa fie bugetul redus. In plus, din pacate, Bogdan Sicleru s-a imbolnavit grav si, dupa o perioada in care a lucrat mai putin, a renuntat complet la activitatea de cercetare (ca si la cea didactica). Oricum n-aveam bani sa-l platesc pentru mai mult timp, dar abia acum se vede ce se intampla cand timpul de lucru este efectiv cel pontat. De aceea, continuarea lucrului in directia O2 este pusa sub semnul intrebării. Cristian Rusu a revenit in tara dupa un stagiu post-doc, dar a plecat in toamna la altul. I-am pastrat formal in echipa proiectului pe BS si CR, cu al doilea colaborarea continuand efectiv.

Pagina web dedicata proiectului este <http://www.schur.pub.ro/Idei2011.htm>.

Articole publicate la conferinte (ISI proceedings)

[1] B.Dumitrescu, B.C.Sicleru - Optimization with sums of exponentials and applications, *Int. Conf. Acoustics Speech Signal Proc. (ICASSP)*, Florence, Italy, pp.370-374, May 2014.

[2] C.Rusu, B.Dumitrescu - An initialization strategy for the dictionary learning problem, *Int. Conf. Acoustics Speech Signal Proc. (ICASSP)*, Florence, Italy, pp.6781-6785, May 2014.

[3] P.Irofti, B.Dumitrescu - GPU Parallel Implementation of the Approximate K-SVD Algorithm Using OpenCL, *European Signal Proc. Conf. (EUSIPCO)*, Lisbon, Portugal, Sep. 2014.

Bibliografie

[AEB06] M. Aharon, M. Elad, A. Bruckstein, K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation, *IEEE Trans. Signal Proc.*, vol.54, no.11, pp. 4311-4322, Nov. 2006.

[Duf07] D. Dufresne, Fitting Combinations of Exponentials to Probability Distributions, *Appl. Stochastic Models Bus. Ind.*, vol. 23, pp. 23–48, 2007.

[RBE10] R. Rubinstein, A. M. Bruckstein, M. Elad, Dictionaries for sparse representation modeling, *Proc. IEEE*, vol.98, no.6, pp.1045–1057, 2010.

[Rusu13] C. Rusu, Design of incoherent frames via convex optimization, *IEEE Signal Proc. Letters*, vol.20, no.7, pp. 673-676, July 2013.

[RuDu13] C.Rusu, B.Dumitrescu, Block Orthonormal Overcomplete Dictionary Learning, *EUSIPCO*, Marrakech, Morocco, Sep. 2013.

[SiDu13] B.C.Sicleru, B.Dumitrescu - POS3POLY - a MATLAB Preprocessor for Optimization with Positive Polynomials, *Optimization and Engineering*, vol.14, no.2, pp.251-273, 2013.

Director proiect,
Prof. Bogdan Dumitrescu

A handwritten signature in black ink, appearing to read 'B. Dumitrescu', is positioned below the typed name.