Ein Verfahren zur automatischen Abstimmung von Einseitenband-Empfängern

Lida Dastager, Jürgen Weith
Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt

Es wird ein Verfahren beschrieben, mit dessen Hilfe eine Korrektur der Demodulation von Einseitenband-Signalen (ESB-Signalen) vorgenommen werden kann. Eine solche Korrektur ist dann oft notwendig, wenn sich die Frequenzen, die zur Modulation und zur Demodulation verwendet worden sind, unterscheiden. Typischerweise tritt ein solcher Frequenzunterschied bei der Übertragung von Sprachsignalen auf, wenn kein festes Kanalraster zugrunde liegt. Das Korrekturverfahren basiert auf einer Analyse stimmhafter Sprachabschnitte mit dem MUSIC-Verfahren.

Kategorie und Themenbeschreibung: Elektrische Nachrichtenübertragung

Zusätzliche Schlüsselwörter: Spektralanalyse mit Subspace-Verfahren, Super-Resolution-Analyse, MUSIC-Verfahren, Sprachqualität

1 EINFÜHRUNG

Zur Übertragung von Sprachsignalen über hochfrequente Kanäle wird in bestimmten Frequenzbereichen oft die Einseitenband-Modulation angewandt. Dabei wird – wie der Name sagt – lediglich ein Seitenband übertragen und das andere unterdrückt . Zusätzlich wird meist die Trägerschwingung ausgesiebt, um die Sendeleistung zu reduzieren. Man spricht dann von einem SSBSC-Signal (SSBSC = Single Side Band Subpressed Carrier).

Ein solches SSBSC-Signal besitzt aus übertragungstechnischer Sicht verschiedene Vorteile wie z. B.

- einen geringen hochfrequenten Bandbreitenbedarf (die hochfrequente Bandbreite ist hier gleich der niederfrequenten Signal-Bandbreite),
- eine hohe spektrale Effizienz,
- lineare Verzerrungen auf dem Übertragungsweg führen lediglich zu linearen Verzerrungen im demodulierten Signal (andere Modulationsverfahren erzeugen nichtlineare Verzerrungen, welche die Signalqualität gravierend verschlechtern können),
- ein SSBSC-Signal kann mit Hilfe eines digitalen Signalprozessors (DSP) in einfacher Weise erzeugt werden. Zu realisieren ist dabei primär ein System, das die Hilbert-Transformation berechnet.

Adresse des Autors:
Lida Dastager, Jürgen Weith, Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Ignaz-Schön-Str. 11, 97421 Schweinfurt
Lida.Dastager@gmx.de, juergen.weith@fhws.de

Der SSBSC-Modulation im Sender, bei der die Trägerfrequenz f_M eingesetzt wird, schließt sich nach der Übertragung über den Kanal im Empfänger die SSBSC-Demodulation mit der Frequenz f_D an. Für eine fehlerfreie Übertragung müssen dabei Sender und Empfänger die gleiche Frequenz verwenden, es muss also $f_D = f_M$ sein.

Die Einhaltung der Bedingung $f_D = f_M$ setzt voraus, dass sich Sender und Empfänger über die zu verwendende Frequenz verständigt haben. Das ist aber aus praktischen Gründen oft nicht möglich. Zum Auffinden der Frequenz f_D wird daher die demodulierende Frequenz solange verändert, bis das zurückgewonnene Sprachsignal „möglichst natürlich" klingt.

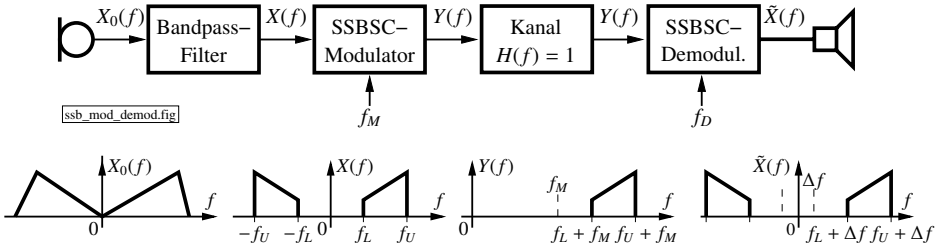

Abbildung 1: Skizze einer SSBSC-Übertragung zusammen mit den auftretenden Spektren der Teilsignale.

