نشریه علمی- پژوهشی، انجمن مهندسان کنترل و ابزار دقیق ایران

جلد ۳، شماره ۱، بهار ۱۳۸۸

مقالات بخش فارسی

بررسی یادگیری تقویتی و خواص سیاست بهینه در مسائل جدولی با استفاده از روش های کنترل دیجیتال ۱
سیدمصطفی کلامی هریس، ناصر پریز، محمدباقر نقیبی سیستانی

طراحی موجک بهینه برای نویززدایی از طیف انرژی رادیوایزوتوپ های گسیلنده پرتو گاما جهت استخراج برهمکنش های غالب ۱۲
علی فیاضی، حسین احمدی نوبری

کاهش زمان نشست در کنترل فیدبک تاخیری سیستم آشوبی روسلر با سنکرون کننده PI ۲۳
حسن فاتحی مرج، رجب اصغریان، ناصر پریز

بکارگیری روش میانگین گیری مرتب وزندار (OWA) در ترکیب داده های ربات مین یاب ۲۹
محمدرضا بادلو، بهزاد مشیری، بابک نجاراعرابی

ارائه یک سیستم آشوبناک بعد بالای جدید همراه با یک نقطه تعادل و پایدارسازی آن با استفاده از کنترل کننده فیدبک حالت خطی ۳۷
علی ابویی، محمدرضا جاهد مطلق، زهرا رحمانی چراتی

مقالات بخش انگلیسی

A Combined DC-Filter and Optimized Modulation to Absorb DC-Link Oscillations of Cascaded H-Bridge Converters 1
Mohammad Tavakoli Bina, Bahman Eskandari

Utilizing Multi-Parametric Programming for Optimal Control of Linear Constrained Systems 12
Farhad Bayat, Ali A. Jalali, Mohammad R. Jahed Motlagh

www.isice.ir

مجله کنترل

نشریه علمی- پژوهشی، انجمن مهندسان کنترل و ابزار دقیق ایران، جلد ۳، شماره ۱، بهار ۱۳۸۸

صاحب امتیاز: انجمن مهندسان کنترل و ابزار دقیق ایران
مدیر مسئول: پروفسور ایرج گودرزنیا
سردبیر: پروفسور علی خاکی صدیق
تلفن: ۸۴۰۶۲۳۱۷
ایمیل: sedigh@kntu.ac.ir
آدرس محل کار: خیابان دکتر شریعتی، پل سیدخندان، دانشکده برق دانشگاه صنعتی خواجه نصیرالدین طوسی
سمت: استاد دانشگاه
شورای سردبیری: پروفسور علی خاکی صدیق، دکتر حمید خالوزاده، دکتر علیرضا فاتحی
دبیر اجرایی: دکتر حمید خالوزاده

هیأت تحریریه:

پروفسور علی خاکی صدیق (استاد)، دکتر حمید خالوزاده (دانشیار)، پروفسور ایرج گودرزنیا (استاد)، پروفسور پرویز جبه دار مارالانی (استاد)، پروفسور علی غفاری (استاد)، دکترهوشنگ حسیبی (استادیار)، دکتر محمدرضا جاهد مطلق (دانشیار)، دکتر کامبیز بدیع (دانشیار)، پروفسور رجب اصغریان (استاد)، دکتر حمیدرضا مومنی (دانشیار)، دکتر کریم صفوی (دانشیار)، پروفسور علی وحیدیان کامیاد (استاد)، پروفسور سهراب خانمحمدی (استاد)، پروفسور سید کمال الدین نیکروش (استاد)، پروفسور مسعود شفیعی (استاد)، دکتر بتول لبیبی(استادیار)، دکتر علیرضا فاتحی(استادیار)، پروفسور بهزاد مشیری (استاد)، پروفسور سراج الدین کاتبی (استاد)، مهندس امیر مقصودی پور (کارشناس ارشد)، مهندس کاووس فلامکی(کارشناس ارشد)، مهندس مهدی برادران مظفری(کارشناس ارشد)، مهندس غلامعباس رمضانی(کارشناس ارشد).

هیأت مشاوران:

دکتر حمیدرضا مومنی، پروفسور علی غفاری، دکتر علی اکبر قره ویسی، دکتر محمد توکلی بینا، دکتر حمیدرضا تقی راد، دکتر محمد بطحایی، دکتر محمدتقی بهشتی، پروفسور بهزاد مشیری، پروفسور مسعود شفیعی، پروفسور رجب اصغریان، پروفسور علی خاکی صدیق، دکتر رضا کاظمی، دکتر سید علی اکبر موسویان، دکتر امیرحسین مرکزی دوایی، پروفسور محمد حایری، دکتر علیرضا خلیلی تهرانی، پروفسور حسین سیفی، دکتر احد کاظمی، دکتر حمید خالوزاده، دکتر علیرضا فاتحی، دکتر محمدرضا اکبرزاده توتونچی، دکتر میرعابدینی، دکتر حسین پدرام، دکتر علی هارون آبادی، پروفسور علی وحیدیان کامیاد، دکتر جعفر حیرانی نوبری، دکتر بتول لبیبی، پروفسور فرامرز حسین بابایی، دکتر پاکنوش کریم آقایی، دکتر بیژن معاونی، دکتر مهدی علیاری شوره دلی، دکتر محمد عاروان

هیأت مدیره انجمن مهندسان کنترل و ابزار دقیق:

مهندس عباس شعری مقدم، دکتر کیوان مسروری، دکتر حمیدرضا مومنی، پروفسور بهزاد مشیری، فرزاد جعفری کاظمی، دکتر حمید خالوزاده، مهندس علیرضا رستگاری

تهران، خیابان انقلاب، میدان فردوسی، خیابان عباس موسوی (فرصت)، پلاک ۷۱، طبقه دوم، اتاق ۲۴۱

صندوق پستی: ۱۵۸۱۵-۳۵۹۵ تلفن: ۸۸۸۱۳۰۰۲ فاکس: ۸۸۳۲۴۹۷۹

http://www.isice.ir

به نام خدا

مقالات بخش فارسی

بررسی یادگیری تقویتی و خواص سیاست بهینه در مسائل جدولی با استفاده از روش‌های کنترل دیجیتال ۱
سید مصطفی کلامی هریس، ناصر پریز، محمدباقر نقیبی سیستانی

طراحی موجک بهینه برای نویززدایی از طیف انرژی رادیوایزوتوپ‌های گسیلنده پرتوگاما جهت استخراج برهمکنش‌های غالب ۱۲
علی فیاضی، حسین احمدی نوبری

کاهش زمان نشست در کنترل فیدبک تاخیری سیستم آشوبی روسلر با سنکرون کننده PI ۲۳
حسن فاتحی مرج، رجب اصغریان، ناصر پریز

بکارگیری روش میانگین گیری مرتب وزندار(OWA) در ترکیب داده های ربات مین یاب ۲۹
محمد رضا بادلّو، بهزاد مشیری، بابک نجار اعرابی

ارائه یک سیستم آشوبناک بعد بالای جدید همراه با یک نقطه تعادل و پایدارسازی آن با استفاده از کنترل‌کننده فیدبک حالت خطی ۳۷
علی ابویی، محمّدرضا جاهدمطلق، زهرا رحمانی چراتی

مقالات بخش انگلیسی

A Combined DC–Filter and optimized Modulation to Absorb DC–Link Oscillations of Cascaded H-Bridge Converters 1
M. Tavakoli Bina, B. Eskandari

Utilizing Multi-Parametric Programming for Optimal Control of Linear Constrained Systems 12
Farhad Bayat, Ali A. Jalali, Mohamad R. Jahed Motlagh

مجله کنترل، مجله ای علمی – پژوهشی است که در برگیرنده تازه ترین نتایج تحقیقات نظری و کاربردی در علوم مختلف مرتبط با مهندسی کنترل و ابزار دقیق میباشد. مقالات ارسالی به مجله کنترل می توانند به زبان فارسی وظ یا انگلیسی باشند. از میان مباحث مورد نظر این مجله میتوان به موارد زیر اشاره نمود:

۱) مدلسازی، شناسایی، شبیه سازی و بهینه سازی سیستمها.
۲) تحلیل و طراحی سیستم های کنترل پیشرفته همچون سیستم های کنترل خطی و غیرخطی، سیستمهای کنترل تطبیقی، کنترل مقاوم و کنترل بهینه، سیستمهای کنترل هوشمند، سیستمهای کنترل تصادفی، سیستمهای کنترل گسسته پیشامد و ترکیبی، سیستمهای ابعاد وسیع.
۳) مکاترونیک و رباتیک.
۴) ابزار دقیق و سیستمهای ترکیب داده و اطلاعات سنسوری.
۵) اتوماسیون صنعتی همچون سیستمهای کنترل گسترده، رابط انسان – ماشین، سیستمهای ایمنی و تشخیص خطا، سیستمهای زمان حقیقی و سیستمهای کنترل سوپروایزری.

کاربردهای مورد علاقه این مجله، وسیع بوده و می تواند در برگیرنده موارد زیر با شد:

۱) سیستمهای هدایت و ناوبری
۲) فرآیندهای صنعتی شامل فرآیندهای شیمیایی و بیوتکنولوژی.
۳) فرآیندهای استخراج و فراوری موارد معدنی.
۴) سیستمهای حمل و نقل و خودروهای خودکار.
۵) تولید و توزیع نیروی برق.
۶) مهندسی محیط زیست و هواشناسی.
۷) مهندسی تکنولوژی تولید.
۸) سیستمهای اقتصادی و مالی.
۹) سیستمهای اطلاعاتی، مخابراتی و شبکه های صنعتی.
۱۰) مهندسی پزشکی.
۱۱) سیستمهای آموزش هوشمند.

از کلیه پژوهشگران و کارشناسان فعال در زمینه های مرتبط با مهندسی کنترل و ابزار دقیق دعوت بعمل می آید تا مقالات و نتایج آخرین دستاوردهای علمی و پژوهشی خود را به این مجله ارسال نمایند. خواهشمند است مقالات خود را به صورت الکترونیکی به آدرس control@isice.ir ارسال فرمایید. برای کسب اطلاعات بیشتر و دریافت نحوه تهیه و ارسال مقالات می توانید به سایت مجله با آدرس www.isice.ir مراجعه نمایید.

مجله کنترل
جلد ۳، شماره ۱، بهار ۱۳۸۸، صفحه ۱۱-۱

بررسی یادگیری تقویتی و خواص سیاست بهینه در مسائل جدولی با استفاده از روش‌های کنترل دیجیتال

سید مصطفی کلامی هریس [۱]، ناصر پریز [۲]، محمدباقر نقیبی سیستانی [۳]

[۱] فارغ التحصیل کارشناسی ارشد برق- کنترل، دانشگاه فردوسی مشهد، دانشکده‌ی مهندسی، گروه مهندسی برق، sm.kalami@gmail.com

[۲] دانشیار، دانشگاه فردوسی مشهد، دانشکده‌ی مهندسی، گروه مهندسی برق، گرایش کنترل، n-pariz@um.ac.ir

[۳] استادیار، دانشگاه فردوسی مشهد، دانشکده‌ی مهندسی، گروه مهندسی برق، گرایش کنترل، mb-naghibi@um.ac.ir

چکیده: فرآیند تصمیم‌گیری مارکوف یا MDP، یکی از مسائلی است که دارای کاربردهای وسیعی در زمینه‌های مختلف علمی، مهندسی، اقتصادی و مدیریت است. بسیاری از فرآیندهای تصمیم‌گیری، دارای خاصیت مارکوف می‌باشند و به صورت یک مسأله‌ی تصمیم‌گیری مارکوف قابل بیان هستند. یادگیری تقویتی یکی از رویکردهایی است که برای حل MDP به کار می‌رود، و به نوبه‌ی خود از برنامه‌ریزی پویا یا DP استفاده می‌کند. در این نوشتار الگوریتم ارزیابی سیاست، که در بحث یادگیری تقویتی و DP برای حل MDP به کار می‌رود، به صورت معادله‌ی دینامیکی یک سیستم دیجیتال یا گسسته-زمان بازنویسی شده است. به این ترتیب این امکان به وجود آمده است که بتوان با بهره‌گیری از روش‌های موجود در کنترل دیجیتال، به بررسی خواص معادلات به دست آمده پرداخت و تحلیل مناسبی از رفتار عامل یادگیرنده، تحت سیاست‌های مختلف، به عمل آورد. روش مذکور برای تحلیل دو مسأله‌ی جدولی استفاده شده است و سپس نتایج کلی در خصوص مسائل جدولی بیان و اثبات شده‌اند. به عنوان مثال، نتایج به دست آمده نشان می‌دهند که سیاست بهینه برای هر مسأله‌ی جدولی، در چارچوب کنترل دیجیتال، به صورت یک سیستم مرده نَوِش یا Dead Beat قابل توصیف است.

کلمات کلیدی: برنامه‌ریزی پویا، سیستم‌های کنترل دیجیتال، فرآیندهای تصمیم‌گیری مارکوف، کنترل تصادفی، یادگیری تقویتی.

Abstract: Markov Decision Process (MDP) has enormous applications in science, engineering, economics and management. Most of decision processes have Markov property and can be modeled as MDP. Reinforcement Learning (RL) is an approach to deal with Markov Decision Processes. RL methods are based on Dynamic Programming (DP) algorithms, such as Policy Evaluation, Policy Iteration and Value Iteration. In this paper, policy evaluation algorithm is represented in the form of a discrete-time dynamical system, namely a Discrete-Time Control system. Hence, using Discrete-Time Control methods, behavior of agent and properties of various policies, can be analyzed. Two grid-world problems are solved and analyzed using this approach. Therefore general case of grid-world problems is addressed, and some important results are obtained for this type of problems, For example, equivalent dynamical system of an optimal policy for a grid-world problem, is always a dead-beat system in the framework of Discrete-Time Control systems.

Keywords: Dynamic Programming, Discrete-Time Control Systems, Markov Decision Process, Reinforcement Learning, Stochastic Control.

۱- مقدمه

مسائل جدولی[۱]، نوع خاص از فرآیند تصمیم‌گیری مارکوف[۲] یا MDP هستند. فرآیندهای تصمیم‌گیری مارکوف برای اولین بار توسط بلمن در سال ۱۹۵۷ و سپس توسط هُوارد در سال ۱۹۶۰ معرفی گردید و مورد بررسی قرار گرفت [1]. اولین کاربرد مشخصی که برای MDP ثبت شده است، استفاده از آن در سازمان‌دهی راه‌های ایالت آریزونا در سال ۱۹۷۸ بوده است [1]. همچنین کاربردهایی نظیر مدیریت حیات وحش، مدیریت تولید و کارخانه، انبارداری و حمل و نقل، از جمله کاربردهایی

[1] Grid-world Problems
[2] Markov Decision Process

هستند که برای MDP پیشنهاد شده‌اند [1]. فهرست کاملی از کاربردهای MDP و مدل‌های نیمه مشاهده‌پذیر MDP در نوشته‌هایی توسط پوترمن [1]، کاساندرا [2]، پایت [3] ، هو و همکارانش [18] و سو و همکارانش [19] آمده است. یافتن یک سیاست بهینه در مسأله‌ای که به صورت MDP مدل شده است، یکی از مباحثی است که در نظریه‌ی بهینه‌سازی و کنترل بسیار مورد توجه بوده است [1,5,7,18,19]. برای حل چنین مسأله‌ای از شیوه‌های برنامه‌ریزی خطی[1] و برنامه‌ریزی پویا[2] استفاده شده است و تغییرات متعددی در این الگوریتم‌ها به وجود آمده‌اند تا سرعت پاسخ‌دهی مناسبی برای این روش‌ها تأمین شود [1,5,8,18,19].

روش‌های اولیه‌ای که برای حل این مسأله با استفاده از برنامه‌ریزی پویا مورد استفاده قرار گرفته‌اند، به نام‌های ارزیابی سیاست[3]، تکرار سیاست[4] و تکرار ارزش[5] شناخته می‌شوند و در منابعی چون [4]، [5]، [6] و [7] مورد بررسی قرار گرفته‌اند. این روش‌ها پاسخ دقیقی را برای هر مسأله‌ی تصمیم‌گیری MDP به دست می‌دهند و می‌توان به قدر نیاز، جواب نهایی را به جواب واقعی نزدیک کرد. وجود جواب برای این روش‌ها، با استناد به قضایای مربوط به آنالیز توابع در فضای‌های برداری اندازه‌پذیر، به خصوص قضیه‌ی نقطه‌ی ثابت [17]، با روش‌های مختلفی مورد بررسی قرار گرفته است [6,7,16]. شرایط دقیقی که برای وجود جواب در یک مسأله‌ی MDP که توسط برنامه‌ریزی پویا حل می‌شود، به خوبی مورد مطالعه قرار گرفته‌اند و گردآوری شده‌اند [5,6,7,16]. چیزی که بلمن از آن به عنوان نفرین ابعاد یاد کرده است، یک مشکل جدی در زمینه‌ی استفاده از این روش‌ها به وجود آورده است، که باعثِ به وجود آمدن روش‌های تقریبی و سریع‌تر شده است که معمولا در قالب مباحث یادگیری تقویتی[6] [5,7,9] و یا برنامه‌ریزی عصبی-پویا[7] [4,8,9] مطرح می‌شوند.

