KOMBINIERTE ANSÄTZE ZEITVARIABLER UND ZEITINVARIANTER MODELLANALYSEN FÜR DIE SPRACHVERARBEITUNG

Karl Schnell, Arild Lacroix

Institut für Angewandte Physik, Goethe-Universität Frankfurt,
Max-von-Laue-Str. 1, 60438 Frankfurt am Main
schnell@iap.uni-frankfurt.de

Abstract: In diesem Beitrag wird für die Sprachanalyse ein Schätzalgorithmus vorgestellt, der eine analytische Lösung einer kombinierten zeitinvarianten und zeitvariablen Schätzung bereitstellt. Durch die Kombination ist es insbesondere möglich, die zeitvariable Schätzung an den Stellen der Modellparameter-Trajektorie zu unterstützen, an denen die Nebenbedingungen der zeitvariablen Analyse keine hinreichend glatten Verläufe fordern können. Das für die Modellschätzung zugrunde gelegte Modell ist ein Kreuzgliedfilter mit Reflexionskoeffizienten. Erste Ergebnisse zeigen auf, dass die kombinierte Analyse für die Schätzung von Modellparameter-Trajektorien vorteilhaft ist.

1 Einleitung

Die zeitinvariante lineare Prädiktion stellt eine etablierte Technik in der Sprachverarbeitung dar, mit der sich Modellschätzungen des Sprechtraktes aus dem Sprachsignal erzielen lassen [1-3]. Ein für Anwendungen großer Vorteil der zeitinvarianten linearen Prädiktion liegt darin begründet, dass sie eine analytische Modellschätzung auf Grundlage eines Nur-Pole-Modells ermöglicht. Da der Sprachproduktionsprozess ein zeitvariabler Prozess ist, wird üblicherweise für die zeitinvariante Analyse das Sprachsignal in kurze als stationär angenommene Segmente zerlegt, die dann einzeln analysiert werden. Soll die Zeitvariabilität des Sprachproduktionsprozesses explizit berücksichtigt werden, so müssen die Modellparameter auch zeitvariabel geschätzt werden. Bei diesem Ansatz sollten die möglichen Trajektorien der Modellparameter in ihrer Dynamik beschränkt werden, um unrealistische Trajektorien auszuschließen. Für die Dynamikbeschränkung sollte allerdings beachtet werden, dass der Sprachproduktionsprozess für bestimmte Laute, wie z. B. Explosivlaute, auch eine recht schnelle Dynamik aufweisen kann. Weiterhin ist es oftmals wünschenswert, den Einfluss der in der Regel eher langsamen Artikulationsbewegungen von dem der schnellen Glottisschwingungen zu separieren. Als Schätzalgorithmen einer zeitvariablen Analyse kommt einmal die Klasse der adaptiven Filter in Betracht [4], die allerdings in der Regel keine analytischen Lösungen bereitstellen. Eine Lösung mittels Kalman-Filtern ist in [5-6] präsentiert worden, während in [7] eine spezielle adaptive Prädiktion einzelner zeitvariabler Resonanzen verwendet wurde. Ein Ansatz, mit dem auch analytische Lösungen möglich sind, ist durch die Entwicklung der Parametertrajektorien nach Basisfunktionen gegeben. Dieses Prinzip wurde für die Analyse eines Segmentes in [8] für Direktform-Koeffizienten und in [9] auch für Reflexionskoeffizienten vorgestellt. Darauf aufbauend sind in [10-13] für unterschiedliche Zwecke verbundene Mehr-Segment-Analysen mit entsprechenden Basisfunktionen vorgestellt worden. Motiviert durch die Analyseergebnisse in [13], die mittels eines iterativen Schätzalgorithmus' erzielt wurden, werden hier auch zeitinvariante Schätzkomponenten in der zeitvariablen Analyse berücksichtigt. In diesem Beitrag wird eine Erweiterung der rein zeitvariablen Schätzalgorithmen in allgemeiner Form vorgestellt, die eine kombinierte Schätzung einer zeitvariablen und zeitinvarianten Analyse beinhaltet.