Ein solches Vorgehen bringt natürlich Nachteile mit sich. Zunächst kann eine Abstimmung „nach Klang" erfahrungsgemäß recht zeitraubend sein, weil immer wieder kleine Frequenzänderungen vorgenommen werden müssen, um den natürlichsten Klangeindruck aufzufinden. Darüber hinaus ist eine derartige Frequenzwahl subjektiv gefärbt, unterschiedliche Personen werden i. a. auch unterschiedliche Demodulationsfrequenzen bevorzugen.

Es sind deshalb Verfahren interessant, welche aus dem demodulierten niederfrequenten Signal die Größe $\Delta f = f_M - f_D$ zu bestimmen vermögen. Ist Δf bekannt, dann kann im Niederfrequenzbereich eine Korrektur des Sprachsignals erfolgen.
Ein Vorschlag zur Offset-Korrektur findet sich z. B. in [5]. Das nachfolgend beschriebene Verfahren hat (im Gegensatz zum Verfahren in der genannten Literatur) zum Ziel, eine Implementierung auf einem DSP, z. B. ausgehend von einem SIMULINk/MATLAB-Modell zu ermöglichen.

2 ESB-MODULATION UND DEMODULATION

Zu übertragen sei ein reelles niederfrequentes Signal $x(t)$, das durch Bandpass-Filterung aus einem Signal $x_0(t)$ entstanden ist. Für das Spektrum $X(f)$ gelte $|X(f)| \equiv 0$ für $|f| > f_U$. Aus $x(t)$ entsteht das Einseitenband-Signal $y(t)$ gemäß

$$y(t) = \mathrm{Re}\{[x(t) \pm j\hat{x}(t)]\, e^{j2\pi f_M t}\} \qquad (1)$$

$$= x(t)\cos(2\pi f_M t) \mp \hat{x}(t)\sin(2\pi f_M t) \qquad (2)$$

$\hat{x}(t)$ bedeutet die Hilbert-Transformierte der Zeitfunktion $x(t)$). Wird das positive Vorzeichen in Glg. (1) verwendet, so liegt ein USB-Signal (USB = Upper Side Band) vor und das obere Seitenband wird übertragen, das untere Seitenband dagegen unterdrückt. Im anderen Fall wird ein LSB-Signal (LSB=Lower Side Band) generiert. Die Gleichungen beschreiben die spektrale Verschiebung des in eckigen Klammern stehenden analytischen Signals um die Modulationsfrequenz f_M.

Die Demodulation erfolgt durch spektrale Rückverschiebung. Entsprechend erhält man das demodulierte Signal als

$$\tilde{x}(t) = 2\,[y(t) \cdot \cos(2\pi f_D t)] * h_{TP}(t). \qquad (3)$$

$h_{TP}(t)$ bezeichnet dabei die Impulsantwort eines Tiefpasses mit der Übertragungsfunktion

$$H_{TP}(f) = \begin{cases} 1 & \text{für} \quad |f| \le f_U + f_M - f_D \\ 0 & \text{für} \quad |f| \ge f_L + f_M + f_D. \end{cases} \qquad (4)$$

Durch dieses dem Multiplizierer (technisch: Mischer) nachzuschaltende Tiefpass-Filter werden solche Frequenzanteile unterdrückt, welche im Bereich $f_L + f_M + f_D\, f \le f_U + f_M + f_D$ liegen[1]. Aus den Gleichungen (1) und (3) ergibt sich schließlich für das demodulierte niederfrequente Signal

$$\tilde{x}(t) = \mathrm{Re}\{[x(t) + j\hat{x}(t)]\, e^{j2\pi\Delta f t}\} \qquad (5)$$

$$= x(t)\cos(2\pi\Delta f t) - \hat{x}(t)\sin(2\pi\Delta f t). \qquad (6)$$

Ein Diagramm des Modulations- und Demodulationsvorgangs zusammen mit den auftretenden Spektren ist in Abb. 1 gezeichnet. Für den Kanal wurde die Übertragungsfunktion $H(f) = 1$ („Durchverbindung“) angenommen.