روش‌های مورد استفاده برای بهتر کردن فرآیند حل MDP به چهار گروه اصلی قابل تقسیم می‌باشند. روش‌های گروه اول از ویژگی‌های ساختاری مسأله، برای تسهیل فرآیند به دست آوردن جواب یا سیاست بهینه، استفاده می‌کنند [11,12,26]. روش‌های بعدی، بر خلاف روش‌های گروه اول، منجر به پیدا شدن جواب بهینه برای مسأله نمی‌شوند. بلکه در این روش‌ها، جواب‌های به دست آمده جواب‌های شبه بهینه هستند و با فرض‌هایی که منجر به ساده شدن مسأله‌ی اصلی شده‌اند، به دست آمده‌اند [10,15]. روش‌های گروه دوم، خود به دو زیر گروه اصلی قابل تقسیم هستند. در روش‌های زیر گروه اول، از مدل‌های ساده شده برای حل مسأله استفاده می‌شود و به این ترتیب می‌توان یک سیاست شبه بهینه را برای مسأله‌ی اصلی پیدا نمود [10,15]. در روش‌های زیر گروه دوم، ساختار سیاست مجهول، به شکلی خاص فرض می‌شود و به این ترتیب نوعی ساده‌سازی در فرآیند حل مسأله به وجود می‌آید [4,6,10,15]. روش‌های گروه سوم، از تقریب توابع ارزش حالت، توابع ارزش حالت-عمل و معادلات برنامه‌ریزی پویا به دست آمده‌اند. در این روش‌ها از شیوه‌هایی همچون تجمیع حالات[8]، نمایش بر اساس توابع پایه[9] و استخراج خواص[10] استفاده شده است [4,5,7,9,10,13,14]. روش‌های گروه چهارم، همانند روش کلاسیک تکرار سیاست، در فضای سیاست تعریف و به کار برده می‌شوند. این روش‌ها، سعی بر این دارند که با شیوه‌های خاصی، شکل‌گیری سیاست بهینه را سریع‌تر کنند و به این ترتیب فرآیند حل مسأله سریع‌تر شود.

فرآیند حل یک MDP، در بحث یادگیری تقویتی، به نام یادگیری شناخته می‌شود و شکل‌گیری شیوه‌های تصمیم‌گیری جدید، همواره در اثر جمع‌آوری اطلاعات جدید می‌باشد. یکی از شیوه‌هایی که برای تسریع فرآیند یادگیری پیشنهاد شده است، استفاده از شیوه‌ی باز-استعمال است که برای تسریع فرآیند یادگیری و همچنین طبقه‌بندی اطلاعات به دست آمده از تجارب قبلی مورد استفاده قرار گرفته است [20-23]. همچنین با بهینه کردن ساختار الگوریتم تکرار سیاست نیز، نتایج مناسبی به دست آمده‌اند که می‌توان برای نمونه به [1]،[6]،[7]،[19]، [24]، [25] و [26] اشاره نمود.

روش‌های بسیاری برای بهبود عملکرد الگوریتم‌های برنامه‌ریزی پویا و همچنین الگوریتم‌های یادگیری تقویتی در محیط‌های مارکوف، ابداع شده‌اند. هر کدام از این روش‌ها با رویکردی خاص، قصد دارند در کمترین تعداد تکرار و کمترین زمان به یک پاسخ بهینه یا شبه بهینه، دسترسی پیدا کنند. مشکل اصلی در بسیاری از مسائل، نبودن شناخت کافی در مورد چگونگی یک پاسخ بهینه است. برای تشخیص یک سیاست بهینه و انجام مقایسه میان دو سیاست، معیاری سریع و ساده وجود ندارد. به این ترتیب که برای ارزیابی یک سیاست، حتما می‌بایست سیاست مذکور، توسط عامل یادگیرنده مورد استفاده قرار بگیرد و بر اساس خروجی به دست آمده، در مورد خوبی یا بدی آن سیاست، قضاوت شود. تا کنون، علی رغم تلاش‌های بسیاری که برای بهبود عملکرد الگوریتم‌های مرتبط با حل مسأله‌ی MDP انجام شده‌اند، روشی برای تحلیلِ ریاضیِ فرآیند حلِ MDP ارائه نشده است. به همین دلیل، معیاری غیر از الگوریتم‌های زمان‌بَر برای تحلیل عملکرد عامل یادگیرنده، که از سیاستی خاص پیروی می‌کند، وجود ندارد.

در این نوشتار فرآیند حل مسائل MDP با استفاده از روش DP، به صورت یک دینامیک گسسته-زمان یا دیجیتال بیان شده است. سیستم دیجیتالی که به دست می‌آید با روش حلی که برای مسأله ارائه شده

[1] Linear Programming
[2] Dynamic Programming
[3] Policy Evaluation
[4] Policy Iteration
[5] Value Iteration
[6] Reinforcement Learning
[7] Neuro-Dynamic Programming
[8] State Aggregation
[9] Basis Function Representation
[10] Feature Extraction

است، متناظر است و می‌توان با تحلیل خصوصیات کنترلی این سیستم، خواص روش حل متناظر با آن را مشخص نمود و کیفیت جواب نهایی را حدس زد. استفاده از این معادل‌سازی، این امکان را به وجود می‌آورد که بتوان به تحلیل عملکرد یادگیری تقویتی در محیط‌های مارکوف پرداخت.

سایر بخش‌های این مقاله به صورت زیر می‌باشند. در بخش ۲، یادگیری تقویتی و ایده‌ی اصلی آن به صورت اجمالی توضیح داده می‌شوند. در بخش ۳، تعاریف ابتدایی در مورد فرآیندهای تصمیم‌گیری مارکوف و روش‌های برنامه‌ریزی پویا برای حل این نوع از مسائل، مورد بررسی قرار می‌گیرند. در بخش ۴، الگوریتم ارزیابی سیاست به صورت یک سیستم دینامیکی گسسته-زمان بیان می‌شود. در بخش ۵، دو مسأله‌ی نمونه با استفاده از مطالب بخش ۴ و روش‌های تحلیل و کنترل سیستم‌های دینامیکی گسسته-زمان مورد بررسی قرار گرفته‌اند. بخش ۶ نیز، حاوی بیان نتایج کلی در مورد مسائل جدولی و سیاست بهینه‌ی مرتبط با این نوع از مسائل است.

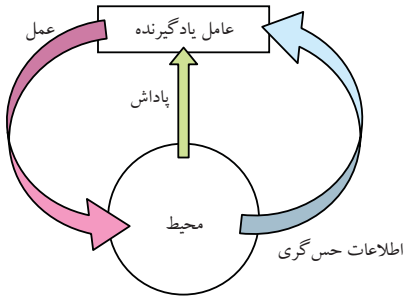
۲- یادگیری تقویتی

هدف اصلی از یادگیری، یافتن شیوه‌ای برای عملکرد در حالات مختلف است که این شیوه در مقایسه با سایرین، با در نظر گرفتن معیارهایی، بهتر است. معمولا این شیوه‌ی عملکرد، از نظر ریاضی، به صورت نگاشتی از فضای حالات به فضای اعمال، قابل بیان است. هنگامی می‌توان گفت یادگیری اتفاق افتاده است که، عاملی بر اساس تجربیاتی که کسب می‌کند به نحوی دیگر، و به احتمال زیاد بهتر، عمل کند. در این صورت می‌بایست نحوه‌ی عملکرد عامل در اثر کسب اطلاعات جدید، متفاوت از نحوه‌ی عملکرد در زمان‌های قبل از کسب این اطلاعات و تجارب باشد.

در یادگیری تقویتی، هدف اصلی از یادگیری، انجام دادن کاری و یا رسیدن به هدفی است، بدون آنکه عامل یادگیرنده، با اطلاعات مستقیم بیرونی تغذیه شود. در این روش، تنها مسیر اطلاع‌رسانی به عامل، از طریق یک سیگنال پاداش یا جریمه می‌باشد. تنها چیزی که از طریق سیگنال پاداش به عامل فهمانده می‌شود، این است که آیا تصمیم مناسبی گرفته است یا نه؟ در بسیاری از حیوانات، یادگیری تقویتی، تنها شیوه‌ی یادگیری مورد استفاده است. همچنین یادگیری تقویتی، بخشی اساسی از رفتار انسان‌ها را تشکیل می‌دهد. هنگامی که دست ما در مواجهه با حرارت می‌سوزد، ما به سرعت یاد می‌گیریم که این کار را بار دیگر تکرار نکنیم. لذت و درد مثالهای خوبی از پاداش‌ها و جرایمی هستند که الگوهای رفتاری ما و بسیاری از حیوانات را تشکیل می‌دهند.

در یادگیری تقویتی، هیچ گاه به عامل گفته نمی‌شود که عمل صحیح در هر وضعیت چیست، و فقط به وسیله‌ی معیاری، به عامل گفته می‌شود که یک عمل چقدر خوب یا چقدر بد است. عامل موظف است، با در دست داشتن این اطلاعات، یاد بگیرد که بهترین عمل کدام است. این موضوع، بخشی از نقاط قوت خاص یادگیری تقویتی است. از این طریق، مسائل پیچیده‌ی تصمیم‌گیری در اغلب اوقات می‌توانند با فراهم کردن کمترین میزان اطلاعات مورد نیاز برای حل مسأله، حل شوند. در این شیوه از یادگیری، حتی در برخی موارد، ماهیت مسأله و هدف از حل آن نیز به طور کامل و مستقیم به عامل تفهیم نمی‌شود. سیگنال پاداش، به طور ضمنی نحوه‌ی عملکرد مناسب را به عامل نشان می‌دهد و هدف از حل مسأله را مشخص می‌کند. در این حالت، هدف عامل از یادگیری به بیشینه کردن میزان پاداش دریافتی در بازه‌ای از زمان، تغییر می‌کند. این دو هدف (نحوه‌ی عملکرد مناسب و بیشینه کردن پاداش‌ها) با هم مترادف هستند و برآورده شدن هرکدام، دیگری را نیز برآورده خواهد کرد. به این طریق، عامل نحوه‌ی عملکرد مناسب را با تمرکز بر پاداش‌های دریافتی، یاد می‌گیرد. این امر به این صورت محقق می‌شود که، نگاشتی میان حالات و اعمالِ قابل انجام توسط عامل، پیدا می‌شود. این نگاشت که به نام سیاست شناخته می‌شود، به عامل می‌گوید که در مواجهه با حالات مختلف، چه عمل یا اعمالی را انجام دهد. تبعیت از یک سیاست خوب، قطعا عامل را به نتیجه‌ای مناسب خواهد رساند.

یک مسأله‌ی یادگیری تقویتی نوعی، در شکل ۱ نشان داده شده است. عامل یادگیرنده از طریق حس‌گرها، توصیفی از حالت محیط اطرافش را به دست می‌آورد. اطلاعات مربوط به محیط در قالب اطلاعات حس‌گری به عامل داده می‌شوند. هنگامی که عامل ، عملی را انجام می‌دهد، پاداشی را دریافت می‌کند که می‌تواند بسته به خوبی یا بدی عمل، پاداشی مثبت یا منفی باشد.



شکل ۱ - یک مسأله‌ی یادگیری تقویتی و نحوه‌ی تعامل محیط و عامل

۳- فرآیندهای تصمیم‌گیری مارکوف

بخش اعظمی از کارهای تحقیقاتی انجام شده بر روی یادگیری تقویتی، توأم با این فرض بوده‌اند که، تعامل بین عامل و محیط اطرافش را می‌توان به صورت یک فرآیند تصمیم‌گیری مارکوف یا MDP گسسته-زمان مدل‌سازی کرد. یک فرآیند تصمیم‌گیری مارکوف، فرآیندی تصادفی و گسسته-زمان می‌باشد که اغلب به صورت دسته‌ی

چهار تایی $(\mathbb{S},\mathbb{A},\mathcal{P},\mathcal{R})$ تعریف می‌شود. اجزای یک فرآیند تصمیم‌گیری مارکوف، عبارتند از:

- یک زمان‌سنج سراسری به صورت $t=0,1,\ldots,T$ برای شمارش زمانِ گسسته. (T می‌تواند نامحدود باشد.)
- $\mathbb{S}$ نشان دهنده‌ی فضای حالت فرآیند می‌باشد و مشتمل بر تمام حالات ممکنی است که عامل تصمیم گیرنده در آن‌ها قرار می‌گیرد و ملزم به تصمیم‌گیری در این حالات می‌باشد. فرض بر این است که $\mathbb{S}$ مجموعه‌ای محدود به صورت $\mathbb{S}=\left\{s^1,s^2,\ldots,s^n\right\}$ می‌باشد.
- $\mathbb{A}$ نشان دهنده‌ی فضای اعمال قابل انتخاب، برای عامل تصمیم‌گیرنده است. این مجموعه حاوی انتخاب‌ها و یا تصمیمات ممکن در هر حالت، برای عامل می‌باشد. فرض بر این است که $\mathbb{A}$ مجموعه‌ی محدودی به صورت $\mathbb{A}=\left\{a^1,a^2,\ldots,a^m\right\}$ می‌باشد.
- $\mathcal{P}$ نحوه‌ی انتقال و تحول حالات را مدل می‌کند. اگر فرآیند در حالت s باشد، و عمل a توسط عامل انتخاب شود، احتمال تغییر حالت فرآیند به حالت s' به صورت $\mathcal{P}^a_{ss'}$ تعریف می‌شود. به عبارت دیگر داریم:

$$\mathcal{P}^a_{ss'}=\Pr\left\{s'\middle|s,a\right\} \tag{۱}$$

در حالت کلی $\mathcal{P}$ نگاشتی به صورت $\mathcal{P}:\mathbb{S}\times\mathbb{A}\times\mathbb{S}\to\left[0,1\right]$ است. این مدل، دارای خاصیت مارکوف می‌باشد. یعنی احتمال فوق، صرفا به حالت و عمل اخیر بستگی دارد و کاملا مستقل از خاطره‌ی قبلی عامل می‌باشد. اگر s_t و a_t به ترتیب نشان دهنده‌ی حالت فرآیند و عمل انتخاب شده در زمان گسسته‌ی t باشند، آن‌گاه خاصیت مارکوف به صورت رابطه‌ی زیر قابل توصیف می‌باشد:

$$\mathcal{P}^{a_t}_{s_ts_{t+1}}=\Pr\left\{s_{t+1}\middle|s_t,a_t\right\}=\Pr\left\{s_{t+1}\middle|s_t,a_t,s_{t-1},a_{t-1},\ldots\right\} \tag{۲}$$

- $\mathcal{R}$ تابع تعریف کننده‌ی امید ریاضی پاداش می‌باشد. اگر عامل در حالت s باشد و با انجام عمل a به حالت s' برود، مقدار پاداشی که دریافت می‌کند به صورت عددی تصادفی و با امید ریاضی $\mathcal{R}^a_{ss'}$ تعریف می‌شود. در حالت کلی تابع پاداش، نگاشتی به صورت $\mathcal{R}:\mathbb{S}\times\mathbb{A}\times\mathbb{S}\to\mathbb{R}$ می‌باشد. تابع $\mathcal{R}$ نیز دارای خاصیت مارکوف می‌باشد و مقدار $\mathcal{R}^a_{ss'}$ صرفا به حالت فعلی (یعنی s) و عمل فعلی (یعنی a) و حالت بعدی (یعنی s') بستگی دارد و کاملا مستقل از حالات یا اعمال قبلی می‌باشد. به عبارت دیگر:

$$\begin{aligned}\mathcal{R}^{a_t}_{s_ts_{t+1}}&=\mathbb{E}\left\{r_{t+1}\middle|s_{t+1},s_t,a_t\right\}\\&=\mathbb{E}\left\{r_{t+1}\middle|s_{t+1},s_t,a_t,s_{t-1},a_{t-1},\ldots\right\}\end{aligned} \tag{۳}$$

فرض کنید عامل در زمان یا مرحله‌ی t در حالت s_t قرار دارد و عمل $a_t\in\mathbb{A}$ را انجام می‌دهد. عامل با احتمال $\mathcal{P}^{a_t}_{s_ts'}$ در زمان $t+1$ به حالت s' می‌رود و خواهیم داشت: $s_{t+1}=s'$. ضمنا عامل در نتیجه‌ی این عمل، پاداشی اسکالر به اندازه‌ی r_{t+1} خواهد گرفت که در حالت کلی، کمیتی تصادفی و با امید ریاضی $\mathcal{P}^{a_t}_{s_ts_{t+1}}$ می‌باشد.

احتمال انتخاب عمل a از طرف عامل، هنگامی که در حالت s قرار دارد، با نگاشتی به صورت $\pi:\mathbb{S}\times\mathbb{A}\to\left[0,1\right]$ تعریف می‌شود و می‌توان نوشت:

$$\Pr\left\{a_t=a\middle|s_t=s\right\}=\pi(s,a) \tag{۴}$$

نگاشت π با نام سیاست شناخته می‌شود و مجهول اصلی یک مسأله‌ی یادگیری تقویتی و هر مسأله‌ی تصمیم‌گیری می‌باشد [1,5,7,14].

برای مقایسه‌ی سیاست‌های مختلف با یکدیگر، می‌توان معیاری را برای سنجش آن‌ها تعریف نمود. این معیار، مقداری است که سیاست در هر حالت از فرآیند برمی‌گرداند و به عنوان خروجی[1] سیاست در حالت مذکور از آن یاد می‌شود. خروجی یک سیاست، میزانی از پاداش است که در اثر اتخاذ تصمیمات متوالی و با تبعیت از آن سیاست به دست آمده است. برای هر کدام از حالت‌ها، ارزشی در نظر گرفته می‌شود که برابر با امید ریاضی خروجی است که با شروع کردن از هر حالت و تبعیت از یک سیاست خاص به دست می‌آید. در حالت کلی، منظور از حل یک مسأله‌ی یادگیری تقویتی، پیدا کردن سیاست π^* است به نحوی که مقدار خروجی سیاست و یا ارزش هر کدام از حالت‌ها، بیشینه شوند [5,7,14]. روش‌های متفاوتی برای تعریف خروجی وجود دارند. روشی که در اکثر کاربردها معمول است و در این مقاله نیز مورد توجه قرار گرفته است، تعریف خروجی به صورت تنزیلی[2] می‌باشد. اگر ضریب تنزیل به صورت $\gamma\in\left[0,1\right]$ باشد، خروجی تنزیلی به صورت زیر خواهد بود:

$$z_t=r_t+\gamma r_{t+1}+\gamma^2r_{t+2}+\cdots=\sum_{k=0}^{\infty}\gamma^kr_{t+k} \tag{۵}$$

در این حالت، ارزش حالت s به صورت زیر قابل بیان است:

$$\begin{aligned}V^\pi(s)&=\mathbb{E}^\pi\left\{z_{t+1}\middle|s_t=s\right\}\\&=\sum_{a\in\mathbb{A}}\pi(s,a)\sum_{s'\in\mathbb{S}}\mathcal{P}^a_{ss'}\left(\mathcal{R}^a_{ss'}+\gamma V^\pi(s')\right)\end{aligned} \tag{۶}$$

که در آن، منظور از $\mathbb{E}^\pi$، عملگر امید ریاضی است. اندیس‌های π نیز، صرفا برای تاکید بر این که عامل از سیاست π پیروی می‌کند، نوشته شده‌اند. رابطه‌ی (۶)، به معادله‌ی (بهینگی) بلمن معروف است [1,4,14,18,19].