2 Lineare Prädiktion

Als Koeffizientendarstellung für die Modellschätzung werden hier die Reflexionskoeffizienten verwendet, da sie bessere Interpolationseigenschaften als die Direktform-Koeffizienten aufweisen. Die Schätzung wird durch eine sukzessive Leistungsminimierung der Ausgänge der einzelnen Kreuzglieder des FIR-Kreuzgliedfilters realisiert. Diese Vorgehensweise ist analog zum Burg-Algorithmus der zeitinvarianten Analyse. Im Folgenden wird exemplarisch die Schätzung eines einzelnen Kreuzgliedes mit dem Reflexionskoeffizienten r behandelt. Ein Kreuzglied mit Ein- und Ausgängen ist in

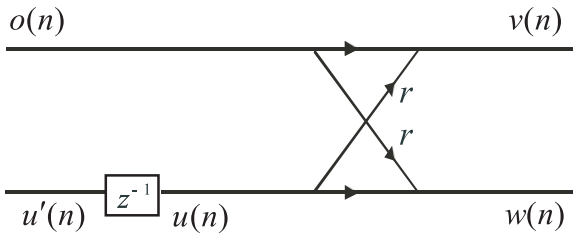

Abbildung 1 – Element des FIR-Kreuzgliedfilters mit Ein- und Ausgängen.

Abb. 1 gezeigt. Es wird nun angenommen, dass die Eingangssignale $o(n)$ und $u'(n)$ bzw. $u(n)$ vorliegen. Die zu verarbeitenden Signale bzw. Segmente können durch Vektoren mit

$$\boldsymbol{x} = (x(1),\ldots,x(L_\mathrm{c}))^\mathrm{T},\ \boldsymbol{o} = (o(1),\ldots,o(L_\mathrm{c}))^\mathrm{T} \text{ und } \boldsymbol{u} = (u(1),\ldots,u(L_\mathrm{c}))^\mathrm{T} \tag{1}$$

dargestellt werden. $x(n)$ stellt für den Abschnitt $n = 1\ldots L_c$ das zu analysierende Segment dar, das für eine einfache Beschreibung bei der Position $n = 1$ anfängt. Für das erste Kreuzglied können die Eingangssignale mit $\boldsymbol{o} = \boldsymbol{x}$ und $\boldsymbol{u}' = \pm\boldsymbol{x}$ initialisiert werden. Das Vorzeichen für die Initialisierung von $\boldsymbol{u}'$ ist für die Prädiktion eigentlich positiv, es kann allerdings für einen Rohrabschluss an den Lippen auch negativ angenommen werden, was sich allerdings nicht auf die Modellschätzung als solches, sondern nur auf die Flächen des zugehörigen Rohrmodells auswirkt. Für die Schätzung werden hier stückweise lineare Trajektorien angenommen, die in ihrem gesamten Verlauf stetig sind. Der Raum dieser Trajektorien kann durch die in [10] definierten Basisfunktionen bzw. -vektoren aufgespannt werden. Seien $m_i = (i-1)\cdot L_a + 1$ für $i = 1\ldots P$ die Anfangspositionen der $P-1$ linearen Abschnitte der Länge L_a, wobei $m_P - 1$ als Endposition fungiert, so ergeben sich $M = P - 2$ Basisvektoren $\boldsymbol{g}_k = (g(1),\ldots,g(L_\mathrm{c}))^\mathrm{T}$ mit

$$g_1(n) = 1 \quad \text{für} \quad n = 1\ldots m_P - 1$$

$$g_k(n) = \begin{cases} 0 & \text{für} \quad n = 1\ldots m_k - 1 \\ \dfrac{n - m_k}{L_a - 1} & \text{für} \quad n = m_k \ldots m_{k+1} - 1 \\ 1 & \text{für} \quad n = m_{k+1}\ldots m_P - 1. \end{cases} \qquad \text{für} \quad k = 2\ldots M \tag{2}$$