3 KORREKTUR DES FREQUENZ-OFFSETS

Das mit unterschiedlichen Frequenzen f_M bzw. f_D modulierte bzw. demodulierte SSBSC-Signal soll nun korrigiert werden. Eine derartige Korrektur wird dadurch erreicht, dass das zum Empfangssignal gehörende analytische Signal bestimmt und anschließend um $f = -\Delta f$ spektral verschoben wird. Abschließend $\tilde{x}(t) + j\hat{\tilde{x}}(t)$ ist noch der Realteil zu bilden. Es gilt dann für die Vorschrift der Frequenzoffset-Korrektur:

$$x(t) = \mathrm{Re}\{[\tilde{x}(t) + j\hat{\tilde{x}}(t)]e^{-j2\pi\Delta f t}\}. \qquad (7)$$

Damit liegt wieder das im Sender durch Bandpass-Filterung gebildete, dem SSBSC-Modulator zugeführte Signal $x(t)$ vor.
Auf die Bestimmung Wertes Δf wird nachfolgend näher eingegangen.

4 SCHÄTZUNG DES FREQUENZOFFSTES

Bei den meisten Übertragungssystemen, die SSBSC-Modulation einsetzen, sind die Frequenzen f_L und f_U nicht oder nicht genau genug bekannt. So werden bei der Sprachübertragung oft f_L = 330 Hz und f_H = 3,4 kHz verwendet. Ob diese Frequenzen aber bei einem aktuell vorliegenden Signal tatsächlich gültig sind, ist nicht bekannt. Eine Bestimmung des Frequenz-Offsets aus diesen Größen scheidet deshalb meist aus.

Eine Bestimmung von Δf ist jedoch auf anderem Weg möglich. Dazu wird ein Ausschnitt (Rahmen, Frame) aus dem Sprachsignal $x_0(t)$ betrachtet, der stimmhaften Charakter hat. Ein solcher stimmhafter Abschnitt wird durch Stimmbandschwingungen erzeugt [1]. Es gilt in diesem Fall

[1]Auf den Tiefpass kann verzichtet werden, wenn alternativ das zum Signal y(t) analytische Signal gebildet und dieses spektral verschoben wird. Hier liegt der Aufwand im System zur Bildung der Hilbert-Transformierten.
[2]Unterstellt wird ein ideales Bandpass-Filter, das im Übertragungsband weder Amplitudenfehler noch Phasenänderungen bewirkt.

$$x_0(t) = \sum_{\mu=1}^{n} a_\mu e^{j\mu 2\pi f_p t}, \quad a_\mu \in \mathbb{C}, \qquad (8)$$

wobei f_p die sog. Sprachgrundfrequenz („Pitch-Frequenz“) darstellt. Nach Bandpass-Filterung[2] ergibt sich

$$x(t) = \sum_{\mu=n_1}^{n_2} a_\mu e^{j\mu 2\pi f_p t}, \quad a_\mu \in \mathbb{C} \qquad (9)$$

und am Ausgang des Demodulators erscheint schließlich das Signal

$$\tilde{x}(t) = \sum_{\mu=n_1}^{n_2} a_\mu e^{j2\pi(\mu f_p + \Delta f)t}, \quad a_\mu \in \mathbb{C}. \qquad (10)$$

Eine Analyse der in $\tilde{x}(t)$ vorkommenden Frequenzen f_α liefert

$$f_\alpha = (\alpha - 1 + n_1) f_p + \Delta f, \quad \alpha = 1(1) n_2 - n_1 + 1. \qquad (11)$$

Damit lässt sich zunächst aus den Frequenzen f_α die Grundfrequenz f_p bestimmen (z. B. durch Differenzbildung benachbarter Frequenzen und Ausschluss von Mehrfachen der Grundfrequenz). Dann gilt für den gesuchten Frequenz-Offset

$$\Delta f = f_1 - \lambda f_p, \quad \lambda \in \mathbb{N} \qquad (12)$$

mit einer ganzzahligen, positiven und noch unbekannten Zahl λ . Es besteht also zunächst eine Mehrdeutigkeit bei der Bestimmung von Δf .