یکی از روش‌هایی که برای حل این معادله و یافتن مقدار ارزش تمام حالات استفاده می‌شود، بازگشتی کردن این معادله است. این روش که مبتنی بر قضیه‌ی نقطه‌ی ثابت[3] [17] است، پیشنهاد می‌کند که معادله‌ی (۶) به صورت زیر بازنویسی شود:

[1] Return
[2] Discounted
[3] Fixed Point Theorem

$$V_{k+1}^{\pi}(s)=\sum_{a\in\mathbb{A}}\pi(s,a)\sum_{s'\in\mathbb{S}}\mathcal{P}_{ss'}^{a}\left(\mathcal{R}_{ss'}^{a}+\gamma V_{k}^{\pi}(s')\right) \quad (۷)$$

که در آن $V_k^{\pi}(s)$، تخمین kاُم از مقدار واقعی $V^{\pi}(s)$ است. با توجه به این که $|\gamma|<1$ $|\gamma|<1$ است، می‌توان استدلال کرد که $V_{k+1}^{\pi}(s)$ با یک نگاشت انقباضی[1] [17] به $V_k^{\pi}(s)$ مرتبط است. طبق قضیه‌ی نقطه‌ی ثابت [17]، این نگاشت دارای نقطه‌ی ثابتِ منحصر به فردی است که جواب معادله‌ی (۶) نیز می‌باشد. با توجه به معادله‌ی (۷)، جواب معادله‌ی (۶) به صورت زیر خواهد بود:

$$V^{\pi}(s)=\lim_{k\to\infty}V_{k}^{\pi}(s) \quad (۸)$$

فرآیند محاسبه‌ی $V^{\pi}(s)$ برای تمام حالت‌ها، در بحث یادگیری تقویتی و برنامه‌ریزی پویا به نام ارزیابی سیاست [1,5,7,14] معروف است.

۴- مدل‌سازی الگوریتم ارزیابی سیاست

فرض کنید برداری به صورت

$$v^{\pi}=\begin{bmatrix}V^{\pi}(s^1) & V^{\pi}(s^2) & \cdots & V^{\pi}(s^n)\end{bmatrix}^T \quad (۹)$$

تعریف شده باشد. این بردار حاوی ارزش تمام حالات یک مدل است. در این صورت می‌توان رابطه‌ی بازگشتی (۷) را، به شکل زیر برای تمام حالت بازنویسی کرد و آن را به صورت یک معادله تبدیل نمود:

$$v_{k+1}^{\pi}=\gamma\mathcal{P}v_{k}^{\pi}+\mathcal{R}=\gamma\begin{bmatrix}\mathcal{P}_{s^1s^1} & \cdots & \mathcal{P}_{s^1s^n}\\ \vdots & \ddots & \vdots\\ \mathcal{P}_{s^ns^1} & \cdots & \mathcal{P}_{s^ns^n}\end{bmatrix}v_{k}^{\pi}+\begin{bmatrix}\mathcal{R}_{s^1}\\ \vdots\\ \mathcal{R}_{s^n}\end{bmatrix} \quad (۱۰)$$

که در آن داریه‌های ماتریس‌های $\mathcal{P}$ و $\mathcal{R}$ عبارتند از:

$$\mathcal{P}_{ss'}=\sum_{a\in\mathbb{A}}\pi(s,a)\mathcal{P}_{ss'}^{a} \quad (۱۱)$$

و

$$\mathcal{R}_{s}=\sum_{a\in\mathbb{A}}\sum_{s'\in\mathbb{S}}\pi(s,a)\mathcal{P}_{ss'}^{a}\mathcal{R}_{ss'}^{a}=\sum_{a\in\mathbb{A}}\pi(s,a)\mathcal{R}_{s}^{a} \quad (۱۲)$$

می‌توان معادله‌ی (۱۰) را به صورت زیر بازنویسی کرد:

$$v_{k+1}^{\pi}=\gamma\mathcal{P}v_{k}^{\pi}+\mathcal{R}u_{k} \quad (۱۳)$$

که در آن فرض شده است که u_k به ازای تمام مقادیر $k\geq 0$، برابر با واحد باشد، که همان تعریف تابع پله‌ی واحد [27] می‌باشد. معادله‌ی (۱۳) معادله‌ی حالت یک سیستم گسسته-زمان یا دیجیتال [27] می‌باشد که متغیرهای حالت آن، ارزش‌های مربوط به حالات فرآیند تصمیم‌گیری می‌باشند. ماتریس حالت این سیستم، از ترکیب اطلاعات مربوط به محیط در قالب $\mathcal{P}_{ss'}^{a}$، اطلاعات مربوط به سیاست در قالب $\pi(s,a)$، و ضریب تنزیل به دست آمده است. بردار وروی این سیستم $\mathcal{R}$ می‌باشد که اطلاعات مربوط به محیط، سیاست و پاداش‌ها را در بر دارد. طبق قرارداد، ورودی این سیستم، همواره برابر با پله‌ی واحد در نظر گرفته می‌شود. شرط پایداری سیستم فوق، عبارت است از این که، همه‌ی مقادیر ویژه‌ی ماتریس $\gamma\mathcal{P}$، که قطب‌های سیستم توصیف شده با (۱۰) هستند ، در داخل دایره‌ی واحد قرار بگیرند [27]. برای تحقق این شرط، می‌بایست ضریب تنزیل γ در نامساوی زیر صدق کند:

$$\gamma<\frac{1}{\max\limits_{1\leq i\leq n}\left|\lambda_i(\mathcal{P})\right|}=\frac{1}{\rho(\mathcal{P})} \quad (۱۴)$$

که در آن، $\lambda_i(\mathcal{P})$ نشان دهنده‌ی مقدار ویژه[2]ی iاُم، و $\rho(\mathcal{P})$ نیز شعاع طیفی[3] ماتریس $\mathcal{P}$ می‌باشد. لذا مشاهده می‌شود که شرط $\gamma\leq 1$، الزاما تضمین کننده‌ی همگرایی سری تعریف شده در معادله‌ی (۵) و یا وجود جواب محدود برای (۶) نمی‌باشد. رابطه‌ی (۱۴)، شرط دقیق‌تری برای γ بیان می‌کند. اگر تمام مقادیر ویژه‌ی ماتریس $\mathcal{P}$، درون دایره‌ی واحد باشند، آن‌گاه سری (۵)، به ازای برخی از مقادیر γ نیز، که بزرگتر از یک هستند، همگرا خواهد بود.

۵- حل و بررسی دو مسأله‌ی نمونه

در این بخش، با استفاده از مطالب مطرح شده در بخش قبل، دو مسأله‌ی نمونه مورد حل و بررسی قرار می‌گیرند. مشاهداتی که در حل این دو مسأله انجام گرفته است، راهگشای نتیجه‌گیری‌های کلی در خصوص مسائل جدولی و مسائل مشابه هستند.

۵-۱- مسأله‌ی اول

یک مسأله‌ی جدولی را، به صورت نشان داده شده در شکل ۲، در نظر بگیرید. عاملی (مثلا یک روبات) در یکی از خانه‌های سفید رنگ این جدول قرار دارد. عامل در هر حالتی می‌تواند به سمت چپ یا راست حرکت کند. هنگامی که عامل به یکی از خانه‌های خاکستری برسد، حرکت او متوقف می‌شود. حرکت به چپ یا راست، پاداشی به اندازه‌ی ۱- در پی دارد که در واقع هزینه‌ای است که عامل برای حرکت کردن می‌پردازد. هدف از حل مسأله، پیدا کردن شیوه‌ای برای حرکت است که عامل از هر کدام از حالات، در کمترین تعداد حرکت به یکی از خانه‌های هدف برساند.

s^0	s^1	s^2	s^3	s^0

شکل ۲ - جدول مربوط به مثال مورد بررسی در بخش ۵-۱

جدول ۱ - احتمال انتخاب حرکات در حالات مختلف برای $\pi^{(p)}$

$\pi^{(p)}$	چپ (L)	راست (R)
s_1	$1-p$	p

[1] Contraction Mapping

[2] Eigenvalue
[3] Spectral Radius

s_2	$\frac{1}{2}$	$\frac{1}{2}$
s_3	p	$1-p$

سیاستی به صورت نشان داده شده در جدول ۱ را برای مسأله‌ی حاضر در نظر بگیرید. توجه کنید که $p \in [0,1]$ پارامتری است که تغییر مقدار آن، باعث ایجاد سیاست‌های مختلف برای مسأله‌ی مورد بررسی می‌شود. به این ترتیب سیاست $\pi^{(\frac{1}{2})}$، یک سیاست کاملا تصادفی و با احتمال مساوی برای حرکت‌های چپ و راست است. سیاست $\pi^{(0)}$ نیز، یک سیاست بهینه برای این مسأله است و عاملی که از این سیاست تبعیت کند، از هر حالت، در کم‌ترین تعداد حرکت به یکی از خانه‌های هدف خواهد رسید. سیستم دینامیکی معادل با ارزیابی سیاست $\pi^{(p)}$، به صورت زیر خواهد بود:

$$\pi^{(p)} : v_{k+1}^{\pi^{(p)}} = \begin{bmatrix} 0 & p & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & p & 0 \end{bmatrix} v_k^{\pi^{(p)}} + \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} u_k \qquad (۱۵)$$

تابع تبدیل معادل با سیستم توصیف شده با معادله‌ی حالت فوق عبارت است از:

$$G(z) = \frac{1}{z^2 - p} \begin{bmatrix} -(z+p) \\ -(z+1) \\ -(z+p) \end{bmatrix} \qquad (۱۶)$$

قطب‌های سیستم فوق $+\sqrt{p}$ و $-\sqrt{p}$ هستند. هنگامی که $p=0$ باشد، یعنی عامل از سیاست $\pi^{(0)}$ تبعیت کند، تمام قطب‌های سیستم فوق در مبدأ صفحه‌ی z قرار می‌گیرند. چنین سیستمی در مبحث کنترل دیجیتال، به نام سیستم مرده نَوِش[1] شناخته می‌شود. یک سیستم مرده نَوِشِ درجه n، در مقایسه با سایر سیستم‌های درجه n، سریع‌ترین پاسخ ممکن را دارد و پاسخ پله‌ی آن، دقیقا در n واحد زمانی گسسته، به مقدار نهایی می‌رسد [27]. از معادله‌ی حالت (۱۵) بر می‌آید که ورودی سیستم مورد بررسی، سیگنال u_k و خروجی آن v_k^{π} است. اگر رابطه‌ی بین ورودی و خروجی را با استفاده از تبدیل z بنویسیم، خواهیم داشت:

$$V(z) = G(z)U(z) \qquad (۱۷)$$

که در آن $V(z)$ و $U(z)$، به ترتیب تبدیل z سیگنال‌های v_k^{π} و u_k هستند. مقدار نهایی v_k^{π}، جواب الگوریتم ارزیابی سیاست است. طبق قضیه‌ی مقدار نهایی برای سیگنال‌های گسسته، مقدار نهایی v_k^{π} به صورت زیر قابل محاسبه است:

$$\begin{aligned} v^{\pi} &= \lim_{k\to\infty} v_k^{\pi} = \lim_{z\to 1} (1-z^{-1})V(z) \\ &= \lim_{z\to 1} (1-z^{-1})G(z)U(z) \end{aligned} \qquad (۱۸)$$

با توجه به این که سیگنال u_k، همواره برابر با پله‌ی واحد در نظر گرفته می‌شود، داریم:

$$v^{\pi} = \lim_{z\to 1} (1-z^{-1})G(z)\frac{1}{1-z^{-1}} = \lim_{z\to 1} G(z) = G(1) \qquad (۱۹)$$

لذا برای سیستم توصیف شده با معادله‌ی حالت (۱۵)، که معادل با ارزیابی سیاست $\pi^{(p)}$ است، مقدار نهایی متغیرهای حالت به صورت زیر قابل محاسبه هستند:

$$v^{\pi} = \begin{bmatrix} -\frac{1+p}{1-p} & -\frac{2}{1-p} & -\frac{1+p}{1-p} \end{bmatrix}^T \qquad (۲۰)$$

این مقادیر نهایی، نشان دهنده‌ی متوسط هزینه‌ای هستند که عامل با شروع از هر یک از حالات، برای رسیدن به خانه‌های هدف، می‌پردازد. کمترین مقدار هزینه‌ای که پرداخت می‌شود، به ازای $p=0$ به دست می‌آید. سیاست معادل با این مقدار، $\pi^{(0)}$ است که یک سیاست بهینه برای این مسأله است.

در محاسبات انجام شده، مقدار ضریب تنزیل γ برابر با یک در نظر گرفته شده است. اگر γ را در محاسبات وارد کنیم، معادله‌ی حالت به دست آمده، به صورت زیر خواهد بود:

$$\pi^{(p)} : v_{k+1}^{\pi^{(p)}} = \gamma \begin{bmatrix} 0 & p & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & p & 0 \end{bmatrix} v_k^{\pi^{(p)}} + \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} u_k \qquad (۲۱)$$

شرط لازم و کافی برای همگرایی الگوریتم ارزیابی سیاست، پایداری سیستم فوق است. این سیستم در صورتی پایدار است که همه‌ی مقادیر ویژه‌ی ماتریس حالت آن در داخل دایره‌ی واحد قرار بگیرند [27]. مقادیر ویژه‌ی ماتریس حالت سیستم توصیف شده با معادله‌ی حالت (۲۲)، عبارتند از:

$$\lambda_1 = 0 \qquad \lambda_{2,3} = \pm\gamma\sqrt{p} \qquad (۲۲)$$

لذا شرط پایداری این سیستم و همچنین همگرایی الگوریتم ارزیابی سیاست $\pi^{(p)}$ به صورت زیر است:

$$\gamma\sqrt{p} < 1 \quad \Rightarrow \quad \gamma < \frac{1}{\sqrt{p}} \qquad (۲۳)$$

شرط فوق بیان می‌کند که اگر سیاست مورد ارزیابی $\pi^{(0)}$ باشد، الگوریتم به ازای تمام مقادیر γ همگرا خواهد بود.

۵-۲- مسأله‌ی دوم

جدولی به صورت شکل ۳ را در نظر بگیرید. در این مسأله نیز، عامل در یکی از خانه‌های سفید رنگ جدول قرار دارد، و می‌بایست با حرکت در یکی از چهار جهتِ بالا، پایین، چپ و راست، خود را به یکی از دو خانه‌ی هدف، که با رنگ خاکستری مشخص شده‌اند، برساند. حرکت

[1] Dead Beat

در هر جهت، پاداشی به اندازه‌ی ۱- در پی دارد، که این پاداش، نشان دهنده‌ی هزینه‌ای است که عامل برای هر حرکت می‌پردازد. حرکت‌هایی که باعث خارج شدن عامل از جدول می‌شوند، بر موقعیت عامل تاثیری ندارند و محل عامل را تغییر نمی‌دهند. عامل باید یاد بگیرد که با دریافت بیشترین پاداش (پرداخت کمترین جریمه)، خود را به یکی از خانه‌های هدف برساند. اگر چنین کاری محقق شود، عامل توانسته است با کمترین تعداد حرکت، به هدف برسد. این مسأله به صورت یک فرآیند تصمیم‌گیری مارکوف، قابل بیان می‌باشد و می‌توان برای حل آن، از روش ارزیابی سیاست استفاده کرد. برای استفاده از روش ارزیابی سیاست، ارزش اولیه‌ی هر کدام از خانه‌ها، برابر با صفر در نظر گرفته می‌شود [1,5,7,14]. طبق قضیه‌ی نقطه‌ی ثابت، نتیجه‌ی نهایی، مستقل از ارزش اولیه‌ی خانه‌ها می‌باشد و الگوریتم همواره به یک نقطه‌ی منحصر به فرد در فضای جستجو، همگرا می‌شود [17].

s^0	s^1	s^2	s^3
s^4	s^5	s^6	s^7
s^8	s^9	s^{10}	s^{11}
s^{12}	s^{13}	s^{14}	s^0

شکل ۳ – جدول مربوط به مسأله‌ی مورد بررسی در بخش ۵-۲

سیاستی که تاپایان حل مسأله مورد استفاده قرار گرفته است، سیاست تصادفی است. به این معنی که، در همه‌ی خانه‌های جدول، احتمال حرکت به تمام جهات، مساوی و همگی برابر با یک چهارم یا ۰/۲۵ می‌باشد. در شکل ۴، چند مرحله از حل تکراری معادله‌ی بلمن، توسط معادله‌ی (۷)، نشان داده شده است. با استفاده از نتایج مربوط به هر مرحله، می‌توان سیاستی را پایه‌ریزی کرد. به این ترتیب که، عامل می‌بایست در هر خانه از جدول، به سمت خانه‌هایی حرکت کند که بیشترین ارزش را دارند. سیاستی که با استفاده از ارزش‌های به دست آمده در مرحله‌ی k اُم به دست می‌آید، به صورت π_k نشان داده شده است. π_0 همان سیاست تصادفی است. π_∞ نیز سیاستی است که با استفاده از ارزش‌های نهایی به دست می‌آید. π_3 و تمام سیاست‌های بعد از آن، همگی معادل هستند و $\pi_3 = \pi_\infty$ می‌باشد. فرض کنید با استفاده از هر کدام از سیاست‌های به دست آمده، معادله‌ی سیستم معرفی شده در معادله‌ی (۱۰) محاسبه شوند، و ماتریس $\mathcal{P}$ در معادله‌ی (۱۰) برای سیاست π_i به صورت $\mathcal{P}_i$ باشد. استفاده از اندیس i، صرفا به دلیل جلوگیری از تداخل اندیس‌ها در معادله‌ی (۱۰) می‌باشد. شعاع طیفی هر کدام از ماتریس‌های مذکور محاسبه شده‌اند و عبارتند از:

$$\begin{aligned} &\rho(\mathcal{P}_0) \simeq 0.9468, \quad \rho(\mathcal{P}_1) \simeq 0.8431, \\ &\rho(\mathcal{P}_2) = 0.5, \qquad \rho(\mathcal{P}_3) = \rho(\mathcal{P}_\infty) \simeq 0 \end{aligned} \tag{۲۳}$$

به وضوح دیده می‌شود که هر چه قدر سیاست به کار رفته در ایجاد مدل، بهینه‌تر باشد، اندازه‌ی بزرگترین مقدار ویژه‌ی ماتریس حالت نیز کوچک‌تر می‌شود. به خصوص به ازای سیاست π_3 یا همان π_∞، تمامی مقادیر ویژه برابر با صفر هستند. به این ترتیب، حد بالای ضریب تنزیل γ، برای همگرایی سری (۵)، به ازای سیاست‌های π_0 تا π_3 به ترتیب عبارت است از: ۱/۰۵۶۲، ۱/۱۸۶۱، ۲ و ∞. به عبارت دیگر، سری تعریف شده با معادله‌ی (۵)، به شرط پیروی از سیاست $\pi_3 = \pi_\infty$، به ازای تمام مقادیر γ، همگرا خواهد بود.

