

Noi algoritmi de tip RLS cu convergență variabilă

SINTEZA FAZEI UNICE / 2012

Contract nr. 7/5.08.2010, Cod TE-50

Algoritmii de tip RLS (*recursive least-squares*) au o aplicabilitate mai redusă comparativ cu algoritmii LMS (*least-mean-square*), în special datorită complexității aritmetice ridicate și a potențialelor probleme de instabilitate numerică. Cu toate acestea, viteza de convergență este net superioară algoritmilor de tip LMS. Parametrul de adaptare din cadrul algoritmilor de tip RLS poartă denumirea de parametru de ponderare sau “factor de uitare” (*forgetting factor*). Este un parametru pozitiv și subunitar, care influențează “memoria” algoritmului. Utilizarea unei valori foarte apropiate de 1 este benefică pentru dezadaptarea algoritmului, dar influențează negativ capacitatea sa de urmărire. Pentru a crește această capacitate de urmărire, valoarea parametrului de ponderare trebuie redusă, ceea ce conduce însă la creșterea dezadaptării. Aceasta a fost motivația ce a stat la baza dezvoltării algoritmilor de tip RLS cu convergență variabilă (VFF - *variable forgetting factor*). O primă soluție are la bază algoritmi de tip Newton, ceea ce implică o complexitate aritmetică foarte ridicată. Mai recent a fost propus un algoritm de tip VFF-RLS ce determină parametrul de ponderare variabil într-o manieră recursivă, prin minimizarea gradientului unei funcții cost definită pe baza energiei erorii. Deși performanțele obținute sunt superioare comparativ cu cele ale algoritmului clasic, procedura de calcul a parametrului de ponderare rămâne destul de complexă.

În cadrul acestei etape de cercetare a proiectului, am dezvoltat o nouă familie de algoritmi adaptivi de tip RLS cu convergență variabilă, pe baza controlului parametrului de ponderare (factorului de uitare). S-au elaborat expresii optimale pentru acest parametru, ținând cont de prezența și nivelul zgomotului ce afectează semnalul de referință al filtrului adaptiv. Prezentarea detaliată a acestor rezultate, precum și extinderea în cazul altor algoritmi au fost publicate în [1], [2], [3], [4]. În această sinteză, se vor prezenta succint principalele contribuții în contextul aplicației de compensare a ecoului acustic stereofonic pe baza modelului WL (*widely linear*) [5].

În cadrul acestei aplicații este necesară identificarea unui sistem cu două intrări și două ieșiri (cu variabile reale). Cele două semnale de intrare vor fi notate cu $x_L(n)$ și $x_R(n)$, iar semnalele de ieșire cu $d_L(n)$ și $d_R(n)$, unde n reprezintă indicele de timp. Prin urmare, obținem:

$$d_L(n) = y_L(n) + v_L(n), \quad (1)$$

$$d_R(n) = y_R(n) + v_R(n), \quad (2)$$

unde $y_L(n)$ și $y_R(n)$ reprezintă semnalele de ecou, iar $v_L(n)$ și $v_R(n)$ sunt semnalele de la capătul apropiat (*near-end*). Semnalele de ecou pot fi exprimate ca:

$$y_L(n) = \mathbf{h}_{t,LL}^T \mathbf{x}_L(n) + \mathbf{h}_{t,RL}^T \mathbf{x}_R(n), \quad (3)$$

$$y_R(n) = \mathbf{h}_{t,LR}^T \mathbf{x}_L(n) + \mathbf{h}_{t,RR}^T \mathbf{x}_R(n), \quad (4)$$

unde $\mathbf{h}_{t,LL}$, $\mathbf{h}_{t,RL}$, $\mathbf{h}_{t,LR}$, $\mathbf{h}_{t,RR}$ sunt răspunsurile la impuls acustice, T este operația de transpunere și

$$\begin{aligned} \mathbf{x}_L(n) &= [x_L(n) \ x_L(n-1) \ \cdots \ x_L(n-L+1)]^T \\ \mathbf{x}_R(n) &= [x_R(n) \ x_R(n-1) \ \cdots \ x_R(n-L+1)]^T \end{aligned}$$

sunt semnalele de difuzor. Pentru a compensa ecoul este necesar să estimăm cele patru răspunsuri la impuls $\mathbf{h}_{t,LL}$, $\mathbf{h}_{t,RL}$, $\mathbf{h}_{t,LR}$, $\mathbf{h}_{t,RR}$, din semnalele de microfon $d_L(n)$ și $d_R(n)$.