Der erste und der letzte Abschnitt der Trajektorie $r(n)$ für $n = 1\ldots m_1 - 1$ und $n = m_{P-1}\ldots m_P - 1$ werden hier als konstant angenommen. Die letzte Marke $m_P = L_c + 1$ entspricht der gesamten Segmentlänge plus Eins. Die stückweise lineare Trajektorie eines Reflexionskoeffizienten kann somit durch die Entwicklung

$$r(n) = \sum_{k=1}^{M} d_k \cdot g_k(n) \quad \text{bzw.} \quad \boldsymbol{r} = \sum_{k=1}^{M} d_k \cdot \boldsymbol{g}_k \tag{3}$$

dargestellt werden; die rechte Darstellung in Gl. (3) liegt als Vektornotation vor. Das obere Ausgangssignal des Kreuzgliedes ergibt sich damit nach Abb. 1 zu

$$\boldsymbol{v} = \boldsymbol{r} \odot \boldsymbol{u} + \boldsymbol{o} = \sum_{k=1}^{M} d_k \cdot \boldsymbol{g}_k \odot \boldsymbol{u} + \boldsymbol{o} = \sum_{k=1}^{M} d_k \cdot \boldsymbol{f}_k + \boldsymbol{o} \,. \tag{4}$$

mit der Definition $\boldsymbol{f}_k = \boldsymbol{g}_k \odot \boldsymbol{u}$. Die Operation $\odot$ stellt eine elementweise Multiplikation mit $\boldsymbol{f} = \boldsymbol{u} \odot \boldsymbol{g} \;\rightarrow\; f(k) = u(k) \cdot g(k)$ dar. Der obere Ausgang $\boldsymbol{v}$ des Kreuzgliedes repräsentiert den Vorwärts-Prädiktionsfehler, der hier auch mit $\boldsymbol{\varepsilon}^{\mathrm{tv}} = \boldsymbol{v}$ bezeichnet wird und für die Schätzung zu minimieren ist. Damit kann die Prädiktion durch

$$\boldsymbol{o} = -\sum_{k=1}^{M} d_k \cdot \boldsymbol{f}_k + \boldsymbol{\varepsilon}^{\mathrm{tv}} \tag{5}$$

dargestellt werden, wobei durch eine Leistungsminimierung von $\varepsilon^{\mathrm{tv}}(n)$ die optimalen Koeffizienten folgen. Gl. (5) stellt die Prädiktionsgleichung für eine rein zeitvariable Modellschätzung dar. Für eine zusätzliche Einbeziehung von zeitinvarianten Schätzkomponenten werden Segmente verwendet, die als Mittelpunkte die Positionen m_i für $i = 3 \ldots P-2$ aufweisen. Damit können die für die Schätzung kritischen Positionen m_i, an denen die zeitvariablen Trajektorien Knicke aufweisen können, die Schätzung unterstützen. Die Positionen für $i < 3$ und $i > P-2$ sind ausgelassen, da sie jeweils ein Ende des ersten oder letzten konstanten Trajektorienabschnittes darstellen. Die Segmente für die zeitinvariante Prädiktion werden durch die Vektoren

$$\boldsymbol{v}_i = (v(m_{i+2} - L_{\mathrm{ti}}/2), \ldots, v(m_{i+2} + L_{\mathrm{ti}}/2 - 1))^{\mathrm{T}} \tag{6}$$

für $i = 1 \ldots M-2$ dargestellt; analoges gilt auch für $\boldsymbol{o}_i$ und $\boldsymbol{u}_i$. Die Gleichungen der Prädiktionsfehler $\boldsymbol{\varepsilon}_i^{\mathrm{ti}} = \boldsymbol{v}_i$ der zeitinvarianten Schätzungen ergeben sich zu

$$\boldsymbol{\varepsilon}_i^{\mathrm{ti}} = r_i^{\mathrm{ti}} \cdot \boldsymbol{u}_i + \boldsymbol{o}_i \,. \tag{7}$$

Die skalaren Reflexionskoeffizienten r_i^{ti} der zeitinvarianten Prädiktion entsprechen den mittigen Positionen der Segmente $\boldsymbol{o}_i$ und sind mit der Parametertrajektorie durch die Beziehung $r_i^{\mathrm{ti}} = r(m_{i+2})$ verbunden. Der Zusammenhang mit den Koeffizienten der Basisvektoren $\boldsymbol{g}_k$ ergibt sich entsprechend den Definitionen von Gl. (2) damit zu

$$r_i^{\mathrm{ti}} = \sum_{k=1}^{i+1} d_k \quad \text{bzw.} \quad \boldsymbol{\varepsilon}_i^{\mathrm{tv}} = \sum_{k=1}^{i+1} d_k \boldsymbol{u}_i + \boldsymbol{o}_i \,, \tag{8}$$

womit die zeitinvariante und zeitvariable Prädiktion miteinander gekoppelt sind. Die Prädiktionsgleichungen der zeitinvarianten Schätzungen ergeben sich zu

$$\boldsymbol{o}_i = -\sum_{k=1}^{i+1} d_k \boldsymbol{u}_i + \boldsymbol{\varepsilon}_i^{\mathrm{ti}} \;. \tag{9}$$