از طرفی، قطب‌های سیستم (۱۰)، با مقادیر ویژه‌ی ماتریس $\gamma\mathcal{P}$ برابر هستند. مشاهده می‌شود که قطب‌های سیستم معادل با سیاست π_∞، همگی در مبدا قرار دارند. به عبارت دیگر، این سیستم نیز یک سیستم مرده نَوِش است. هدف از حل مسأله‌ی فوق نیز، رسیدن به یکی از خانه‌های هدف در کمترین تعداد حرکت می‌باشد. لذا کاملا طبیعی است که پاسخ بهینه، متناظر با یک سیستم مرده نَوِش باشد، که سریع‌ترین پاسخ را در بین سیستم‌های هم‌درجه‌اش دارد. می‌توان استدلال کرد که، الگوریتم ارزیابی سیاست برای π_∞، حد اکثر در ۱۴ تکرار همگرا می‌شود و پس از آن، هیچ تغییری در ارزش حالات ایجاد نخواهد شد.

با تعریف ارزش همه‌ی حالات به عنوان خروجی، می‌توان تابع تبدیل این سیستم‌ها را به صورت زیر به دست آورد:

$$G_i(z) = (zI - \gamma\mathcal{P}_i)^{-1}\mathcal{R}_i \tag{۲۵}$$

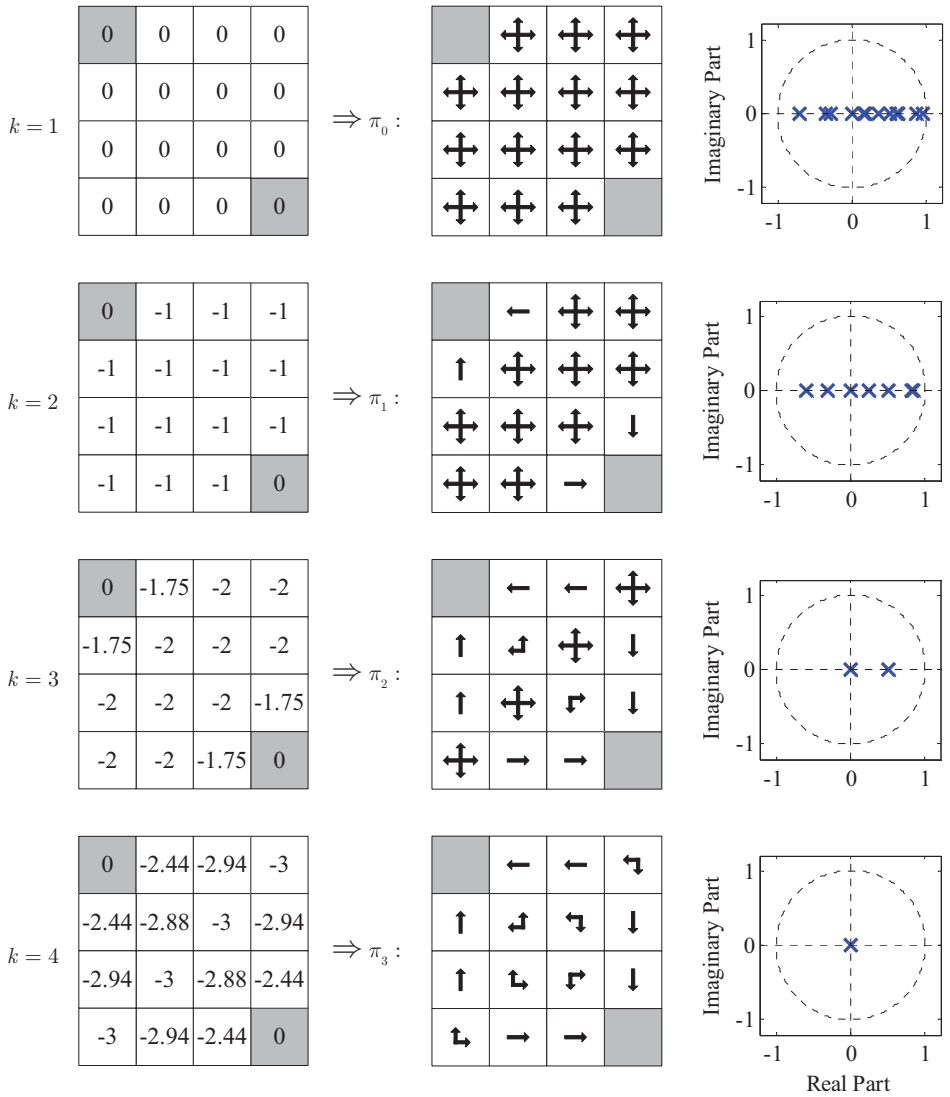
تابع تبدیل فوق، متناظر با سیستمی با یک ورودی و ۱۴ خروجی می‌باشد. هر کدام از خروجی‌ها، متناظر با ارزش یکی از خانه‌های جدول مربوط به مسأله‌ی مورد بررسی می‌باشند.

به عنوان نمونه، پاسخ فرکانسی هر یک از سیستم‌ها را به ازای s^3 در شکل ۵ مشاهده می‌کنید. با توجه به تقارن موجود در مسأله، این پاسخ فرکانسی، مربوط به خانه‌ی s^{12} نیز می‌باشد. توجه کنید که درجه‌بندی محور عمودی، به صورت دِسی بِل (dB) انتخاب نشده است و مقادیر نشان داده شده، مقادیر واقعی هستند.

با توجه به این که در کنترل دیجیتال، رابطه‌ی فرکانس موهومی ω با فرکانس مختلط z به صورت $z = e^{j\omega T_s}$ است، در شکل ۵، فرکانس موهومی ω به بازه‌ی $[0,\pi]$ محدود شده است. منظور از T_s، زمان نمونه‌برداری سیستم دیجیتال است. توجه نمایید که پاسخ فرکانسی سیستم، که تبدیل فوریه‌ی گسسته از سیگنالی گسسته است، یک سیگنال متناوب و پیوسته است و دوره‌ی تناوب آن 2π می‌باشد و به دلیل حقیقی بودن سیستم، تابعی زوج بر حسب ω می‌باشد [27].

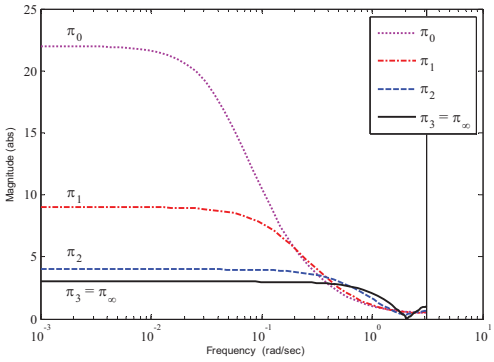
پاسخ فرکانسی نشان داده شده در شکل ۵، حاوی اطلاعات مهمی در مورد محیط و سیاست به کار رفته از طرف عامل، می‌باشد. در این شکل، عملکرد حالت ماندگار سیستم در خانه‌ی s^3، با توجه به پاسخ فرکانسی، قابل مشاهده است. عملکرد حالت ماندگار این سیستم‌ها، متناظر با مقدار پاسخ فرکانسی در فرکانس صفر است. دیده می‌شود که

دامنه‌ی پاسخ فرکانسی در فرکانس صفر، به ازای سیاست‌های π_0، π_1، π_2 و π_3 به ترتیب برابر با ۲۲، ۹، ۴ و ۳ است.



شکل ۴ – مراحل الگوریتم ارزیابی سیاست برای π_0 و سیاست‌های استخراج شده از اطلاعات هر مرحله به همراه نمودار قطب‌های سیستم دینامیکی معادل با ارزیابی هر کدام از سیاست‌ها

این مقادیر، متوسط هزینه‌ای هستند که عامل برای رسیدن به هر یک از خانه‌های هدف و با شروع از s^3 می‌پردازد. با توجه به تعریف تابع پاداش برای این مسأله، مقدار جریمه برابر با تعداد حرکت‌هایی است که عامل برای رسیدن به خانه‌های هدف، انجام می‌دهد. مشاهده می‌شود که به ازای π_3، کمترین تعداد حرکت‌ها به دست آمده است. در شکل ۵ مشاهده می‌شود که سیستم‌های معادل با سیاست‌های بهتر، پهنای باند وسیع‌تری دارند. با توجه به این که پهنای باند، معیاری از سرعت پاسخ‌دهی هر سیستم است، می‌توان استدلال کرد که، برای سیاست‌های بهتر، سرعت پاسخ‌دهی سیستم‌ها بیشتر است. این نکته همان چیزی است که در مورد مسأله‌ی حاضر، از یک سیاست خوب انتظار داریم.



شکل ۵ – پاسخ فرکانسی سیستم‌های تبعیت کننده از سیاست‌های π_0 تا π_3 در خانه‌ی s^3 از جدول نشان داده شده در شکل

۶- بررسی جامع مسائل جدولی

نتایج به دست آمده در این بخش، به مسائل جدولی و یا مسائل مشابه مربوط می‌شوند. خواص مشترک مسائلی که نتیجه‌گیری‌های این بخش در مورد آن‌ها صدق می‌کنند، عبارتند از:

- می‌بایست مسأله در قالب فرآیند تصمیم‌گیری مارکوف قابل توصیف باشد.
- یک یا چندین حالت نهایی[1] وجود داشته باشند که با رسیدن به آن حالات، وظیفه‌ی عامل تمام می‌شود.
- تمام حرکات، هزینه‌ای به صورت پاداش منفی داشته باشند.
- هدف از حل مسأله، یافتن شیوه‌ای برای رسیدن به حالت نهایی است که مستلزم پرداخت کمترین هزینه، و یا اخذ کمترین میزان پاداش منفی، باشد. هزینه‌ی کل به صورت مجموع تنزیلی پاداش‌های دریافت شده توسط عامل، تعریف می‌شود.

یک مسأله‌ی جدولی را، با یک یا چند خانه‌ی هدف در نظر بگیرید. از نظر مدل مارکوفی، تمام خانه‌های هدف، به عنوان یک حالت واحد در نظر گرفته می‌شوند. ارزش حالتی که متناظر با خانه‌های هدف است، همواره برابر با صفر است. لذا جهت اختصار، ارزش این حالت از بردار ارزش حالت حذف گردیده است.

بدون کاسته شدن از کلیت مسأله، فرض شده است که، هر حرکتی، هزینه‌ای دارد که به صورت پاداشی به اندازه‌ی ۱- مدل شده است. همچنین ضریب تنزیل به صورت $\gamma = 1$ در نظر گرفته شده است. هر یک از خانه‌های عادی جدول، معادل با یک حالت مانند s هستند. برای هر حالت s، قطعا می‌توان مسیری به سمت یکی از خانه‌های هدف پیدا کرد که مستلزم دریافت کمترین جریمه باشد. این مسیر دارای کمترین تعداد حرکت ممکن است و برای حالت s، این تعداد حرکت با $m(s)$ نمایش داده می‌شود. با در نظر گرفتن فرض‌های یاد شده، می‌توان قضیه‌ای را در خصوص مسائل جدولی و به صورت زیر بیان نمود.

قضیه. سیستم دینامیکی معادل با ارزیابی سیاست برای سیاست بهینه‌ی یک مسأله‌ی جدولی، دارای خواص زیر است:

- یک سیستم مرده نَوِش است و مولفه‌ای از آن که متناظر با حالت s است، دقیقا $m(s)$ قطب دارد که همگی در مبدأ صفحه‌ی z قرار دارند.
- هنگامی که $\gamma = 1$ اختیار می‌شود، صفرهای مولفه‌ی مربوط به حالت s، به همراه نقطه‌ی $z = 1$، همگی ریشه‌های $m(s)$اُم واحد هستند و محیط دایره‌ی واحد را به $m(s)$ قسمت مساوی تقسیم می‌کنند.
- این سیستم در مقایسه با سیستم‌های معادل با ارزیابی سیاست‌های دیگر، سریع‌ترین پاسخ ممکن را دارد.
- این سیستم به ازای تمامی مقادیر ضریب تنزیل γ، پایدار است. به عبارت دیگر، الگوریتم ارزیابی سیاست برای سیاست بهینه‌ی یک مسأله‌ی جدولی، مستقل از مقدار γ، همواره همگرا است.
- اندازه‌ی پاسخ فرکانسی در فرکانس موهومی $\omega = 0$، برای این سیستم، کمترین مقدار ممکن را دارد. در صورتی که $\gamma = 1$ باشد، اندازه‌ی پاسخ فرکانسی متناظر با حالت s، در فرکانس $\omega = 0$ یا $z = 1$، برابر با $m(s)$ خواهد بود. این مقدار برابر با قرینه‌ی ارزش نهایی حالت s است.

□

اثبات. اگر عاملی از سیاست بهینه تبعیت کند، با شروع از حالت s، مسیر بهینه‌ای را به سمت خانه‌های هدف طی خواهد کرد، دقیقا با انجام $m(s)$ حرکت به خانه‌ی هدف خواهد رسید، و نهایتا پاداشی به اندازه‌ی $-m(s)$ دریافت خواهد نمود.

اگر $\gamma = 1$ در نظر گرفته شود، ارزش خانه‌ی s، پس از انجام دادن حرکت kاُم، برابر با مجموع پاداش‌هایی است که تا زمان انجام حرکت kاُم، توسط عامل دریافت شده است. به این ترتیب، ارزش حالت s، به شرطی که عامل از سیاست بهینه تبعیت کند، به صورت تابعی از زمانِ گسسته قابل تعریف است:

$$v_s^*[k] = \begin{cases} 0 & ,k < 0 \\ -k & ,0 \le k \le m(s) \\ -m(s) & ,k > m(s) \end{cases} \qquad (۲۶)$$

توجه کنید که زمان در تعریف فوق، متناظر با شماره‌ی تکرار در الگوریتم ارزیابی سیاست می‌باشد. تابع فوق، یکی از خروجی‌های سیستم معادل با ارزیابی سیاست بهینه می‌باشد. تبدیل z تابع ارزش فوق، عبارت است از:

$$\mathcal{Z}\left\{v_s^*[k]\right\} = V_s^*(z) = -\sum_{k=1}^{m(s)} kz^{-k} - m(s)\sum_{k=m(s)+1}^{\infty} z^{-k} \qquad (۲۷)$$

از طرفی، طبق مطالب مطرح شده، ورودی سیستم معادل با ارزیابی سیاست، همواره برابر با پله‌ی واحد اختیار می‌شود. لذا تابع تبدیل متناظر با ارزش حالت s، با توجه به رابطه‌ی زیر قابل محاسبه است:

$$V_s^*(z) = G_s^*(z)U(z) = \frac{G_s^*(z)}{1 - z^{-1}} \qquad (۲۸)$$

لذا تابع تبدیل متناظر با حالت s، یعنی $G_s^*(z)$، عبارت است از:

$$G_s^*(z) = (1 - z^{-1})V_s^*(z) \qquad (۲۹)$$

با جایگذاری عبارت به دست آمده برای $V_s^*(z)$ در رابطه‌ی اخیر، $G_s^*(z)$ به صورت زیر به دست خواهد آمد:

$$\begin{aligned} G_s^*(z) &= -\sum_{k=1}^{m(s)} z^{-k} = -\frac{z^{m(s)-1} + \cdots + z + 1}{z^{m(s)}} \\ &= -\frac{z^{m(s)} - 1}{z^{m(s)}(z-1)} \end{aligned} \qquad (۳۰)$$

[1] Terminal State

مشاهده می‌شود که، تابع تبدیل متناظر با حالت s، دقیقا $m(s)$ قطب دارد که همگی در مبدأ صفحه‌ی z قرار دارند. به عبارت دیگر، $G_s^*(z)$ همواره یک سیستم مرده نَوِش است. همچنین صفرهای تابع تبدیل $G_s^*(z)$، همواره بر روی دایره‌ی واحد قرار دارند و همگی ریشه‌های $m(s)$اُم واحد هستند. توجه کنید که $z=1$، صفر تابع تبدیل $G_s^*(z)$ نمی‌باشد. نقطه‌ی $z=1$ و صفرهای تابع تبدیل $G_s^*(z)$، محیط دایره‌ی واحد را به $m(s)$ قسمت مساوی تقسیم می‌کنند. مطالب مطرح شده، مربوط به یک مولفه‌ی اختیاری از تابع تبدیل سیستم معادل با ارزیابی سیاست بهینه بودند. همه‌ی مولفه‌های این تابع تبدیل، مرده نَوِش هستند و بنابراین تمام قطب‌های این تابع تبدیل در مبدأ صفحه‌ی z قرار دارند. با توجه به نظریه‌ی کنترل سیستم‌های گسسته-زمان، سیستم مرده-نوش، سریع‌ترین پاسخ ممکن را در بین تمام سیستم‌های هم‌درجه دارد [27]. همچنین با توجه به این که قطب‌های سیستم در مبدأ صفحه‌ی z قرار دارند، و با توجه به نحوه‌ی تاثیر ضریب γ بر روی محل قطب‌ها در رابطه‌ی (۱۳)، می‌توان استدلال کرد که γ تاثیری بر محل قطب‌های سیستم ندارد و از این رو، سیستم تعریف شده با (۳۰)، مستقل از مقدار γ، همواره پایدار است.