În [4] am propus o modificare a schemei clasice cu două intrări și două ieșiri, ca o un sistem cu o intrare și o ieșire, dar cu variabile complexe. Prin urmare, putem scrie:

$$d(n) = d_L(n) + jd_R(n) = y(n) + v(n), \quad (5)$$

unde $j = \sqrt{-1}$, $y(n) = y_L(n) + jy_R(n)$ și $v(n) = v_L(n) + jv_R(n)$. De asemenea, definim:

$$\mathbf{x}(n) = \mathbf{x}_L(n) + j\mathbf{x}_R(n). \quad (6)$$

Astfel rezultă semnalul de ecou (complex)

$$y(n) = \mathbf{h}_t^H \mathbf{x}(n) + \mathbf{h}_t^H \mathbf{x}^*(n), \quad (7)$$

unde H și $*$ reprezintă operația hermitică și respectiv de conjugare, iar

$$\mathbf{h}_t = \mathbf{h}_{t,1} + j\mathbf{h}_{t,2}, \quad (8)$$

$$\mathbf{h}'_t = \mathbf{h}'_{t,1} + j\mathbf{h}'_{t,2}, \quad (9)$$

cu $\mathbf{h}_{t,1} = (\mathbf{h}_{t,LL} + \mathbf{h}_{t,RR})/2$, $\mathbf{h}_{t,2} = (\mathbf{h}_{t,RL} - \mathbf{h}_{t,LR})/2$, $\mathbf{h}'_{t,1} = (\mathbf{h}_{t,LL} - \mathbf{h}_{t,RR})/2$, și $\mathbf{h}'_{t,2} = -(\mathbf{h}_{t,RL} + \mathbf{h}_{t,LR})/2$. În acest context, putem exprima (7) ca

$$y(n) = \tilde{\mathbf{h}}_t^H \tilde{\mathbf{x}}(n), \quad (10)$$

unde $\tilde{\mathbf{h}}_t = [\mathbf{h}_t^T \ \mathbf{h}_t^{*T}]^T$ și $\tilde{\mathbf{x}}(n) = [\mathbf{x}^T(n) \ \mathbf{x}^H(n)]^T$. În consecință, rezultă că

$$d(n) = \tilde{\mathbf{h}}_t^H \tilde{\mathbf{x}}(n) + v(n). \quad (11)$$

Se observă că avem acum un răspuns la impuls complex, de lungime $2L$, $\tilde{\mathbf{h}}_t$, cu intrarea și ieșirea complexe, $x(n)$ și $d(n)$. Din (7) sau (10), se poate recunoaște modelul WL pentru variabile complexe. De asemenea, această abordare este în concordanță cu principiul dualității.

Pentru a compensa ecoul trebuie estimat sistemul $\tilde{\mathbf{h}}_t$. Notăm cu $\tilde{\mathbf{h}}(n)$ un filtru adaptiv și cu

$$e(n) = d(n) - \tilde{\mathbf{h}}^H(n-1)\tilde{\mathbf{x}}(n) = d(n) - \hat{y}(n) \quad (12)$$

semnalul de eroare la momentul de timp n . Vom redefini vectorul semnalului de eroare (de lungime $2L$) ca

$$\tilde{\mathbf{x}}(n) = [\boldsymbol{\chi}^T(n) \quad \boldsymbol{\chi}^T(n-1) \quad \cdots \quad \boldsymbol{\chi}^T(n-L+1)]^T, \quad (13)$$

unde $\boldsymbol{\chi}(n) = [x(n) \quad x^*(n)]^T$. Astfel obținem

$$\begin{aligned} \tilde{\mathbf{h}}_t &= [h_{t,0} \quad h'_{t,0} \quad \cdots \quad h_{t,L-1} \quad h'_{t,L-1}]^T, \\ \tilde{\mathbf{h}}(n) &= [h_0(n) \quad h'_0(n) \quad \cdots \quad h_{L-1}(n) \quad h'_{L-1}(n)]^T, \end{aligned}$$

unde $h_{t,l}$, $h'_{t,l}$, $h_l(n)$ și $h'_l(n)$, cu $l = 0, 1, \dots, L-1$, sunt elementele vectorilor \mathbf{h}_t , \mathbf{h}'_t , $\mathbf{h}(n)$ și respectiv $\mathbf{h}'(n)$. Aceste noi definiții nu modifică expresiile semnalelor $d(n)$ și $e(n)$ din relațiile (11) și respectiv (12).