Diese Mehrdeutigkeit lässt sich auflösen, wenn weitere $q - 1$ Bestimmungen der Frequenzkomponenten vorgenommen werden, *wobei sich aber die Grundfrequenzen voneinander unterscheiden müssen*. Man hat jetzt

$$\Delta f = f_1^{(\nu)} - \lambda^{(\nu)} f_p^{(\nu)}, \quad \nu = 1(1) q, \quad \lambda \in \mathbb{N} \qquad (13)$$

(ν bedeutet die Nummer der Messung). Der gesuchte Offset ergibt sich daraus als

$$\Delta f = \arg\min_f \sum_{\nu=1}^{q} \left| \min_{\lambda^{(\nu)} \in \mathbf{N}} \{ f - f_1^{(\nu)} + \lambda^{(\nu)} f_p^{(\nu)} \} \right|. \qquad (14)$$

Damit ist der gesuchte Frequenz-Offset bekannt und es kann eine entsprechende Korrektur erfolgen.

5 BESTIMMUNG DER FREQUENZEN DER SPEKTRALKOMPONENTEN

Zur Bestimmung von Δf gemäß Glg. (14) werden die Werte $f_p^{(\nu)}$ und $f_1^{(\nu)}$ benötigt. $f_p^{(\nu)}$ kann als Differenz zweier unmittelbar benachbarter Frequenzen aus Glg. (11) gewonnen werden:

$$f_p^{(\nu)} = f_{\alpha_0+1}^{(\nu)} - f_{\alpha_0}^{(\nu)}. \qquad (15)$$

Bei gestörten Signalen kann auch an eine Mittelung benachbarter Frequenzen gedacht werden ($\alpha_0 = 1, 2, \cdots$).

Da sich die Frequenz f_p von Messung zu Messung (über mehrere Rahmen hinweg) zum Teil nur wenig ändert, ist eine möglichst genaue Bestimmung der Frequenzen f der stimmhaften Sprachabschnitte notwendig. Verschiedene Versuche führten zum Ergebnis, dass hierzu die Anwendung des MUSIC-Verfahrens[3] am besten geeignet ist.

[3]MUSIC = Multiple Signal Classification

Beim MUSIC-Verfahren handelt es sich um eine parametrische Methode zur Bestimmung der Frequenzen eines sich aus harmonischen Komponenten zusammensetzenden Signals $y(t)$. Details dazu finden sich in [2]. Verwendet werden dazu die Abtastwerte $y[k] = y(kT_A)$. Die Anzahl der Teilkomponenten muss dabei bekannt sein. Eine Beschreibung des Verfahrens findet sich z. B. in [3]. Das beim MUSIC-Verfahren unterstellte Signalmodell hat die Form

$$y[k] = \sum_{\mu=1}^{N} b_\mu e^{j\Omega_\mu k} + n[k], \qquad (16)$$

wobei die b_μ komplexe Koeffzienten und $n[k]$ eine Störung mit konstanter Leistungsdichte, gaußförmiger Verteilungsdichtefunktion und der Varianz σ^2 darstellen. Aus der Folge $y[k]$ der Abtastwerte wird die Kovarianzmatrix $\boldsymbol{R} = E\{\boldsymbol{y}\boldsymbol{y}^{\mathrm{H}}\}$ mit $\boldsymbol{y} = [\boldsymbol{y}[0]\ \boldsymbol{y}[1]\cdots\boldsymbol{y}[\mathrm{M}-1]]^{\mathrm{T}}$ gewonnen. Die hermitesche Matrix $\boldsymbol{R}$ besitzt M Eigenwerte, die positiv sind und von denen $M - N$ den Wert σ^2 aufweisen. N Eigenwerte sind größer als σ^2. Man bestimmt nun die $M - N$ zu den Eigenwerten σ^2 gehörenden Eigenvektoren $u_{N+1}, u_{N+2}, \cdots, u^M$ und fasst diese Spaltenvektoren in der Matrix $\boldsymbol{U}$ zusammen. Mit dem Vektor $\boldsymbol{a}(\Omega) = [1\ e^{-j\Omega}\ \ e^{-j2\Omega} \cdots e^{-j(M-1)\Omega}]^T$ gilt dann die Aussage: *Die gesuchten N Frequenzen sind Lösungen der Gleichung*

$$\boldsymbol{a}^H(\Omega)\boldsymbol{U}\boldsymbol{U}^H\boldsymbol{a}(\Omega) = 0. \qquad (17)$$

In der Praxis wird man die N kleinsten lokalen Minima dieser Funktion suchen, indem man die Frequenz Ω in kleinen Schritten von Null an stetig vergrößert. Diejenigen Frequenzen, bei denen $\boldsymbol{a}^H(\Omega)\boldsymbol{U}\boldsymbol{U}^H\boldsymbol{a}(\Omega)$ jeweils den Wert Null erreicht, sind dann die gesuchten Werte Ω_μ.