اگر عامل یادگیرنده از سیاست بهینه پیروی نکند، مقدار تابع ارزش آن در تمام لحظات، قطعا کمتر از یا مساوی با مقدار تابع ارزش مربوط به سیاست بهینه خواهد بود. فرض کنید سیاست π، نشان دهنده‌ی سیاستی باشد که در تمام حالت‌ها دقیقا مثل سیاست بهینه باشد و فقط در یک حالت به خصوص مانند s'، متفاوت از سیاست بهینه باشد. همچنین فرض کنید، عاملی که از این سیاست تبعیت می‌کند، با شروع از حالت s، پس از k_0 حرکت، به s' می‌رسد. در این حالت می‌توان تابع ارزش حالت s را در حرکت kاُم و به شرط تبعیت از سیاست π، به صورت زیر تعریف نمود:

$$v_s^\pi[k] = \begin{cases} 0 & ,k<0 \\ -k & ,0 \le k < k_0 \\ -k-\sum_{i=k_0}^{k}\delta_i & ,k \ge k_0 \end{cases} \qquad (۳۱)$$

که در آن، δ_i مقداری از جریمه‌ی اضافه‌ای است که در اثر عدم تبعیت از سیاست بهینه، در حرکت k_0 دریافت می‌شود. اگر تبدیل z مربوط به این تابع ارزش محاسبه شود و طبق رابطه‌ی (۲۹)، تابع تبدیل سیستم محاسبه شود، خواهیم داشت:

$$G_s^\pi(z) = -\sum_{k=1}^{k_0-1} z^{-k} - \sum_{k=k_0}^{\infty}(1+\delta_k)z^{-k} \qquad (۳۲)$$

سیستم فوق فقط به شرطی می‌تواند مرده-نوش باشد که بتوان k_1 را پیدا کرد، به نحوی که:

$$\forall k > k_1,\ \ 1+\delta_k = 0 \qquad (۳۳)$$

در غیر این صورت سیستم مرده-نوش نیست و کندتر از سیستم (۳۰) عمل خواهد کرد. البته درجه‌ی توابع تبدیل (۳۲) و (۳۰) نیز یکسان نخواهد بود و حتی در صورت مرده-نوش بودن (۳۲)، باز هم سیستم (۳۰) دارای سرعت پاسخ‌دهی بیشتری خواهد بود.

برای سیستم (۳۲)، اندازه‌ی پاسخ فرکانسی در $\omega=0$ یا $z=1$ بزرگتر از سیستم (۳۰) است و داریم:

$$\begin{aligned} &G_s^\pi(1) = -(k_0-1) - \sum_{k=k_0}^{\infty}(1+\delta_k) \\ &G_s^\pi(1) \le -m(s) - \sum_{k=k_0}^{\infty}\delta_k \le -m(s) = G_s^*(1) \\ &\Rightarrow \left|G_s^\pi(1)\right| \ge \left|G_s^*(1)\right| \end{aligned} \qquad (۳۴)$$

■

۷- نتیجه‌گیری

در این مقاله با ارائه‌ی مروری بر یادگیری تقویتی، فرآیندهای مارکوف و برنامه‌ریزی پویا، معادلات مربوط به حل یک فرآیند مارکوف با استفاده از برنامه‌ریزی پویا، به صورت یک دینامیک گسسته-زمان جمع‌بندی و دوباره‌نویسی شدند. این روش برخورد، امکان بررسی فرآیند حل یک مسأله‌ی یادگیری تقویتی در محیط مارکوف را، در قالب یک سیستم دیجیتال فراهم می‌آورد. به این ترتیب، می‌توان شیوه‌های مرسوم در کنترل دیجیتال را برای تحلیل یک فرآیند یادگیری استفاده نمود. موضوع بحث این نوشتار، بر روی مسائلی موسوم به مسائل جدولی می‌باشد. نتایج حاکی از آن هستند که یک سیاست بهینه برای این نوع از مسائل، در قالب کنترل دیجیتال به صورت یک سیستم مرده نَوِش قابل توصیف می‌باشد. تعمیم این نتیجه به انواع دیگر مسائل و تعریف دوگانی بین فضای تصمیم‌گیری و فضای سیستم‌های کنترل دیجیتال، از مطالعات و تحقیقات تکمیلی در ادامه‌ی این مقاله هستند.

مراجع

[1] M. L. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming, Wiley, 2005.
[2] A. Cassandra, "Exact and Approximate Algorithms for Partially Observable Markov Decision Processes," as Ph.D. Thesis, Brown University, 1998.
[3] L. Pyeatt, "Integration of Partially Observable Markov Decision Processes and reinforcement Learning for Simulated Robot Navigation," as Ph.D. Thesis, Colorado State University, 1999.
[4] D. P. Bertsekas and J. N. Tsitsiklis, Neuro-Dynamic Programming, Athena Scientific, 1996.

[5] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, The MIT Press, 1998.
[6] A. Lew and H. Mauch, Dynamic Programming: A Computational Tool, Springer-Verlag, Berlin, 2007.
[7] S. I. Reynolds, "Reinforcement Learning with Exploration," as Ph.D. Thesis, School of Computer Science, The University of Birmingham, UK, 2002.
[8] B. Van Roy, "Neuro-Dynamic Programming: Overview and Recent Trends," chapter of, E. A. Feinberg and A. Schwartz, Handbook of Markov Decision Processes: Methods and Applications, Kluwer Academic, 2002.
[9] J. Si et al., Handbook of Learning and Approximate Dynamic Programming, Wiley Inter-Science, 2004.
[10] Hyeong Soo Chang et al, "A survey of some Simulation-Based Algorithms for Markov Decision Processes," in Communications in Information and Systems, Vol. 7, No. 1, pp. 59-92, 2007.
[11] J. E. Smith and K. F. Mc Cardle, "Structural Properties of Stochastic Dynamic Programs," in Operations Research, Vol. 50, pp. 796-809, 2002.
[12] M. C. Fu et al., "Monotone optimal policies for queuing staffing problem," in Operations Research, Vol. 46, pp. 327-331, 2000.
[13] R. Givan et al. "Bounded Markov Decision Processes," in Artificial Intelligence, Vol. 122, pp. 71-109, 2000.
[14] L. P. Kaelbling, M. L. Littman and A. W. Moore, "Reinforcement Learning: A Survey," Journal of Artificial Intelligence Research, Vol. 4, pp. 237-285, 1996.
[15] G. J. Gordon, "Approximate Solution to Markov Decision Processes," Ph.D. Thesis, School of Computer Science, Carnegie Mellon University, 1999.
[16] D. P. de Farias and B. Van Roy, "On the Existance of Fixed Points for Approximate Value Iteration and Temporal-Difference Learning," in Journal of Optimization theory and Applications, Vol. 105, No. 3, pp. 589-608, 2000.
[17] H. Royden, Real Analysis (3rd Edition), Prentice Hall, 1988.
[18] Qiying Hu and Wuyi Yue, Markov Decision Processes with Their Applications, Springer Science+Busines Media, LLC, 2008.
[19] Heyeong Soo Chang et al., Simulation-based Algorithms for Markov Decision Processes, Springer-verlag, London, 2007.
[20] F. Fernandez and M. Veloso, "Exploration and Policy Reuse," as Technical Report, School of Computer Science, Carnegie Mellon University, 2005.
[21] F. Fernandez and M. Veloso, "Probabilistic Reuse of Past policies," as Technical Report, School of Computer Science, Carnegie Mellon University, 2005.
[22] F. Fernandez and M. Veloso, "Building a Library of Policies through Policy Reuse," as Technical Report, School of Computer Science, Carnegie Mellon University, 2005.
[23] D. S. Bernstein, "Reusing Old Policies to Accelerate Learning on New Markov Decision Processes," as Technical Report, Department of Computer Science, University of Massachusetts, Amherst Tech. Rep. No. 99-26, 1999.
[24] N. L. Zhang and W. Zhang, "Speeding Up the Convergence of Value Iteration in Partially Observable Markov Decision Processes," in Journal of Artificial Intelligence Research, Vol. 14, pp. 29-51, 2001.
[25] E. A. Hansen, "An Improved Policy Iteration for Partially Observable Markov Decision Processes," in Proceedings of 10th Neural Information Processing Systems Conference, 1997.
[26] B. Sallans, "Reinforcement Learning for Factored Markov Decision Processes," as Ph.D. Thesis, Graduate Department of Computer Science, University of Toronto, 2002.
[27] K. Ogata, Discrete-Time Control Systems (2nd Edition), Prentice Hall, 1994.

مجله کنترل جلد ۳، شماره ۱، بهار ۱۳۸۸، صفحه ۲۲-۱۲

طراحی موجک بهینه برای نویززدایی از طیف انرژی رادیوایزوتوپ‌های گسیلنده پرتوگاما جهت استخراج برهمکنش‌های غالب

علی فیاضی [۱]، حسین احمدی نوبری [۲]

[۱] دانشجوی کارشناسی ارشد برق- کنترل، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران، دانشکده فنی و مهندسی، a.fayazi@yahoo.co.uk

[۲] دانشیار، قطب علمی کنترل و پردازش هوشمند، دانشکدهٔ برق و کامپیوتر، دانشگاه تهران، noubari@ece.ubc.ca

چکیده: تحلیل طیف پرتوگاما با استفاده از تبدیل موجک در سالهای اخیر به یک روش جدید برای کاهش نویز جهت شناسایی برهمکنش های غالب فوتون با ماده نظیر جذب فتوالکتریک، لبه کامپتون و پیک‌های ناشی از پرتوهای بازگشتی در طیف‌های پرتو گاما پیشنهاد شده است. در این مقاله روشی جدید برای نویززدایی از طیف‌های پرتو گاما(^{60}Co و ^{137}Cs) معرفی گردید. این روش بر مبنای انتخاب پایه هایی از موجک است که بهترین تقریب غیر خطی را از طیف داده شده، فراهم می‌کنند. نشان داده شد که چنین پایه‌هایی از موجک بهترین عملکرد نویززدایی را به لحاظ خطای تقریب طیف خواهند داشت. نتایج شبیه سازی نشان داد که استفاده از موجک بهینه طراحی شده در نویززدایی طیف، باعث بهبود ۱۵ درصدی نسبت سیگنال به نویز نسبت به موجک های استاندارد هم طول می شود.

کلمات کلیدی: آنالیز موجک، طیف پرتو گاما، برهمکنش های غالب، موجک بهینه، نویززدایی.

Abstract: In recent years analysis of γ-ray spectra using wavelet transform has offered a new approach for an improved noise reduction resulting in an accurate identification of dominant interactions of photon with material in γ-ray spectras. Examples of such interaction are: Photoelectric effect, Compton edge, Scattared photon's. In this paper, a novel approach is presented for denosing of γ-ray spectras (^{60}Co, ^{137}Cs). The approach is based on searching for a wavelet basis that provides the best nonlinear approximation of a given spectra. It is shown that such a basis will have the best wavelet denoising performance in the sense of spectrum estimation error. The result of simulation indicate that an improvement of up to 15% is achieved in SNR using optimally designed wavelet over the same length Daubechies wavelet in denoising of the spectra.

Keywords: Wavelet analysis, γ-ray spectra, Dominant Interactions, Optimum Wavelet, Denoising.

۱- مقدمه

اکثر سیگنال های کاربردی، سیگنال هایی در حوزه زمان هستند. اما در اکثر کاربردهای مرتبط با پردازش سیگنال نمایش زمانی همیشه بهترین نمایش از یک سیگنال نیست. در بسیاری از موارد اطلاعات مهم و تشخیصی سیگنال در محتوای فرکانسی آن نهفته است. برای دست یابی به محتوای فرکانسی سیگنال از تبدیل فوریه سیگنال استفاده می شود. اما نقص و مشکل جدی که تحلیل فوریه در انتقال به حوزه فرکانس دارد، اینست که اطلاعات زمانی سیگنال گم می شود. آنالیز موجک[۱] بدلیل قابلیت تحلیل همزمان در حوزه زمان و فرکانس به‌طور گسترده در پردازش سیگنال استفاده شده است. نشان داده شده که تخمین سیگنال از سیگنال نویزی بااستفاده از موجک‌ها عملکرد بهتری نسبت به روش-های معمول نظیر فیلتر کردن در حوزه فرکانس و فیلتر وینر دارد[۱].

ویژگی مهم تبدیل موجک[۲] که به قابلیت آن در نویززدایی کمک می-کند، توانایی تمرکز انرژی سیگنال روی تعداد محدودی از ضرایب است. روش نویززدایی معرفی شده توسط داناهو و جانستون[۳] [۲] می-تواند با بهره‌گیری از اطلاعات قبلی سیگنال جهت طراحی یک موجک بهینه متناسب باویژگیهای سیگنال، بهبود یابد.

روش‌های زیادی برای نویززدایی موجک های بهینه پیشنهاد شده است. در [۳] نسبت سیگنال به نویز[۴] (SNR)یک سیگنال پس از نویززدایی با موجک به عنوان یک معیار بررسی شده است. در [۴] خطای بین سیگنال اصلی و سیگنال بازسازی شده بااستفاده از ضرایب مقیاس به حداقل رسیده است. در این روش، انرژی سیگنال غالباً بر روی ضرایب

[1] Wavelet Analysis

[2] Wavelet Transform

[3] Danaho and Johnstone

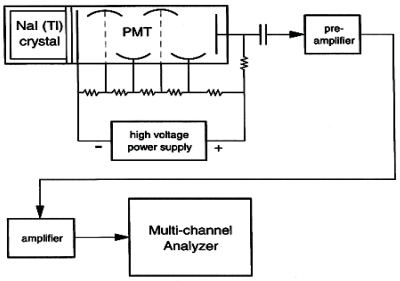
[4] Signal to Noise Ratio

مکانی مقیاس متمرکز شده است. چاپا و رائو[1] روشی بر مبنای انطباق موجک به یک سیگنال به خصوص در حوزه فرکانس معرفی کرده‌اند[۵].

در این مقاله یک موجک بهینه بر مبنای حداقل‌سازی خطای تخمین غیرخطی برای یک سیگنال مشخص (طیف رادیوایزوتوپ‌های گسیلنده پرتو گاما) طراحی شده است. این حداقل سازی منجر به بهبود حداکثری نسبت سیگنال به نویز (SNR) می‌شود. علاوه بر این، چنین موجکی برای استخراج مشخصه‌های مهم یک سیگنال کارآمد بوده و می‌تواند مورد استفاده قرارگیرد. روش طراحی موجک بهینه به منظور بهبود SNR و مشخص کردن ویژگیهای بخصوص(برهمکنش‌های غالب فوتون باماده) در طیف رادیوایزوتوپ‌های گسیلنده پرتو-گاما(^{60}Co و ^{137}Cs) اعمال شد. طیف رادیوایزوتوپ‌های گسیلنده پرتو-گاما نوعا نویزی بوده و برای شناسایی مشخصه‌ها و ویژگیهای مهم درآنها بایستی کیفیت طیف بهبود یابد [۸-۶]. داده‌های این مقاله (^{60}Co و ^{137}Cs) به‌طور حقیقی بااستفاده از دستگاه اسپکترومتری گاما مجهز به آشکارساز سنتیلاسیون یدیرو سدیم (NaI) در آزمایشگاه آنالیز هسته-ای بدست آمده است. رادیوایزوتوپ‌های گسیلنده پرتو گاما کاربردهای فراوانی دارند، به عنوان مثال ^{137}Cs در انرژی های کوچک برای کالیبره کردن تجهیزات آشکار سازی تشعشع و در انرژی های بزرگ به عنوان وسیله پرتودرمانی برای درمان سرطان استفاده می شود. در اندازه گیری های صنعتی برای آشکار کردن جریان مایع درون لوله و در دیگر تجهیزات صنعتی برای اندازه گیری ضخامت مواد، نظیر کاغذ، فیلم عکاسی یا ورق های فلز بکار می رود.

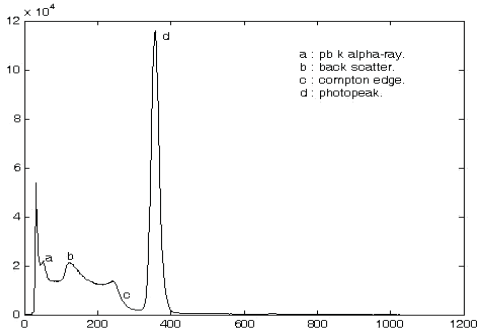
۲- دستگاه‌اسپکترومتری گاما

مشخصات و قسمت های مختلف یک دستگاه اسپکترومتری گامانوعا شامل یک آشکارساز ژرمانیوم (Ge) مخزن نیتروژن مایع پیش تقویت کننده منبع ولتاژ معین برای آشکارساز تقویت کننده خطی مبدل قیاسی به عددی (ADC) تحلیلگر چند کاناله (MCA) وسیستم جمع آوری اطلاعات (MCB) می باشد. شمای دستگاه اسپکترومتری گاما در شکل ۱ نمایش داده شده است.



شکل ۱: شمای دستگاه اسپکترومتری گاما

طیف انرژی رادیو ایزوتوپ های گسیلنده پرتوگاما که تعداد را در انرژی مربوطه نشان می‌دهد. در حالت ایده آل به صورت یک تابع شبیه به تابع دیراک است. البته برای بعضی از رادیو ایزوتوپ‌ها همانند ید می‌تواند شامل چندین پیک بسیار تیز باشد ولی این طیف دستخوش عواملی چون لبه کامپتون در آشکارساز، قله پراکندگی ناشی از پرتوهای بازگشتی، قله اشعه ایکس سرب در کالیماتور، اسکتر و غیره حالت تیز بودن خود را از دست داده پهن می‌شود. مهمترین برهمکنش‌های فوتون با ماده عبارتند از: جذب فوتوالکتریک، پراکندگی کامپتون. این برهمکنش‌ها در اثر برخورد فوتون با الکترون‌های مداری، هسته، اتمها یا میدان‌های الکتریکی اطراف آنها صورت می‌پذیرد. دراین برخوردها فوتون قسمتی یا تمام انرژی خود را از دست می‌دهد. هر اثر در انرژی خاص و برای کاربردهای خاص حائز اهمیت است. در محدوده انرژی-های تشخیصی در پزشکی هسته‌ای جذب فوتوالکتریک و پراکندگی کامپتون دو برهمکنش غالب هستند. در شکل ۲ نمونه‌ای از یک طیف انرژی رادیوایزوتوپ پرتو گاما(^{137}Cs) همراه با مشخصات آن که بوسیله دستگاه اسپکترومتری گاما بدست آمده، نمایش داده شده است.