În acest context, funcția cost în sens LS (*least-squares*) este definită ca

$$J[\tilde{\mathbf{h}}(n)] = \sum_{i=1}^n \lambda^{n-i} |d(i) - \tilde{\mathbf{h}}^H(n)\tilde{\mathbf{x}}(i)|^2, \quad (14)$$

unde λ ($0 \ll \lambda < 1$) este factorul de uitare (parametrul de ponderare). Astfel, putem exprima (14) ca

$$J[\tilde{\mathbf{h}}(n)] = \sum_{i=1}^n \lambda^{n-i} |d(i)|^2 - \tilde{\mathbf{h}}^H(n)\mathbf{p}_{\tilde{\mathbf{x}}d}(n) - \mathbf{p}_{\tilde{\mathbf{x}}d}^H(n)\tilde{\mathbf{h}}(n) + \tilde{\mathbf{h}}^H(n)\mathbf{R}_{\tilde{\mathbf{x}}}(n)\tilde{\mathbf{h}}(n),$$

unde $\mathbf{R}_{\tilde{\mathbf{x}}}(n) = \sum_{i=1}^n \lambda^{n-i}\tilde{\mathbf{x}}(i)\tilde{\mathbf{x}}^H(i)$ și $\mathbf{p}_{\tilde{\mathbf{x}}d}(n) = \sum_{i=1}^n \lambda^{n-i}\tilde{\mathbf{x}}(i)d^*(i)$. Minimizarea funcției $J[\tilde{\mathbf{h}}(n)]$ în raport cu $\tilde{\mathbf{h}}(n)$ conduce la setul de ecuații normale:

$$\mathbf{R}_{\tilde{\mathbf{x}}}(n)\tilde{\mathbf{h}}(n) = \mathbf{p}_{\tilde{\mathbf{x}}d}(n). \quad (15)$$

Presupunând că $\mathbf{R}_{\tilde{\mathbf{x}}}(n) > 0$ (matricea este pozitiv definită), filtrul optimal în sens LS poate fi dedus ca $\tilde{\mathbf{h}}(n) = \mathbf{R}_{\tilde{\mathbf{x}}}^{-1}(n)\mathbf{p}_{\tilde{\mathbf{x}}d}(n)$. De asemenea, se poate evalua recursiv:

$$\mathbf{p}_{\tilde{\mathbf{x}}d}(n) = \lambda\mathbf{p}_{\tilde{\mathbf{x}}d}(n-1) + \tilde{\mathbf{x}}(n)d^*(n), \quad (16)$$

$$\mathbf{R}_{\tilde{\mathbf{x}}}(n) = \lambda\mathbf{R}_{\tilde{\mathbf{x}}}(n-1) + \tilde{\mathbf{x}}(n)\tilde{\mathbf{x}}^H(n). \quad (17)$$

Pe baza identității Woodbury aplicată în (17), inversa matricei $\mathbf{R}_{\tilde{\mathbf{x}}}(n)$ se poate exprima

$$\mathbf{R}_{\tilde{\mathbf{x}}}^{-1}(n) = \lambda^{-1}\mathbf{R}_{\tilde{\mathbf{x}}}^{-1}(n-1) - \lambda^{-1}\tilde{\mathbf{k}}(n)\tilde{\mathbf{x}}^H(n)\mathbf{R}_{\tilde{\mathbf{x}}}^{-1}(n-1), \quad (18)$$

unde

$$\tilde{\mathbf{k}}(n) = \frac{\mathbf{R}_{\tilde{\mathbf{x}}}^{-1}(n-1)\tilde{\mathbf{x}}(n)}{\lambda + \tilde{\mathbf{x}}^H(n)\mathbf{R}_{\tilde{\mathbf{x}}}^{-1}(n-1)\tilde{\mathbf{x}}(n)} \quad (19)$$

este vectorul câștig Kalman. Pe baza (19) și ținând cont de (18), vectorul câștig Kalman se poate exprima de asemenea ca

$$\tilde{\mathbf{k}}(n) = \mathbf{R}_{\tilde{\mathbf{x}}}^{-1}(n)\tilde{\mathbf{x}}(n). \quad (20)$$