Die Prädiktionsgleichungen (9) werden jeweils mit einem Hamming-Fenster $w(k)$ gewichtet, das durch einen Gewichtungsvektor $\boldsymbol{w} = c_w (w(1), \ldots w(L_{\mathrm{ti}}))^{\mathrm{T}}$ repräsentiert wird. Der Faktor c_w ist in der Weise gestaltet, sodass der Mittelwert von $w^2(n)$ gleich Eins ist. Die Gleichungen der zeitinvarianten Analysen enthalten dann die gewichteten Vektoren

$$\widehat{\boldsymbol{o}}_i = \boldsymbol{o}_i \odot \boldsymbol{w} \quad \text{und} \quad \widehat{\boldsymbol{u}}_i = \boldsymbol{u}_i \odot \boldsymbol{w} \,. \tag{10}$$

Die Prädiktionsgleichungen der zeitvariablen und zeitinvarianten Schätzungen können mithilfe von zusammengesetzten Vektoren durch eine einzige Vektorgleichung dargestellt werden. Hierfür werden die Vektoren

$$\bar{\boldsymbol{o}} = (\boldsymbol{o}^{\mathrm{T}}, \hat{\boldsymbol{o}}_1^{\mathrm{T}}, \hat{\boldsymbol{o}}_2^{\mathrm{T}} \ldots, \hat{\boldsymbol{o}}_{M-2}^{\mathrm{T}})^{\mathrm{T}} \quad \text{und} \quad \bar{\boldsymbol{u}} = (\boldsymbol{u}^{\mathrm{T}}, \hat{\boldsymbol{u}}_1^{\mathrm{T}}, \hat{\boldsymbol{u}}_2^{\mathrm{T}} \ldots, \hat{\boldsymbol{u}}_{M-2}^{\mathrm{T}})^{\mathrm{T}} \tag{11}$$

definiert, die aus dem vollständigen Segment für die zeitvariable Prädiktion und den einzelnen Segmenten für die zeitinvariante Prädiktion zusammengesetzt werden. Mit Berücksichtigung von Gl. (8) werden erweiterte Basisvektoren $\bar{\boldsymbol{g}}_k$ durch

$$\begin{aligned}
\bar{\boldsymbol{g}}_k &= (\boldsymbol{g}_k^{\mathrm{T}}, \boldsymbol{1}^{\mathrm{T}}, \ldots \boldsymbol{1}^{\mathrm{T}})^{\mathrm{T}} && \text{für } k = 1,2 \\
\bar{\boldsymbol{g}}_k &= (\boldsymbol{g}_k^{\mathrm{T}}, \underbrace{\boldsymbol{0}^{\mathrm{T}}, \ldots \boldsymbol{0}^{\mathrm{T}}}_{(k-2)\times \boldsymbol{0}^{\mathrm{T}}}, \underbrace{\boldsymbol{1}^{\mathrm{T}}, \ldots \boldsymbol{1}^{\mathrm{T}}}_{(M-k)\times \boldsymbol{1}^{\mathrm{T}}})^{\mathrm{T}} && \text{für } M-1 \geq k > 2 \\
\bar{\boldsymbol{g}}_k &= (\boldsymbol{g}_k^{\mathrm{T}}, \boldsymbol{0}^{\mathrm{T}}, \ldots \boldsymbol{0}^{\mathrm{T}})^{\mathrm{T}} && \text{für } k > M-1
\end{aligned} \tag{12}$$