شکل ۲: طیف انرژی رادیوایزوتوپ پرتو گاما(^{137}Cs)

۳- تبدیل موجک گسسته

در یک تحلیل با تفکیک چندگانه متعامد یکه[2] (OMRA)، بااستفاده از تبدیل موجک گسسته[3] (DWT). یک سیگنال $f(t) \in V_0$ به یک سری نامحدود توابع جزئی $\{g_i(t)\}$ تجزیه می‌شود به‌طوریکه:

$$f(t)=\sum_{j=-\infty}^{0} g_i(t) \qquad (۱)$$

اولین مرحله تجزیه با تصویر کردن $f(t)$ روی دو زیر فضای متعامد W_{-1}, V_{-1} انجام می‌گیرد که در آن $V_0 = V_{-1} \oplus W_{-1}$ و $(\oplus)$ عملگر جمع مستقیم می‌باشد. این تجزیه باعث ایجاد $f_{-1}(t) \in V_{-1}$، که یک تقریب با تفکیک پایین تر از $f(t)$ است، و

[1] Chapa and Rao

[2] Orthogonal Multiresolution Analysis

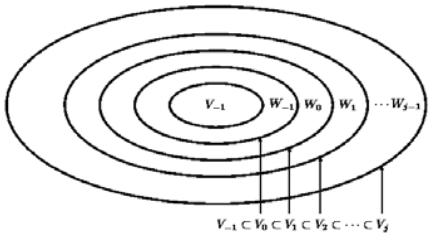
[3] Discrete Wavelet Transform

$g_{-1} \in W_{-1}$، کـه خطـای جزئـی ناشـی از تقریـب $f_{-1}(t)$ از $f(t)$ اسـت، مـی‌شـود. عمـل تجزیـه بـه دو زیـر فضـای متعامـد W_{-2}, V_{-2} و همین‌طور تصویر کردن‌های بعدی ادامه می‌یابد. زیر-فضاهای V_j و خواص آن به‌صورت روابط (۲) و (۳) بیان می‌شود:

$$\ldots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \ldots \qquad (۲)$$

$$\bigcup_{-\infty}^{\infty} V_j = L^2(\Re); \qquad (۳)$$

شکل ۳ این زیر فضاهای تودرتو را نمایش می دهد.



شکل ۳: شمایی از زیر فضاهای تودرتو شامل زیر فضاهای W_j, V_j

توابع پایه متعامد یکه در زیر فضاهای W_j و V_j توسط روابط زیر بیان می‌شود:

$$\psi_{j,k} = 2^{j/2}\psi(2^j t - k) \quad k, j \in Z \qquad (۴)$$

$$\phi_{j,k} = 2^{j/2}\phi(2^j t - k) \quad k, j \in Z \qquad (۵)$$

که در آن $\psi(t)$، موجک مادر و $\phi(t)$، تابع مقیاس می‌باشد به-طوریکه:

$$\int \psi(t)dt = 0 \Leftrightarrow \Psi(0) = 0 \qquad (۶)$$

$$\int \phi(t)dt = 1 \Leftrightarrow \Phi(0) = 0 \qquad (۷)$$

و $\Psi(\omega), \Psi(\omega)$ به ترتیب تبدیل فوریه $\psi(t), \phi(t)$ می‌باشند. معادلات تصویر کردن $f(t)$ روی زیرفضاها به صورت زیر بیان می-شود:

$$f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a_{j,k}.2^{j/2}\phi(2^j t - k) \qquad (۸)$$

که،

$$a_{j,k} = \int_{-\infty}^{\infty} f(t).2^{j/2}\phi(2^j t - k)dt. \qquad (۹)$$

$a_{j,k}$ ضرایب مقیاس را بیان می‌کند. اما از آن‌جایی‌که زیر فضاهای V_j تودرتو می باشند لذا بسط $f(t)$ متعامد نیست. همچنین توابع $\left\{ 2^{j/2}\phi(2^j t - k)\big|_{j,k\in z} \right\}$ در سرتاسر مقیاس ها متعامد نخواهند بود. یعنی:

$$\left\langle 2^{j/2}\phi(2^j t - k), 2^{j'/2}\phi(2^{j'} t - k) \right\rangle\big|_{k,k'\in z} \neq 0, \quad when \quad j \neq j', \qquad (۱۰)$$

بنابراین این توابع نمی توانند پایه های متعامد را برای فضای $L^2(\Re)$ تشکیل دهند. زیرفضاهای تودرتو بدین معنی است کـه زیـر فضـای V_1 شامل زیرفضای V_0 و زیر فضای V_2 شامل زیرفضای V_1 و همین‌طور زیـر فضـای V_j شـامل زیرفضـای V_{j-1} مـی باشـد. بـدین ترتیـب $\phi(t)$ که در زیرفضای V_0 قرار دارد، می توانـد بـر حسـب توابـع پایـه مربوط به زیر فضای V_1 مطابق رابطه زیر (معادله اتساع) بیان شود.

$$\phi(t) = \sqrt{2}\sum_k h(k)\phi(2t - k) \qquad (۱۱)$$

که در آن $h(k)$ تصویر $\phi(t)$ برروی توابع پایه زیرفضای V_1 می-باشد، و از رابطه زیر بدست می آید:

$$h(k) = \sqrt{2}\int_{-\infty}^{\infty} \phi(t)\phi(2t - k)dt. \qquad (۱۲)$$

به منظور رسیدن به یک تجزیـه متعامـد از فضـای $L^2(\Re)$، فضـاهای تفاضـل W_j را بـه عنـوان مکمـل فضـاهای V_j در V_{j+1} تعریـف مـی شوند. شکل ۳ این زیرفضاهای تفاضل را نشان می دهد. به‌گو نه ای‌که:

$$V_{j+1} = V_j \oplus W_j \qquad and \qquad V_j \cap W_j = \phi \qquad (۱۳)$$

زیرفضاهای W_j فضای $L^2(\Re)$ را به زیر فضاهای متعامد تجزیه می کنند. به‌گونه‌ای‌که:

$$W_j \perp W_{j'} \quad if \quad j \neq j', \quad and \quad \bigoplus_{j=-\infty}^{\infty} W_j = L^2(\Re). \qquad (۱۴)$$

زیـــر فضـــای W_j بوســـیله دســـته ای ازتوابـــع پایـه $\left\{ 2^{j/2}\psi(2^j t - k)\big|_{j,k\in z} \right\}$، پوشـش داده مـی شـود کـه شیفت یافته و مقیاس یافته دودویی سیگنال موجـک $\psi(t)$ مـی باشـند. بنـابراین هرسـیگنال $f(t) \in L^2(\Re)$ را مـی تـوان بـه صـورت ترکیب خطی از این توابع پایه نوشت:

$$f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} b_{j,k}.2^{j/2}\psi(2^j t - k) \qquad (۱۵)$$

$$b_{j,k} = \int_{-\infty}^{\infty} f(t).2^{j/2}\psi(2^j t - k)dt \qquad (۱۶)$$

ازآنجاییکه، زیر فضاهای W_j متعامد هستند. این بسط، یک بسط متعامد از سیگنال $f(t)$ می باشد. $b_{j,k}$ ضرایب موجک نامیده می-شود. معادلات (۱۵) و (۱۶) با هم معادلات ترکیب و تحلیل تبدیل موجک گسسته را تشکیل می دهند. $\psi(t)$ در زیر فضای W_0 قرار گرفته است، که بخشی از زیرفضای V_1 می باشد. بنابراین می توان تابع موجک $\psi(t)$ را بر حسب توابع پایه مربوط به زیر فضای V_1 به صورت معادله زیر(معادله موجک) بیان نمود:

$$\psi(t) = \sqrt{2}\sum_k g(k)\phi(2t - k) \qquad (۱۷)$$

که $g(k)$ تصویر $\psi(t)$ برروی توابع پایه زیرفضای V_1 می باشد و ازرابطه زیر بدست می آید.

$$g(k) = \sqrt{2}\int_{-\infty}^{\infty}\psi(t)\,\phi(2t-k)\,dt. \qquad (۱۸)$$

زیرفضاهای تودرتو در شکل ۳ دارای توابع پایه ایی هستند که از مقیاس های مختلف سیگنال $\phi(t)$ مشتق شده اند. بنابراین نگاشت هر سیگنال $f(t)$ برروی این زیرفضاها درواقع امکان دیدن این سیگنال را در مقیاس ها و تفکیک های مختلف می دهد. در مقیاس های بالا ضرایب موجک $(b_{j,k})$ جزئیات نرم $f(t)$ را نشان می دهد، در حالیکه در مقیاس های پایین تر ضرایب موجک ساختار های سخت $f(t)$ را اندازه گیری می کند. بدین ترتیب است که یک سیگنال $f(t)\in L^2(\Re)$ در یک تحلیل با تفکیک چندگانه متعامد یکه[1]، به یک سری نامحدود از توابع جزئی تجزیه می شود. در حالت کلی، هرزیرفضای V_M از جمع مستقیم $V_N\big|_{N<M}$ و $W_j\big|_{N\le j<M}$ به صورت زیر بدست می آید:

$$V_M = V_N \oplus W_N \oplus W_{N+1} \oplus W_{N+2} \oplus \ldots \oplus W_{M-1}. \qquad (۱۹)$$

بنابراین تبدیل موجک گسسته هر سیگنال $f(t)\in L^2(\Re)$ می‌تواند بوسیله تعداد محدودی از ضرایب موجک بیان شود. باتوجه به معادله بالا هر سیگنال $f(t)\in V_M$ می تواند بر حسب ضرایب موجک در مقیاس های مورد علاقه $(N\le j<M)$، و ضرایب مقیاس در مقیاس N، $(<M)$ به صورت زیربسط داده شود.

$$f(t)=2^{N/2}\sum_{k=-\infty}^{\infty}a_{N,k}\,\phi(2^N t-k)+\sum_{j=N}^{M-1}2^{j/2}\sum_{k=-\infty}^{\infty}b_{j,k}\,\psi(2^j t-k). \qquad (۲۰)$$

اگر سیگنال $f(t)$ محدود باشد. حدود برروی متغییر شیفت K در عبارت بالا محدود است. بنابراین سیگنال های محدود با تعداد محدودی ضرایب موجک و مقیاس مشخص می شوند. از آن‌جایی‌که $\phi(t)\in V_{-1}\subset V_0$ بوده و $\psi(t)\in W_{-1}\subset V_0$ می‌باشد، این توابع می‌توانند به صورت ترکیب خطی از توابع پایه فضای V_0 نمایش داده شوند:

$$\phi(t)=\sqrt{2}\sum_{k=-\infty}^{\infty}h[k]\phi(2t-k) \qquad (۲۱)$$

$$\psi(t)=\sqrt{2}\sum_{k=-\infty}^{\infty}g[k]\phi(2t-k) \qquad (۲۲)$$

در فضای تحلیل با تفکیک چندگانه متعامد یکه دنباله های $h[k]$ و $g[k]$ در معادلات (۲۱) و (۲۲) بیانگر پاسخ ضربه فیلترهای آینه‌ای چهارتایی[2] (QMF) هستند[۹]. فیلتر پایین گذر $h[k]$ با شرط برقراری معادله (۲۱) و (۱۰) دارای خواص زیر در حوزه زمان است[۱۰]:

$$\sum_k h[k]=\sqrt{2} \qquad (۲۳)$$

$$\sum_k h[k]\,h[k-2n]=\delta[n] \qquad (۲۴)$$

$$\sum_k (-1)^k k^m h[k]=0 \quad for \quad m=0,\ldots,M-1 \qquad (۲۵)$$

کمیت M تعداد گشتاور های محوشونده[3] موجک متناظر با ϕ است[۱۱] و [۱۲]. موجک توسط فیلتر پایین گذر $h[k]$ به تنهایی مشخص می شود و فیلتر بالا گذر $g[k]$ از روی آن قابل محاسبه است. اگر $g[k]$ به صورت رابطه زیر محاسبه شود آنگاه $\psi_{j,n}$ یک پایه متعامد برای فضای W_j خواهد بود[۱]:

$$g[k]=(-1)^{1-k}h[1-k] \Leftrightarrow G(\omega)=e^{-j\omega}H^*(\omega+\pi) \qquad (۲۶)$$

که $H(\omega)$ و $G(\omega)$ به ترتیب تبدیل فوریه $h[k]$ و $g[k]$ می‌باشند.

۴- تخمین طیف بوسیله تبدیل موجک

از میان انبوه کاربردهای تبدیل موجک، تخمین سیگنال اصلی از سیگنال نویزی به عنوان یکی از کاربرد های مهم موجک محسوب می شود. در [۶] به منظور اعمال تکنیک های موجک طیف انرژی که تعداد را در کانال انرژی نشان می‌دهد، معادل سیگنال فرض شده است. لذا در این تحقیق هدف استخراج بهترین تخمین طیف از میان طیف نویزی دنبال می شود. فرض کنید

$$y[n]=f[n]+w[n] \qquad (۲۷)$$

که در آن $w[n]$ بیانگر نویز گوسی است. توزیع نویز در تابش های هسته ای پواسون است اما در صورتی‌که میانگین شمارش ها از ۲۰ تجاوز کند، توزیع به سمت توزیع گوسی نرمال میل می کند[۱۳]. دراینجا چون طیف پرتو گامای بدست آمده با میانگین شمارش بیش از ۲۰ بدست آمده است، لذا $w[n]$ تعیین کننده نویز گوسی سفید می‌باشد. $y[n]$ طیف نویزی مشاهده شده و $f[n]$ طیفی است که باید تخمین زده شود. تخمین گر $\hat{y}$ از f یک بردار تصادفی است که با تجزیه y در یک پایه متعامد، و انتخاب یک زیر مجموعه آستانه سازی شده از ضرایب بسط برای باز سازی طیف بدست می آید. این مجموعه از ضرایب بر اساس آستانه ای که روی ضرایب بسط اعمال می‌شود بدست می‌آیند. این تخمین گر های آستانه ای بوسیله داناهو و جانستون معرفی شدند [۲]. در حالت کلی این روش، تخمین گر $\hat{y}$ از f، که هر ضریب نویزی $|\langle y,g_m\rangle|$ را بطور مستقل با تابع غیرخطی $\alpha(x)$ تضعیف می کند، چنین بیان می‌شود:

$$\hat{y}=\sum_{m=0}^{N-1}\alpha\left(|\langle y,g_m\rangle|\right)g_m \qquad (۲۸)$$

که g_m بردار های پایه هستند و

[1] Orthogonal Multiresolution Analysis
[2] Quadrature Mirror Filter
[3] Vanishing moments

$$\alpha(x)=\begin{cases} x, & if\ |x|>\lambda \\ 0, & if\ |x|\le\lambda \end{cases} \qquad (۲۹)$$

این نوع انتخاب تابع $\alpha(x)$ استراتژی آستانه سخت[1] را پیاده سازی می‌کند. که برای بدست آوردن مقدار آستانه λ استراتژی های مختلف وجود دارد اما در حالت کلی مقدار آستانه جهانی که بوسیله داناهو و جانستون معرفی شد، مطابق رابطه زیر بیان می‌شود[۱۴]:

$$\lambda=\sqrt{2\ln(n)}\times\sigma_w \qquad (۳۰)$$

که در آن n طول طیف و σ_w انحراف معیار نویز است.

در پایه موجک، N نمونه نویزی $y[n]$ در پایه موجک گسسته‌ای که روی $[0, N-1]$ تعریف شده است تجزیه می‌شوند:

$$\left[\{\psi_{j,m}\}_{0<j\le J,0\le m<N2^{-j}}, \{\phi_{J,m}\}_{0\le m<N2^{-J}}\right] \qquad (۳۱)$$

در چنین پایه‌ای، تخمین‌گر آستانه سخت (۲۸) به صورت رابطه زیر تبدیل می‌شود:

$$\hat{y}=\sum_{j=1}^{J}\sum_{m=0}^{N2^{-j}}\alpha_h\left(\langle\psi_{j,m},y\rangle\right)\psi_{j,m}+\sum_{m=0}^{N2^{-J}}\alpha_h\left(\langle\phi_{j,m},y\rangle\right)\phi_{j,m} \qquad (۳۲)$$

که $\alpha_h(\langle .,. \rangle)$ همان استراتژی آستانه سخت (۲۹) است.

بنابراین، در این روش طیف نویزی بر روی یک پایه متعامد، که در اینجا یک پایه موجک است، بسط داده می شود سپس ضرایبی از بسط که بزرگتراز یک مقدار آستانه باشند حفظ و بقیه برابر صفر قرار داده می‌شوند. سپس طیف از روی ضرایب باقی مانده باز سازی می شود. برای بهبود عملکرد نویززدایی بوسیله این روش تخمین و طراحی موجک بهینه، لازم است متوسط مربع خطا حداقل گردد:

$$\varepsilon=E\{\|f-\hat{y}\|^2\} \qquad (۳۳)$$

اگر تخمین‌گر $\hat{y}$ بوسیله آستانه سخت محاسبه شده باشد. تخمین‌گر را می‌توان این‌گونه بیان نمود: اگر $\|\langle y,g_m\rangle\|<\lambda$ باشد ضریب مربوطه برابر صفر قرار داده می‌شود که منجر به مربع خطایی برابر با $|\langle f, g_m\rangle|^2$ خواهد شد. و اگر $\|\langle y,g_m\rangle\|\ge\lambda$، این ضریب به همین صورت حفظ می‌شود حفظ می شود که متوسط مربع خطایی برابر با واریانس نویز σ_w^2 ایجاد می‌کند. لذا برای کوچک بودن این خطا لازم است تعداد ضرایبی که اندازه بزرگتر از آستانه دارند کم باشند و سایر ضرایب که اندازه‌ای کمتر از آستانه دارند مجموع مربعات کوچکی داشته باشند.