Astfel, soluția pentru (15) rezultă ca

$$\tilde{\mathbf{h}}(n) = \lambda \mathbf{R}_{\tilde{\mathbf{x}}}^{-1}(n) \mathbf{p}_{\tilde{\mathbf{x}}d}(n-1) + \mathbf{R}_{\tilde{\mathbf{x}}}^{-1}(n) \tilde{\mathbf{x}}(n) d^*(n). \quad (21)$$

Înlocuind (18) în primul termen al părții drepte din (21) și utilizând (12), relația de reactualizare a algoritmului RLS rezultă ca

$$\tilde{\mathbf{h}}(n) = \tilde{\mathbf{h}}(n-1) + \tilde{\mathbf{k}}(n) e^*(n). \quad (22)$$

Se poate observa că acest algoritm poate fi interpretat ca un algoritm RLS cu două canale cu intrările $x(n)$ și $x^*(n)$.

Algoritmul RLS prezentat anterior utilizează o valoare constantă a factorului de uitare λ , ceea ce conduce la un compromis între criteriile de performanță ale filtrului adaptiv. Pentru a dezvolta o variantă VFF a acestui algoritm, se pleacă de la premisa că semnalul de referință $d(n)$ din (11) conține ecoul complex $y(n)$ [a se vedea (10)] și semnalul de la capătul apropiat (*near-end*), $v(n)$. În acest context, scopul filtrului adaptiv nu este de a anula eroarea, ci de a recupera semnalul de la capătul apropiat în eroarea filtrului adaptiv.

După cum se poate observa, semnalul $e(n)$ din (12) reprezintă eroarea a priori a filtrului, deoarece este calculată pe baza coeficienților de la momentul de timp $n-1$. În contextul modelului WL, eroarea a posteriori (rezultată pe baza filtrului adaptiv de la momentul de timp n) este definită ca

$$\epsilon(n) = d(n) - \tilde{\mathbf{h}}^H(n) \tilde{\mathbf{x}}(n). \quad (23)$$

Utilizând (12) și (22) în (23), relația dintre erorile a posteriori și a priori rezultă ca

$$\epsilon(n) = e(n) \left[1 - \tilde{\mathbf{k}}^H(n) \tilde{\mathbf{x}}(n) \right]. \quad (24)$$

Fie $u(n) = \tilde{\mathbf{x}}^H(n) \mathbf{R}_{\tilde{\mathbf{x}}}^{-1}(n-1) \tilde{\mathbf{x}}(n)$, care este o variabilă reală deoarece matricea $\mathbf{R}_{\tilde{\mathbf{x}}}^{-1}(n-1)$ este hermitică. Astfel, ținând cont de (19), relația (24) se poate rescrie ca

$$\epsilon(n) = e(n) \frac{\lambda}{\lambda + u(n)}. \quad (25)$$

Condiția de recuperare a semnalului de la capătul apropiat în eroarea filtrului adaptiv poate fi impusă în termenii estimațiilor de putere astfel:

$$E \left[|\epsilon(n)|^2 \right] = \sigma_v^2(n), \quad (26)$$

unde $\sigma_v^2(n) = E \left[|v(n)|^2 \right]$ este varianța semnalului $v(n)$. Multiplicând (25) cu conjugatul ei, apoi aplicând operatorul de mediere statistică și presupunând că semnalul de intrare și eroarea sunt necorelate (ceea ce este adevărat în faza de convergență), condiția (26) devine

$$\lambda^2 \left[\sigma_e^2(n) - \sigma_v^2(n) \right] - 2\lambda \sigma_v^2(n) \sigma_u(n) - \sigma_v^2(n) \sigma_u^2(n) = 0, \quad (27)$$

unde $\sigma_e^2(n) = E \left[|e(n)|^2 \right]$ și $\sigma_u^2(n) = E \left[u^2(n) \right]$ sunt varianțele lui $e(n)$ și respectiv $u(n)$. În continuare, să presupunem că factorul de uitare este deterministic și dependent de timp. Astfel, rezolvând ecuația (27), rezultă expresia pentru VFF:

$$\lambda(n) = \frac{\sigma_u(n) \sigma_v(n)}{\sigma_e(n) - \sigma_v(n)}. \quad (28)$$