Die Bestimmung der Frequenzen Ω_μ kann manchmal einfacher durch Anwendung des sogenannten Root-MUSIC-Verfahrens erfolgen. Dazu definiert man zunächst die Vektoren $\boldsymbol{z}^1 = [1\ \ z\ \ z^2\ \ \ z^{M-1}]^T$ sowie $\boldsymbol{z}^2 = [1\ \ z^{-1}\ \ z^{-2}\ \ \ z^{-(M-1)}]^T$ und setzt anschließend beide in die Glg. (17) ein. Dann stellt $z^{M-1} \cdot z^{\mathrm{T}}_1\, \boldsymbol{U}\boldsymbol{U}^{\boldsymbol{H}}\, \boldsymbol{z}_2$ ein Polynom $P(z)$ vom Grad $2M–2$ dar. Von dessen $2M–2$ Nullstellen $z_{0,\lambda}$ liegen (im Idealfall) $2N$ Nullstellen auf dem Einheitskreis (für sie gilt $z_{0,\lambda} = e^{j\Omega_\mu}$). Diese Nullstellen – und damit die interessierenden Frequenzen – müssen nun durch ein geeignetes Suchprogramm gefunden werden.

6 ANMERKUNG ZUR PRAKTISCHEN AUSLEGUNG

Das hier beschriebene Korrektur-Verfahren wurde in Form einer Simulation mit Hilfe der Software OCTAVE untersucht. Dazu wurden zunächst SSBSC-Sprachsignale aus dem Kurzwellenbereich demoduliert. Es folgte eine Abtastung mit der Rate $f_A = 8$ kHz bei einer Wortbreite von 16 Bit.

Das Korrekturverfahren verwendet eine Rahmenlänge von 62,5 ms, was 500 Abtastwerten entspricht. Zur Bestimmung der Kovarianzmatrix $\boldsymbol{R}$ in Form einer Näherung $\hat{\boldsymbol{R}}$ wurden aus den 500 Abtastwerten 349 Vektoren mit jeweils 160 Elementen gebildet, die um jeweils einen Abtastwert gegeneinander verschoben waren. Durch „Subspace Embedding“ wurde die Kovarianz-Näherung mit 160×160 Elementen gebildet.

Zur Unterscheidung von stimmhaften und (für die Korrektur ungeeigneten) stimmlosen Segmenten wurde ein eigenes Verfahren angewandt, über das an dieser Stelle nicht berichtet werden soll. Zusammenfassend ist festzustellen, dass das beschriebene Verfahren eine Frequenzkorrektur liefert, die dem subjektivem Empfinden nach eine natürliche und angenehme Klangfarbe liefert. Dabei gelingt eine Korrektur auch bei stärker verrauschten Sprachsignalen. Über objektive Messergebnisse kann momentan noch nicht berichtet werden. Sie sind Gegenstand gegenwärtiger Untersuchungen.

LITERATUR

[1] W. Hess P. Vary, U. Heute. Digitale Sprachverarbeitung. Teubner-Verlag, 1998.
[2] Harry L. Van Trees. Detection, Estimation, and Modulation Theory, Optimum Array Processing (Part IV). Wiley-Interscience, 2002.
[3] Sophocles J. Orfanidis. Optimum Signal Processing. McGraw-Hill International Edition, 1990.
[4] A. Jansen P. Clark, S. H. Mallidi and H. Hermansky. Frequency oset correction in speech without detecting pitch. IEEE ICASSP, May 2013.
[5] H. J. Kolb T. Gülzow, U. Heute. ssb carrier mismatch detection from speech characteristics: Extensions beyond the range of uniqueness. EUSIPCO Vol 1, 2002.
[6] H. Voelcker. Demodulation of single side-band signals via envelope detection. IEEE Trans. Communication Technology, Feb. 1966.