تقریب غیر خطی طیف در پایه های موجک: تقریب‌های خطی، سیگنال را بر روی M بردار از پیش تعیین شده تصویر می‌کنند. دقت این تقریب را می‌توان با انتخاب M بردار متعامد، متناسب با ویژگی‌های سیگنال، بهبود بخشید. در این بخش عملکرد کلی این تقریب‌های خطی بررسی شده و نتایج آن بر روی پایه‌های موجک اعمال می‌گردد. سیگنال $f\in G$ بوسیله M بردار که بصورت وفقی[2] از یک پایه متعامد $B=\{g_m\}_{m\in N}$ از G انتخاب شده‌اند، تقریب زده می‌شود. فرض کنید f_m تصویر f روی M بردار است که اندیس آنها در I_M باشد:

$$f_M=\sum_{m\in I_m}\langle f,g_m\rangle g_m \qquad (۳۴)$$

خطای تقریب، مجموع ضرایب باقیمانده است:

$$\varepsilon[m]=\|f-f_M\|^2=\sum_{m\notin I_M}|\langle f,g_m\rangle|^2 \qquad (۳۵)$$

برای اینکه این خطا حداقل شود، باید اندیس‌های موجود در I_M متناظر با M برداری باشند که بیشترین شباهت (بزرگترین اندازه ضرب داخلی $|\langle f, g_m\rangle|$) را با سیگنال دارا هستند. این بردارها را می‌توان به مشخصه‌های[3] اصلی f تعبیر نمود. خطای $\varepsilon[m]$ بدست آمده، ضرورتاً کوچکتر از خطای تقریب خطی است که M بردار مستقل از f انتخاب می‌کند.

فرض کنید ضرایب $\{|\langle f, g_m\rangle|\}_{m\in N}$ به صورت نزولی مرتب شده باشند، یعنی $|\langle f, g_{m_k}\rangle|\ge|\langle f, g_{m_{k+1}}\rangle|, m\in N$. بهترین تقریب غیرخطی عبارت است از

$$f_M=\sum_{k=1}^{+\infty}\langle f,g_{m_k}\rangle g_{m_k} \qquad (۳۶)$$

این تقریب را می‌توان بصورت زیر نوشت:

$$f_M=\sum_{m=0}^{+\infty}\alpha_T\left(\langle f,g_m\rangle\right)g_m \qquad (۳۷)$$

و $T=|\langle f,g_{m_M}\rangle|$ که در آن α_T همان رابطه (۲۹) است. خطای غیر خطی عبارت است از

$$\varepsilon[M]=\|f-f_M\|^2=\sum_{k=M+1}^{+\infty}|\langle f,g_{m_k}\rangle|^2 \qquad (۳۸)$$

که این کمیت همان مقداری است که یکی از پارامتر های تعیین کننده در بخش قبل را تشکیل می داد. اگر مقادیر مرتب شده $|\langle f,g_{m_k}\rangle|$ سرعت کاهش[4] زیادی با افزایش k داشته باشند، این خطا با افزایش M به سرعت به صفر میل می کند. این کاهش را می توان با محاسبه نرم l^p ضرایب بسط سیگنال در پایه B بصورت کمی بیان نمود:

$$\|f\|_{B,p}=\left(\sum_{m=}^{+\infty}|\langle f,g_m\rangle|^p\right)^{\frac{1}{p}} \qquad (۳۹)$$

سرعت کاهش $\varepsilon[M]$ را می توان به $\|f\|_{B,p}$ وابسته نمود. نشان داده شده است که اگر $\|f\|_{B,p}<+\infty$ و $p<2$ آنگاه[۱]:

1. Hard Threshold

2 Adaptive
3 Features
4 Decay

$$\varepsilon[M] \le \frac{\|f\|_{B,p}^2}{\frac{2}{p}-1} M^{1-\frac{2}{p}} \qquad (۴۰)$$

۵- طراحی موجک بهینه

با توجه به هدف تعیین شده از طراحی موجک بهینه، لازم است معیاری اختیار شود که حداکثر میزان حذف نویز یا به عبارتی بیشترین افزایش نسبت سیگنال به نویز را ایجاد نماید. برای این منظور باید حداقل خطای باز سازی سیگنال (۳۳) حداقل گردد. حداقل کردن این خطا به معنی ایجاد بهترین تقریب غیرخطی تابع f است. حد بالای خطای این تقریب توسط رابطه (۳۹) تعیین می‌گردد. لذا با کاهش این خطا، می‌توان موجکی بهینه برای بهبود نسبت سیگنال به نویز طراحی نمود. بر این اساس، معیار طراحی موجک را چنین در نظر می‌گیریم:

$$\eta = \min_B \|f\|_{B,p}^2 \qquad (۴۱)$$

پایه B که η را حداقل نماید پایه مطلوب خواهد بود و بیشترین میزان بهبود سیگنال به نویز را خواهد داشت.

پارامتری سازی ضرایب موجک: پایه‌های موجک بصورت منحصر بفرد توسط فیلترهای تجزیه $h[k]$ و $g[k]$ مشخص می‌گردند. از آنجایی‌که فیلتر $g[k]$ از روی $h[k]$ قابل محاسبه است، کافیست تنها فیلتر $h[k]$ برای موجک بدست آید که شرایط ذکر شده در بخش ۳ را دارا بوده و معیار (۴۱) را حداقل نماید. لذا برای طول فیلتر $2N$ مساله طراحی مورد نظر تبدیل به یک مساله بهینه سازی با $2N$ پارامترخواهد شد. اما ضرایب فیلتر پارامتر آزاد نیستند و باید شرایط ذکر شده در روابط (۲۳) تا (۲۵) را اقناع نمایند. در واقع این مساله یک مساله بهینه سازی همراه با محدودیت خواهد بود. برای سهولت حل عددی، می توان توسط پارامترسازی ضرایب فیلتر، مساله را به یک مساله بهینه سازی با N پارامتر آزاد و بدون محدودیت تبدیل نمود [۱۵]. این روش پارامترسازی در این بخش به اختصار شرح داده می‌شود.

فرض کنید $H(z)$ و $G(z)$ تبدیل z دنباله های $h[k]$ و $g[k]$ باشند. در این صورت نمایش چند فازی[1] فیلترهای $H(z)$ و $z^{2(N-1)}G(z)$ را می‌توان چنین نوشت:

$$\begin{bmatrix} H(z) \\ z^{2(N-1)}G(z) \end{bmatrix} = \sqrt{2}E(z^2)\begin{bmatrix} 1 \\ z^{-1} \end{bmatrix} \qquad (۴۲)$$

که ماتریس $E(z)$ ماتریس چندفازی خوانده می‌شود. اگر دنباله $h[k]$ دارای $2N$ عبارت باشد، با استفاده از نتایج [۱۶] می‌توان نشان داد که $E(z)$ دارای نمایش پارامتری زیر است[۱۵]:

$$E(z) = V_{N-1}(z)V_{N-2}(z)\ldots V_1(z)V_0(z) \qquad (۴۳)$$

که

$$V_0(z) = \begin{bmatrix} -\cos\beta_0 & \sin\beta_0 \\ \sin\beta_0 & \cos\beta_0 \end{bmatrix} \qquad (۴۴)$$

$$V_n(z) = I + (z-1)v_n v_n^T, \quad 1 \le n \le N-1 \qquad (۴۵)$$

در رابطه بالا v_n یک بردار حقیقی ۱×۲ با طول واحد است($v_n v_n^T = \|v_n\|^2 = 1$).بدون از دست دادن کلیت مساله، می‌توان فرض نمود که v_n به فرم زیر باشد:

$$v_n = \begin{bmatrix} \cos\beta_n \\ \sin\beta_n \end{bmatrix} \qquad (۴۶)$$

رابطه‌های (۴۲) و (۴۳) یک پارامتری سازی کامل برای g و h هستند، چراکه هر دنباله $h[k]$ و $g[k]$ متناظر آن که شرط فیلتر باز سازی کامل[2] را اقناع کند را می‌توان بصورت (۴۲) و (۴۳) بیان نمود. در رابطه (۴۲) فیلتر $z^{2(N-1)}G(z)$ بجای خود $G(z)$ بکار رفته تا شرط $g[k] = (-1)^n h[1-k]$ را ارضا نماید[۱۵]. رابطه (۴۳) همچنین نشان می‌دهد که دنباله $h[k]$ به طول $2N$ و با شرط تعامد را می توان توسط N پارامتر آزاد بیان نمود. اما با اضافه کردن شرط regularity(رابطه(۲۵))، با فرض کمترین مقدار ممکن برای M یعنی ۱، تعداد پارامترهای آزاد یکی کاهش می‌یابد. نشان داده شده است که موجک $\psi(x)$ حداقل یک گشتاور محو شونده دارد، یعنی حداقل صفری از مرتبه۱ در $\omega = \pi$ دارد اگر و تنهااگر $\beta_0 = 3\pi/4$ [۱۵]. بر این اساس می‌توان طراحی موجک $\psi(x)$ با گستره $2N-1$ را بوسیله $N-1$ پارامتر انجام داد. استفاده از این روش پارامتری سازی منجر به کاهش تعداد مجهولات و ساده‌تر شدن فرآیند طراحی می‌گردد. به خصوص از آنجایی‌که مساله مورد بررسی دارای حل تحلیلی نیست و با روش‌های بهینه‌سازی حل می‌شود، کاهش تعداد مجهولات تاثیر بسزایی در حل مساله دارد.

طراحی و تعیین ضرایب فیلتر مربوط به موجک بهینه: همان طور که در ابتدای این بخش شرح داده شد، در روش پیشنهادی، به دنبال پایه B می‌گردیم به گونه‌ای که معیار η که در رابطه (۴۱) داده شده حداقل‌گردد. از آنجایی که این پایه یک پایه موجک است توسط فیلتر $h[k]$ بطور کامل توصیف می‌گردد. لذا یافتن این پایه بهینه، معادل یافتن $h[k]$ متناظر آن است. یعنی فیلتر $h[k]$ باید به گونه‌ای انتخاب شود که پایه B متناظر با آن معیار η را حداقل نماید. برای حصول اطمینان از برقراری شروط داده شده در روابط (۲۳) تا (۲۵) از پارامتری سازی ضرایب $h[k]$ که در بخش قبل بررسی شد استفاده می‌گردد. به این ترتیب یافتن فیلتر بهینه با شرایط معین، تبدیل به یافتن

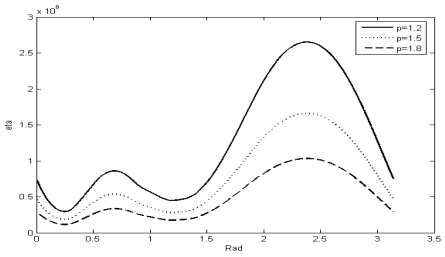
[1] Polyphase

[2] Perfect Reconstruction Quadrature Mirror Filter(PR QFM)

پارامترهای β بدون محدودیت می‌گردد. برای طراحی موجک بهینه با طول مشخص L، معیار η را چنین بازنویسی می‌کنیم:

$$\eta = \min_{\{\beta_i\}} \|f\|_{B,p}^2 \qquad (۴۷)$$

که β_iها همان پارامترهای β در رابطه (۴۶) هستند و تعداد عناصر مجموعه $\{\beta_i\}$ به طول L (طول موجک) بستگی دارد. برای حداقل‌سازی این معیار می‌توان از روش های بهینه سازی مختلفی استفاده نمود. در این مقاله از روش کمترین مربعات برای یافتن حداقل η استفاده شده است. چنان‌که از رابطه (۴۰) برمی‌آید، مقدار کمیت p تا زمانی که در بازه مشخص شده باشد، اثری برروی فرآیند بهینه سازی نخواهد داشت. این بدان معناست که حداقل‌های رابطه (۴۷) وابسته به p نیستند. این مطلب به صورت شهودی برای فیلتری به طول ۴ و ۳ مقدار مختلف p و برای بهینه سازی طیف ^{137}Cs نشان داده شده است.



شکل ۴: نمودار منحنی معیار بهینه η برای سه مقدار مختلف p

۶- نسبت سیگنال به نویز

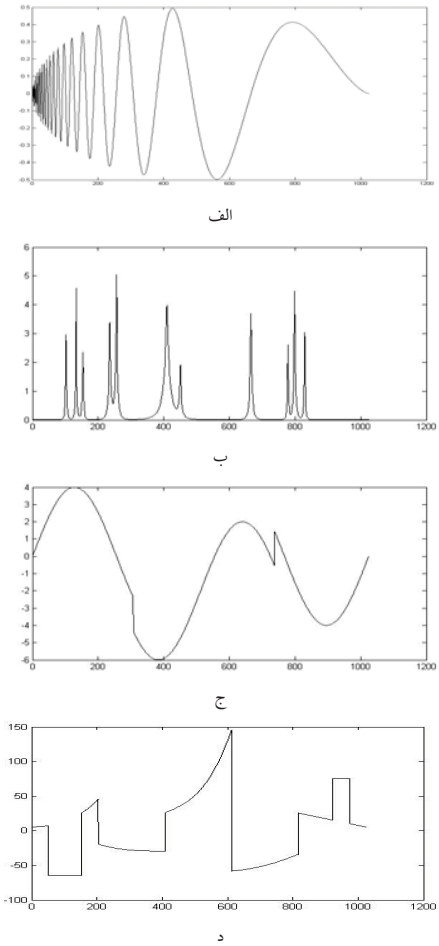
نسبت سیگنال به نویز به صورت رنج میان سطح پایینی نویز و سطح اسمی سیگنال تعریف می شود. تعریف لگاریتمی نسبت سیگنال به نویز به صورت فرمول(۴۸) می باشد.

$$SNR = 10\log_{10}(\frac{P_{signal}}{P_{noise}}) = 20\log_{10}(\frac{A_{signal}}{A_{noise}}) \qquad (۴۸)$$

که P نماینده متوسط توان سیگنال و A نماینده دامنه(rms) است.

۷- نتایج شبیه سازی

سیگنال‌های استاندارد: برای بررسی چگونگی عملکرد این روش، ابتدا عملکرد آن را برروی چند نمونه سیگنال استاندارد بررسی شده است. سیگنال‌های مورد استفاده سیگنال‌های Doppler، Bumps، و Piecewise Polynomial و HeaviSine هستند. هر کدام از این سیگنال‌های انتخاب شده دارای ویژگی‌های متغییر با زمان هستند و وجه تشابه زیادی در پردازش سیگنال‌های واقعی دارند [۲]. ابتدا برای طول فیلتر مشخص L، با روش بالا موجکی بهینه برای هر یک از سیگنال‌ها طراحی می‌گردد. سپس میزان بهبود نسبت سیگنال به نویز بررسی می‌گردد. برای مقایسه و نشان دادن برتری روش پیشنهادی، این سیگنال‌ها بوسیله یک موجک استاندارد نیز نویززدایی شده‌اند. از آنجایی که موجک‌های خانواده Daubechies به ازای طول فیلتر معین، بیشترین تعداد گشتاور محو شونده را دارا هستند، لذا دارای بیشترین قابلیت نویززدایی به ازای طول مشخص هستند و به همین جهت این خانواده به عنوان مرجعی برای مقایسه نتایج انتخاب شده است. سیگنال‌های مورد بررسی در شکل ۵ نشان داده شده‌اند.



شکل ۵: سیگنال های آزمایشی برای بررسی کیفیت عملکرد روش طراحی موجک
الف) Doppler ب) Bumps ج) HeaviSine د) Piecewise Polynomil

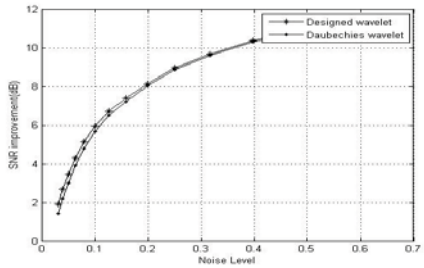
شایان ذکر است که رابطه بهینه‌سازی (۴۷) باید برای مقیاس‌هایی که نویززدایی در آنها انجام می‌گیرد استفاده شود. این مقیاس‌ها به نوع سیگنال و میزان هموار بودن آن بستگی دارد. برای سیگنال‌هایی که مولفه‌های فرکانس بالای زیادی ندارند می‌توان مقیاس‌ها را کوچک‌تر انتخاب نمود. برای سیگنال‌هایی که تغییرات سریع دارند و در نتیجه حاوی اطلاعات فرکانس بالا هستند، عمل بهینه سازی در مقیاس‌های بالاتر انجام می‌گیرد. برای ارائه مثالی از چگونگی عملکرد این روش، طول فیلتر برابر ۶ در نظر گرفته شده که منجر به دو پارامتر آزاد می‌گردد. موجک هم طول استاندارد آن از خانواده Daubechies موجک db3 خواهد بود. پارامتر های β_1 و β_2 به همراه مقدار معیار η برای

موجک Daubechies (η_{db3}) و موجک بهینه طراحی شده ($\eta_{optimal}$) در جدول ۱ نمایش داده شده‌اند.

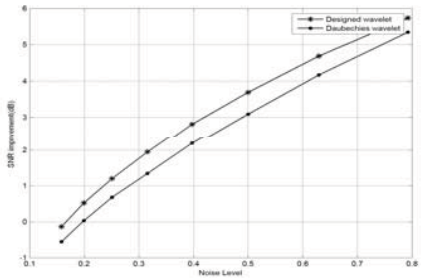
جدول ۱: مقادیر پارامترهای بهینه β_1 و β_2 برای سیگنال های مورد بررسی به ازای طول فیلتر ۶.

سیگنال های استاندارد	پارامتر بهینه $\beta_1(rad)$	پارامتر بهینه $\beta_2(rad)$	$\eta_{optimal}$	η_{db3}
Doppler	۱/۲۰۵۸	۱/۶۸۶۸	۳۶/۱۲۲۶	۴۵/۸۴۱۹
Bumps	۱/۰۵۲۵	۲/۹۵۲۵	۱۰/۸۴۱۱	۱۱/۳۶۷۳
HeaviSine	۱/۲۴۲۴	۳/۰۸۳۱	۱۵/۰۵۸۳	۲۲/۶۳۹۲
Piecewise Polynomial	۱/۵۴۱۵	۱/۵۷۰۴	۸/۴۰۸۸	۱۸/۳۶۹۴

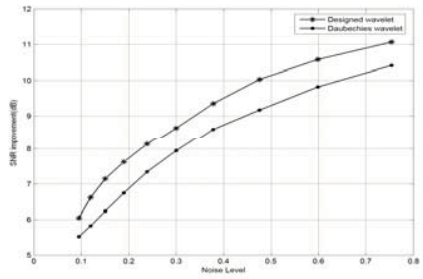
در شکل ۶ عملکرد موجک طراحی شده با روش پیشنهادی با عملکرد موجک Daubechies مقایسه شده است.