Pentru a evalua (28) în practică avem nevoie de estimății parametrilor $\sigma_e^2(n)$, $\sigma_u^2(n)$ și $\sigma_v^2(n)$. Deoarece semnalele $e(n)$ și $u(n)$ sunt disponibile, estimății lor de putere pot fi evaluați recursiv ca

$$\hat{\sigma}_e^2(n) = \alpha \hat{\sigma}_e^2(n-1) + (1-\alpha) |e(n)|^2, \quad (29)$$

$$\hat{\sigma}_u^2(n) = \alpha \hat{\sigma}_u^2(n-1) + (1-\alpha) u^2(n), \quad (30)$$

unde $\alpha = 1 - 1/(2KL)$, cu $K \geq 1$; valorile inițiale sunt $\hat{\sigma}_e^2(0) = 0$ și $\hat{\sigma}_u^2(0) = 0$.

Estimarea parametrului $\sigma_v^2(n)$ este mai dificilă. De exemplu, în cazul vorbirii simultane (*double-talk*), semnalul de la capătul apropiat este o combinație între zgomotul de fond și semnalul vocal. În acest caz, parametrul $\sigma_v^2(n)$ poate fi estimat plecând de la presupunerea că filtrul adaptiv a convers deja, astfel încât

$$\hat{\sigma}_v^2(n) = |\hat{\sigma}_d^2(n) - \hat{\sigma}_y^2(n)|, \quad (31)$$

unde $\hat{\sigma}_d^2(n)$ și $\hat{\sigma}_y^2(n)$ sunt estimății de putere a lui $d(n)$ și respectiv $\hat{y}(n)$ [a se vedea (12)], calculați similar cu (29).

În acest moment, având la dispoziție toți parametrii necesari în (28), se poate observa comportarea globală a factorului de uitare variabil $\lambda(n)$. Se observă că la numitorul relației (28) avem teoretic $\sigma_e(n) \geq \sigma_v(n)$. Cu toate acestea, deoarece se utilizează estimății de putere, această condiție poate să nu fie adevărată în orice moment de timp; pentru a evita această situație, se poate utiliza modulul numitorului. În faza de convergență inițială sau când are loc o schimbare de sistem, $\sigma_e(n)$ este mult mai mare decât $\sigma_v(n)$. Prin urmare, numitorul relației (28) crește, astfel încât $\lambda(n)$ scade, oferind astfel o convergență rapidă și o capacitate de urmărire crescută. Atunci când algoritmul se află în faza de convergență, $\sigma_e(n) \approx \sigma_v(n)$. Pentru a evita probleme de natură numerică în (28), propunem să limităm valoarea lui $\lambda(n)$ la o limită maximă λ_{\max} (foarte apropiată de 1). De asemenea, pentru a evita calcule suplimentare, putem impune ca $\lambda(n) = \lambda_{\max}$ atunci când algoritmul se află în faza de convergență, prin verificarea condiției

$$\hat{\sigma}_e(n) \leq \gamma \hat{\sigma}_v(n), \quad (32)$$

unde $1 < \gamma \leq 2$. Astfel, atunci când (32) este îndeplinită, valoarea factorului de uitare atinge valoarea maximă, conducând astfel la o dezadaptare scăzută. În consecință, expresia (28) poate fi rescrisă astfel:

$$\lambda(n) = \begin{cases} \min \left[\frac{\hat{\sigma}_u(n) \hat{\sigma}_v(n)}{\varepsilon + |\hat{\sigma}_e(n) - \hat{\sigma}_v(n)|}, \lambda_{\max} \right], & \text{dacă } \hat{\sigma}_e(n) > \gamma \hat{\sigma}_v(n) \\ \lambda_{\max}, & \text{dacă } \hat{\sigma}_e(n) \leq \gamma \hat{\sigma}_v(n) \end{cases} \quad (33)$$

unde ε este o constantă pozitivă mică ce prevede împărțirea prin zero. Menționăm că volumul de calcule necesar evaluării VFF este mult mai redus comparativ cu complexitatea globală a algoritmului RLS.