definiert, welche die Basisvektoren der Trajektorie für die zeitvariable Schätzung und zusätzlich die Kopplung mit den zeitinvarianten Schätzungen berücksichtigen. $\boldsymbol{0} = (0, \ldots 0)^{\mathrm{T}}$ und $\boldsymbol{1} = (1, \ldots 1)^{\mathrm{T}}$ stellen Vektoren der Länge L_{ti} dar, die als Werte ausschließlich Nullen bzw. Einsen enthalten. Die Nullen bzw. Einsen resultieren aus Gl. (8), da die dortige Summe sich nicht über alle d_k erstreckt. Damit ergibt sich die Gesamtgleichung zu

$$\bar{\boldsymbol{o}} = -\sum_{k=1}^{M} d_k \cdot \bar{\boldsymbol{g}}_k \odot \bar{\boldsymbol{u}} + \bar{\boldsymbol{\varepsilon}} \,. \tag{13}$$

Der Fehlervektor

$$\bar{\boldsymbol{\varepsilon}} = ((\bar{\boldsymbol{\varepsilon}}^{\mathrm{tv}})^{\mathrm{T}}, (\bar{\boldsymbol{\varepsilon}}^{\mathrm{ti}})^{\mathrm{T}})^{\mathrm{T}} \tag{14}$$

setzt sich aus einem Anteil für die zeitvariable Schätzung und einem Anteil der zeitinvarianten Schätzung zusammen. Diese beiden Anteile können durch einen zusätzlichen Faktor α unterschiedlich gewichtet werden. Der Faktor α liegt zwischen Null und Eins, wobei $\alpha = 1$ einen reinen zeitvariablen Anteil und $\alpha = 0$ einen reinen zeitinvarianten Anteil bedeutet. Da die Anzahl der einzelnen Prädiktionsgleichungen für die zeitinvariante und zeitvariable Schätzung in der Regel unterschiedlich ist, sind die beiden Komponenten der Schätzung schon im Vorhinein unterschiedlich gewichtet. Diese Ungleichheit wird durch Faktoren λ_{tv} und λ_{ti} ausgeglichen, die jeweils mit den reziproken Längen der Vektoren $\bar{\boldsymbol{\varepsilon}}^{\mathrm{tv}}$ und $\bar{\boldsymbol{\varepsilon}}^{\mathrm{ti}}$ korrespondieren. Für eine Gewichtung zwischen den zeitinvarianten und zeitvariablen Schätzkomponenten werden die Vektoren mit dem Gewichtungsvektor $\boldsymbol{\gamma}$, dessen Werte sich zu

$$\gamma(n) = \begin{cases} \sqrt{\alpha \cdot \lambda_{\mathrm{tv}}} & \text{für} \quad n \leq L_{\mathrm{c}} \\ \sqrt{(1-\alpha)\lambda_{\mathrm{ti}}} & \text{für} \quad n > L_{\mathrm{c}} \end{cases} \tag{15}$$

ergeben, multiplikativ gewichtet. Die gewichteten Vektoren ergeben sich zu

$$\tilde{\boldsymbol{o}} = \boldsymbol{\gamma} \odot \bar{\boldsymbol{o}} \quad \text{und} \quad \tilde{\boldsymbol{u}} = \boldsymbol{\gamma} \odot \bar{\boldsymbol{u}} \,, \tag{16}$$

wodurch sich die Prädiktionsgleichung

$$\tilde{\boldsymbol{o}} = -\sum_{k=1}^{M} d_k \cdot \bar{\boldsymbol{g}}_k \odot \tilde{\boldsymbol{u}} + \tilde{\boldsymbol{\varepsilon}} \quad \text{bzw.} \quad \tilde{\boldsymbol{o}} = -\sum_{k=1}^{M} d_k \cdot \tilde{\boldsymbol{f}}_k + \tilde{\boldsymbol{\varepsilon}} \tag{17}$$

mit der Definition $\tilde{\boldsymbol{f}}_k = \bar{\boldsymbol{g}}_k \odot \tilde{\boldsymbol{u}}$ ergibt. Für die Analyse wird der Vektor $\tilde{\boldsymbol{o}}$ nach den Vektoren $\tilde{\boldsymbol{f}}_k$ entwickelt, wodurch die Norm von $\tilde{\boldsymbol{\varepsilon}}$ minimiert wird. Für diesen Zweck wird die Basis $\{\tilde{\boldsymbol{f}}_k\}$ mittels des Gram-Schmidt-Verfahrens orthogonalisiert. Die Entwicklungskoeffizienten können dann unmittelbar durch Skalarprodukte mit den orthogonalen Basisvektoren berechnet