O serie de simulări au fost efectuate în contextul unei aplicații de compensare a ecoului acustic stereofonic, în contextul modelului WL. Frecvență de eșantionare este de 8 kHz, iar filtrul adaptiv are lungimea $2L = 1024$. Semnalul de intrare $x(n)$ este o secvență vocală. Zgomotul de fond este presupus alb și Gaussian, cu SNR = 30 dB. Performanțele algoritmilor au fost evaluate prin prisma dezalinierei normate (în dB) și a MSE. Rezultatele prezentate în figurile anexate justifică validitatea soluției VFF propuse.

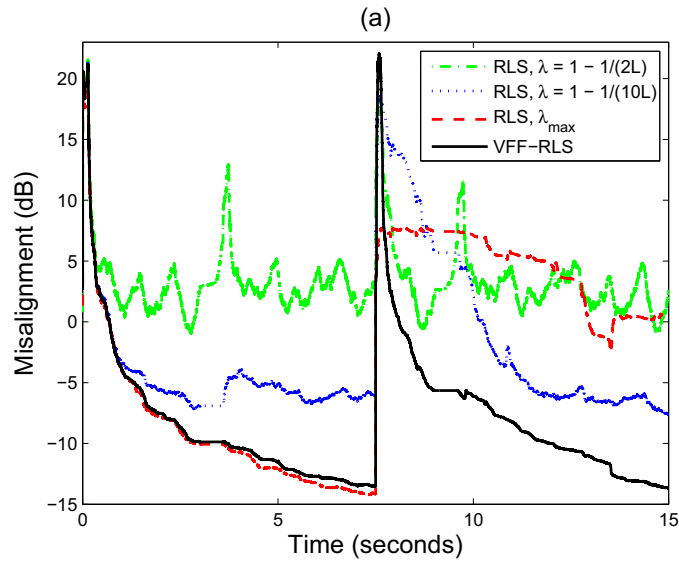


Figure 1: Dezalinierea algoritmilor RLS cu $\lambda = 1 - 1/(2L)$, $\lambda = 1 - 1/(10L)$, $\lambda_{\max} = 0,99999$ și VFF-RLS. Scenariu *single-talk*, calea de ecou se schimbă după 7,5 secunde.

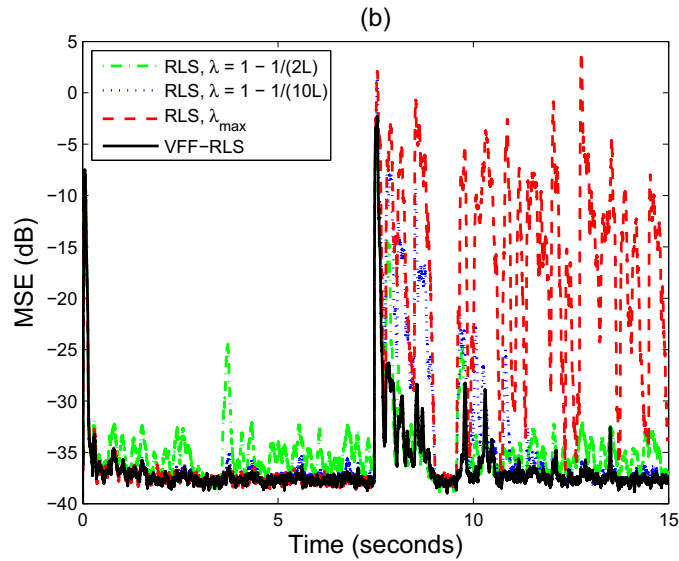


Figure 2: MSE pentru algoritmii RLS cu $\lambda = 1 - 1/(2L)$, $\lambda = 1 - 1/(10L)$, $\lambda_{\max} = 0,99999$ și VFF-RLS. Scenariu *single-talk*, calea de ecou se schimbă după 7,5 secunde.

References

- [1] C. Stanciu, J. Benesty, C. Paleologu, T. Gänslér, and S. Ciocinã, "A novel perspective on stereophonic acoustic echo cancellation," in *Proc. ICASSP*, 2012, pp. 25–28.
- [2] C. Stanciu, C. Paleologu, J. Benesty, T. Gänslér, and S. Ciocinã, "An efficient RLS algorithm for stereophonic acoustic echo cancellation with the widely linear model," in *Proc. Inter-Noise*, 2012, 12 pagini.
- [3] C. Stanciu, C. Paleologu, J. Benesty, S. Ciocinã, and F. Albu, "A variable-forgetting factor RLS algorithm for stereophonic acoustic echo cancellation with the widely linear model," in *Proc. EUSIPCO*, 2012, pp. 1960–1964.

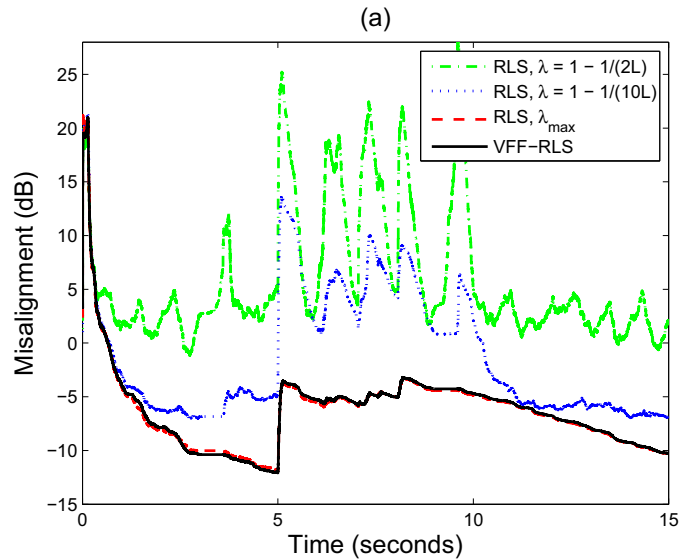


Figure 3: Dezaliniera algoritmiilor RLS cu $\lambda = 1 - 1/(2L)$, $\lambda = 1 - 1/(10L)$, $\lambda_{\max} = 0,99999$ și VFF-RLS. Scenariu *double-talk*, vocea de la *near-end* apare între 5 și 10 secunde.

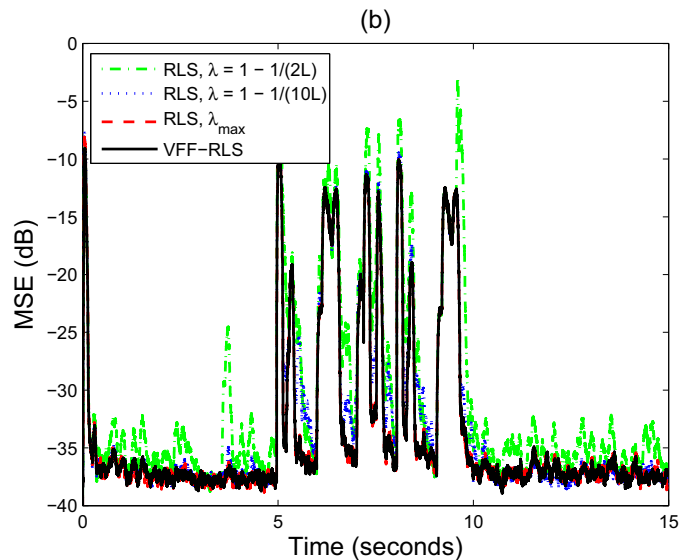


Figure 4: MSE pentru algoritmiile RLS cu $\lambda = 1 - 1/(2L)$, $\lambda = 1 - 1/(10L)$, $\lambda_{\max} = 0,99999$ și VFF-RLS. Scenariu *double-talk*, vocea de la *near-end* apare între 5 și 10 secunde.

[4] C. Stanciu, J. Benesty, C. Paleologu, T. Gänslér, and S. Ciochină, "A widely linear model for stereophonic acoustic echo cancellation," *Signal Processing*, vol. 93, pp. 511–516, Feb. 2013.

[5] J. Benesty, C. Paleologu, T. Gänslér, and S. Ciochină, *A Perspective on Stereophonic Acoustic Echo Cancellation*. Berlin, Germany: Springer-Verlag, 2011.

Director de proiect

Conf. dr. ing. Constantin Paleologu