werden. Abschließend müssen die Koeffizienten noch in die ursprüngliche nichtorthogonale Basis transformiert werden. Nach der Berechnung der zeitvariablen Trajektorie $r(n)$ können die Ausgangssignale $v(n)$ und $w(n)$ berechnet werden, die die Eingangssignale $o(n)$ und $u'(n)$ bzw. $u(n)$ des nächsten Kreuzgliedes repräsentieren. Damit können sukzessive alle Kreuzglieder berechnet werden.

3 Analyse von Sprachsignalen

Für die Analyse von Sprachsignalen werden diese zuvor mit einer adaptiven Präemphase vorgefiltert, wodurch der Einfluss der Anregung und Abstrahlung auf die spektrale Hülle weitgehend eliminiert wird. In Abb. 2 sind Analyseergebnisse der Äußerung „audio“ gezeigt, die mit einer Abtastrate von 16 kHz aufgenommen wurde. In Abb. 2(a) sind die Schätzergebnisse einer rein zeitvariablen Analyse präsentiert, während in Abb. 2(b) die einer

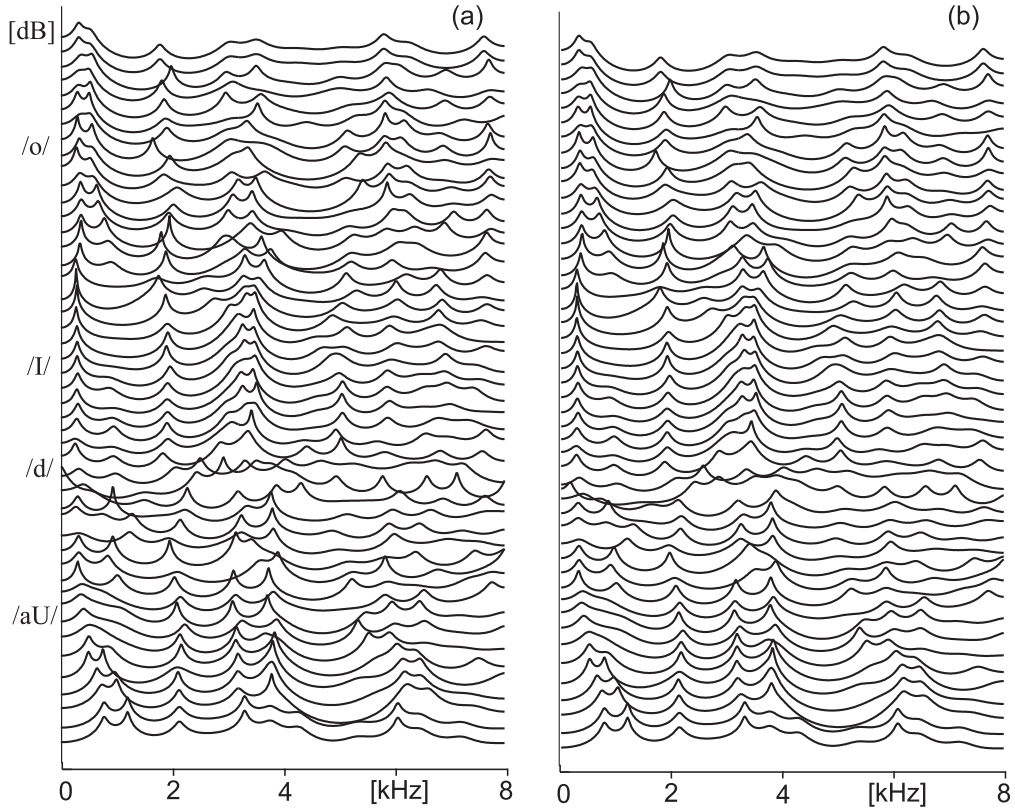

Abbildung 2 - Analyse der Äußerung „audio“: (a) rein zeitvariable Analyse; (b) kombinierte zeitvariable und zeitinvariante Analyse.

kombinierten zeitvariablen und zeitinvarianten Analyse gezeigt sind. Es sind jeweils die Betragsgänge im Abstand von 200 Abtastwerten (12,5 ms) an den Positionen m_i gezeigt. Es ist zu erkennen, dass die kombinierte Analyse einen glatteren Verlauf aufweist und dennoch im Bereich des stimmhaften Explosivs auch eine hohe Zeitauflösung erzielt. Die Länge L_a der linearen Abschnitte der Trajektorie betragen 200 Abtastwerte. Die Länge L_{ti} der Segmente der zeitinvarianten Schätzkomponenten beträgt 400 Abtastwerte (25 ms). Werden mit L_a die Bereiche der linearen Abschnitte für die Analyse länger gewählt, so ergeben sich zwischen der rein zeitvariablen und der kombinierten Analyse geringere Unterschiede.

4 Zusammenfassung

In diesem Beitrag wird ein nichtiterativer Prädiktionsalgorithmus vorgestellt, der auf zeitvariablen und zeitinvarianten Modellschätzungen basiert. Die optimale Lösung kann hierfür analytisch gewonnen werden. Die Analysen von Sprachsignalen zeigen auf, dass durch die Einbeziehung von zeitinvarianten Schätzkomponenten insbesondere für Parameter-Trajektorien mit kurzen linearen Abschnitten glattere Trajektorien erzielt werden können, die für Anwendungen der Sprachanalyse und -synthese vorteilhaft sind.

Literatur

[1] Makhoul, J.: "Linear Prediction: A Tutorial Review", in Proc. IEEE, vol. 63, pp. 561–580, Apr. 1975.

[2] Markel, J.; Gray, A.: Linear Prediction of Speech. New York: Springer-Verlag, 1976.

[3] Burg, J.: "A New Analysis Technique for Time Series Data", NATO Advanced Study Inst. on Signal Processing, Enschede, 1968.

[4] Haykin, S.: Adaptive Filter Theory. New Jersey: Prentice-Hall, Inc., 3 ed., 1996.

[5] Malladi, K. M.; Rajakumar, R. V.: "Estimation of Time-Varying AR Models of Speech through Gauss-Markov Modeling", Proc. IEEE Conf. ICASSP, Hong Kong, pp. 305–308, 2003.

[6] Deng, Li; Lee, K. J.; Attias, H.; Acero A.: "Adaptive Kalman Filtering and Smoothing for Tracking Vocal Tract Resonances Using a Continuous-Valued Hidden Dynamic Model", IEEE Trans. ASSP-15, Issue 1, 2007 pp.13-23.

[7] Vargas, J.; McLaughlin, S.: "Cascade prediction filters with adaptive zeros to track the time-varying resonances of the vocal tract", IEEE Trans. ASSP-16, no. 1, pp. 1–7, Jan. 2008.

[8] Subba Rao, T.: "The Fitting of Non-stationary Time-series Models with Time-dependent Parameters," in J. Roy. Statist. Soc. Series B, vol. 32, no. 2, pp. 312-322, 1970.

[9] Grenier, Y.: "Time-Dependent ARMA Modeling of Non-stationary Signals", IEEE Trans. ASSP-31, no. 4, pp. 899–911, August 1983.

[10] Schnell, K.; Lacroix, A.: "Time-Varying Linear Prediction for Speech Analysis and Synthesis", Proc. IEEE Conf. ICASSP, Las Vegas, pp. 3941-3944, 2008.

[11] Schnell, K.: "Time-Varying Burg Method for Speech Analysis", Proc. EURASIP Conf. EUSIPCO, Lausanne Switzerland, 2008.

[12] Schnell, K.; Lacroix, A.: "Model-Based Analysis of Speech and Audio Signals for Real-Time Processing Based on Time-Varying Lattice Filters", Proc. IEEE Conf. ICASSP, Taipei Taiwan, pp. 3973-3976, 2009.

[13] Schnell, K.; Lacroix, A.: "Iterative Inverse Filtering by Lattice Filters for Time-Varying Analysis and Synthesis of Speech", Proc. IEEE Conf. ICASSP, Taipei Taiwan, pp. 4017-4020, 2009.