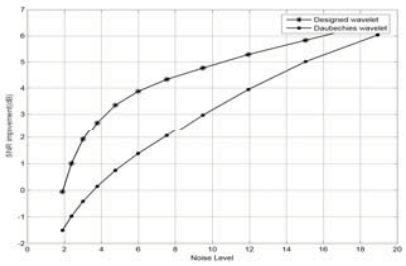
الف



ب



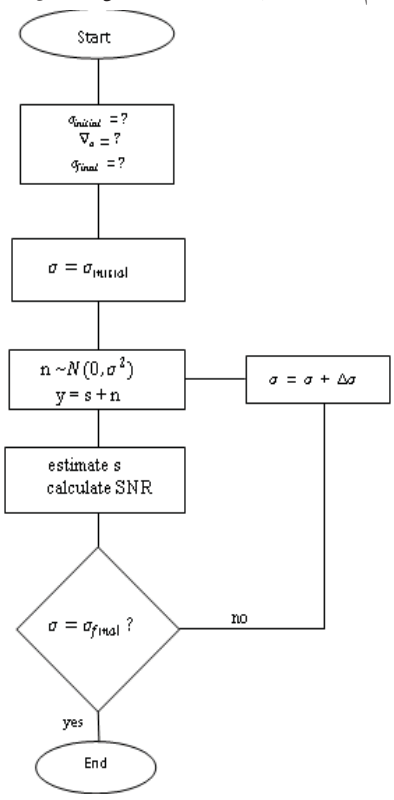
پ



ت

شکل ۶: نمودار مقایسه میزان بهبود نسبت سیگنال به نویز بوسیله موجک بهینه طراحی شده و موجک استاندارد db3 به روش نویز زدایی Visu برای سیگنال های استاندارد الف) نمودار مقایسه میزان بهبود نسبت سیگنال به نویز برای سیگنال Doppler در سطوح مختلف نویز ب) نمودار میزان بهبود نسبت سیگنال به نویز برای سیگنال Bumps در سطوح مختلف نویز پ) نمودار میزان بهبود نسبت سیگنال به نویز برای سیگنال Heavisine در سطوح مختلف نویز ت) نمودار میزان بهبود نسبت سیگنال به نویز برای سیگنال Piecewise Polynomial در سطوح مختلف نویز

الگوریتم مقایسه عملکرد دو موجک در شکل ۷ نمایش داده شده است.



شکل ۷: فلوچارت محاسبه بهبود نسبت سیگنال به نویز

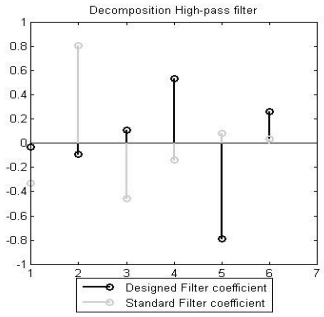
طیف‌انرژی رادیوایزوتوپ‌های (^{60}Co و ^{137}Cs): نحوه کار مشابه بخش قبل است. یعنی ابتدا برای طول L، موجک بهینه طراحی می شود و سپس میزان بهبود نسبت سیگنال به نویز طیف انرژی رادیو ایزوتوپ-

های گسیلنده پرتو گاما (^{60}Co و ^{137}Cs) بوسیله روش پیشنهادی و موجک استاندارد (db3) مقایسه می‌گردند. مقادیر پارامترها و معیار η برای بهینه‌سازی در مقیاس‌های ۶، ۷، ۸، ۹ در جدول نمایش داده شده است.

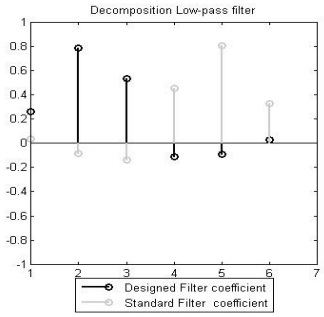
جدول ۲: مقدار پارامتر بهینه یرای طیف مورد بررسی به ازای طول فیلتر ۶.

طیف پرتو گاما	پارامتر بهینه $\beta_1(rad)$	پارامتر بهینه $\beta_2(rad)$	$\eta_{optimal}$	η_{db3}
^{60}Co	۰/۴۶۲۲	۳/۰۲۳۷	۴۴۴۲۰۰	۶۱۰۹۸۰
^{137}Cs	۱/۱۸۵۹	۱/۶۸۵۴	۲۱۳۲۱۶۰۰	۲۹۳۲۷۰۴۰

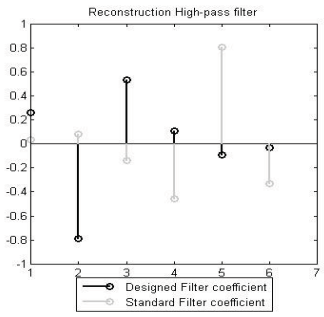
شکل ۸ به عنوان نمونه‌ای از فیلتر بهینه طراحی شده متناسب با ویژگی‌های طیف ^{60}Co (ضرایب فیلترآینه‌ای چهارتایی مربوط به موجک بهینه طراحی شده) همراه با ضرایب فیلتر مربوط به موجک استاندارد db3 نمایش داده شده است.



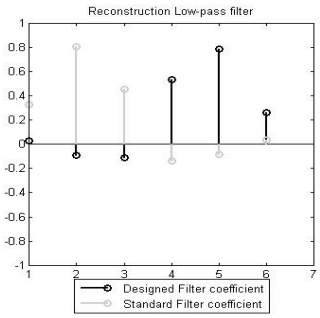
الف



ب



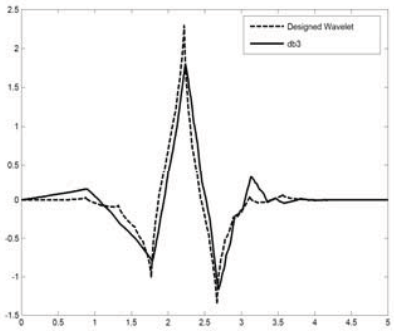
ج



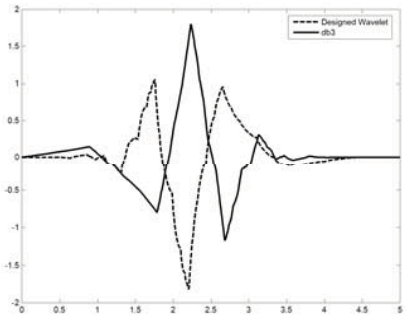
د

شکل ۸ : ضرایب فیلتر آینه‌ای چهارتایی موجک طراحی شده و موجک استاندارد db3 به ازای طول فیلتر شش برای طیف ^{60}Co الف) ضرایب فیلتر بالاگذر تجزیه ب) ضرایب فیلتر پایین گذر تجزیه ج) ضرایب فیلتر بالاگذر بازسازی د) ضرایب فیلتر پایین گذر بازسازی

در شکل‌های ۹ و ۱۰ به ترتیب موجک های بهینه طراحی شده برای طیف رادیوایزوتوپ‌های گسیلنده پرتوگاما سزیم ۱۳۷ (^{137}Cs) وکبالت ۶۰ (^{60}Co) و موجک استاندارد هم طول آنها به ازای طول ۶ نمایش داده شده است.

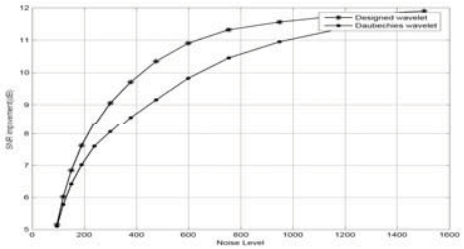


شکل ۹: موجک طراحی شده برای طیف ^{137}Cs و موجک استاندارد برای طول فیلتر شش(L=6)

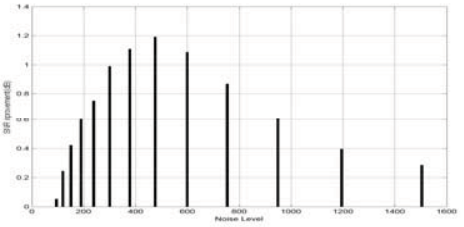


شکل ۱۰: موجک طراحی شده و و موجک استاندارد برای طول فیلترشش(L=6)

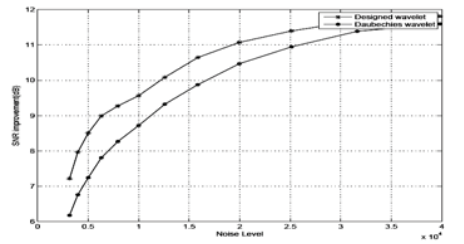
در شکل‌های ۱۳-۱۱ نیز میزان بهبود نسبت سیگنال به نویز بوسیله موجک طراحی شده و موجک استاندارد با هم مقایسه شده‌اند.



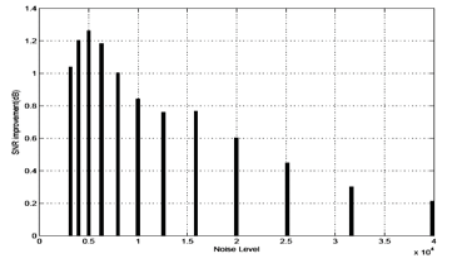
شکل ۱۱: نمودار مقایسه میزان بهبود نسبت سیگنال به نویز بوسیله موجک بهینه طراحی شده و موجک استاندارد db3 به روش نویز زدایی Visu برای طیف پرتو گاما (^{60}Co) در سطوح مختلف نویز



شکل ۱۲: نمودار میزان افزایش نسبت سیگنال به نویز بوسیله موجک بهینه طراحی شده و موجک استاندارد db3 در سطوح مختلف نویز با روش نویز زدایی Visu

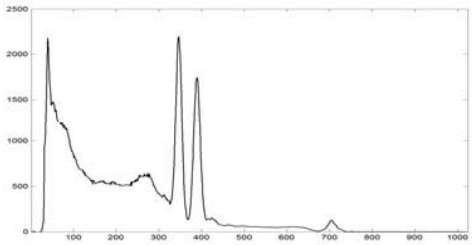


شکل ۱۳: نمودار مقایسه میزان بهبود نسبت سیگنال به نویز بوسیله موجک بهینه طراحی شده و موجک استاندارد db3 به روش نویززدایی Visu برای طیف پرتو گاما (^{137}Cs) در سطوح مختلف نویز

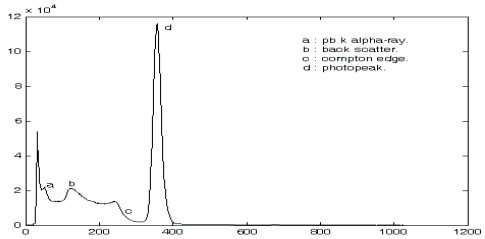


شکل ۱۴: نمودار میزان افزایش نسبت سیگنال به نویزبوسیله موجک بهینه طراحی شده و موجک استاندارد db3 در سطوح مختلف نویز به روش نویز زدایی Visu

در شکل ۱۵ و ۱۶ طیف انرژی رادیو ایزوتوپ‌های گسیلنده پرتو گاما (^{60}Co و ^{137}Cs) نمایش داده شده است.

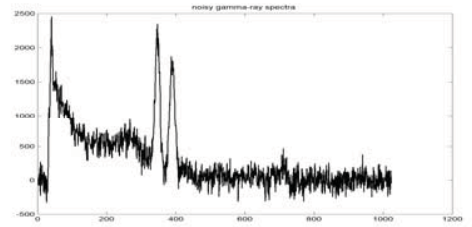


شکل ۱۵: طیف انرژی رادیو ایزوتوپ ^{60}Co

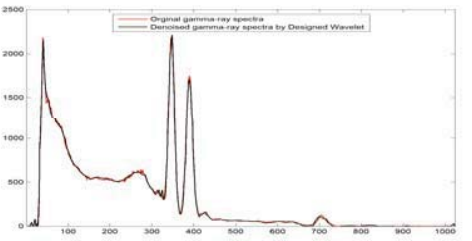


شکل ۱۶: طیف انرژی رادیو ایزوتوپ ^{137}Cs

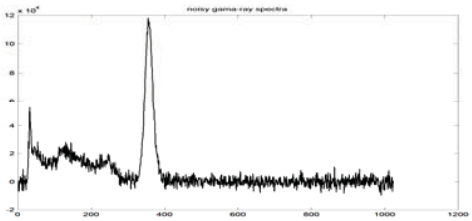
شکل ۱۷ یک نمونه از طیف نویزی و نویززدایی شده ^{60}Co و ^{137}Cs با استفاده از روش بهینه سازی پیشنهاد شده (اعمال موجک بهینه طراحی شده جهت کاهش نویز از طیف نویزی) را نمایش می‌دهد.



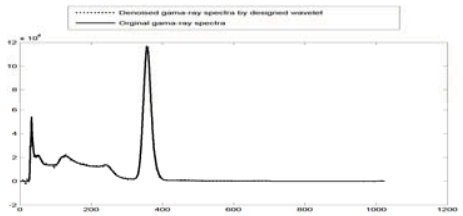
الف



ب



پ



ت

شکل ۱۷: طیف‌های اصلی، نویزی و نویززدایی شده بوسیله موجک بهینه. الف) طیف نویزی ^{60}Co ب) طیف اصلی و نویززدایی شده ^{60}Co پ) طیف نویزی ^{137}Cs ب) طیف اصلی و نویززدایی شده ^{137}Cs

۷- نتیجه گیری

در این مقاله روشی برای طراحی موجک بهینه متناسب با سیگنال مورد پردازش ارائه گردید. معیار انتخابی برای بهینه بودن موجک ایجاد بهترین تقریب غیرخطی سیگنال در پایه موجک انتخاب شده است. نشان داده شد که این معیار منجر به بیشترین میزان بهبود نسبت سیگنال به نویز می‌گردد. نتایج شبیه سازی بر روی سیگنال‌های استاندارد و نیز طیف‌های ^{60}Co و ^{137}Cs که شناسایی برهمکنش‌های غالب در آنها در محدوده‌ی انرژی های تشخیصی در پزشکی هسته‌ای مهم است نشان‌گر برتری روش پیشنهادی بر روش‌های استاندارد است.

با توجه به نتایج بدست‌آمده از بخش ۷ ملاحظه می‌شود که برای طیف ^{60}Co روش پیشنهادی از ۰/۱ دسی بل (۱/۵ درصد افزایش سیگنال به نویز) تا ۱/۱۹ دسی بل(۱۱/۵ درصدافزایش سیگنال به نویز) و ^{137}Cs از ۰/۲۱ دسی بل(۱/۷درصد افزایش نسبت سیگنال به نویز) تا ۱/۲ دسی‌بل (۱۵درصدافزایش نسبت سیگنال به نویز) بهتر از موجک استاندارد db3 عمل می‌کند. بنابراین در این مقاله موجک‌های بهینه‌ای برای طیف‌های ^{60}Co و ^{137}Cs جهت شناسایی برهمکنش‌های غالب در آنها طراحی شد. بنابراین طراحی موجک های بهینه برای طیف‌های مختلف رادیوایزوتوپ‌های گسیلنده پرتو گاما می تواند راهکاری جهت بهبود شناسایی برهمکنش‌های غالب و تفکیک مشخصات آنها باشد.

مراجع

[1] S. Mallat, A Wavelet Tour of Signal Processing. Academic Press, 1998.

[2] D. Donoho and I. Johnstone, "Ideal spatial adaptation via waveletshrinkage," Biometrika, vol. 81, pp. 425–455, December 1994.

[3] A. Das, U. B. Desai, and P. P. Vaidya, "Search for optimal basis forsignal denoising in the space of n-tap wavelets," in Symposium on SignalProcessing and its Applications, August 2001.

[4] G. Shi, A. Ding, and L. Jiaom, "A new approach for constructingmatch wavelet to signal detection," in IEEE International Conferenceon Communications, Circuits and Systems, June 2004.

[5] J. O. Chapa and R. M. Rao, "Algorithms for designing wavelets to matcha specific signal," IEEE Trans. Signal Processing, vol. 48, no. 12, pp.3395–3406, December 2000

[6] C. J. Sullivan, "Generation of customized wavelets for the analysis of gama-ray spectra," Published by Elsevier B.V. Nucl. Instr. and Meth, Vol. 579, pp. 275-278, 2007.

[7] Q. Zhang, "Denoising of gamma-ray signals by interval-dependent thresholds of wavelet analysis," Meas. Sci. Technol, Vol. 17, pp. 731-735, 2006.

[8] C.J. Sullivan, "Wavelet analysis of gamma-ray spectra. IEEE, Nuclear Science Symposium Conference Record, Vol. 1, pp. 281-286, 2004.

[9] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Pattern Anal Machine Intell, vol. 11, No. 7, pp. 674–93, 1989.

[10] C. S. Burrus, R. A. Gopinath, and H. Guo, "Introduction to Wavelet and Wavelet Transforms," Prentice Hall, pp. 31-32, 1998.

[11] I. Daubechies, "Orthonormal bases of compactly supported wavelets," Commun. Pure Appl. Math, vol. 41, pp. 909-996, 1988.

[12] G. Strang, "Wavelet and dilation equations: A brief introduction," SIAM Rev, vol. 31, No. 4, pp. 611-642, December 1989.

[13] N. T soulfanidis, Measurement Detection of Radiation, pp. 20-21, 1995.

[14] D.L. Donoho, "Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and noisy data," in Proc. of Symposia in Applied Mathematics, vol. 47, pp. 173-205, AMS, 1993.

[15] H. Zou and A. H. Tewfik, "Parametrization of compactly supported orthonormal wavelets," IEEE Trans.Signal Processing, Vol. 41, No. 3, 1993.

[16] P. Vaidyanathan, "Multirate digital filters, filter bank, polyphase, networks and applications: A tutorial," Proc. IEEE, vol. 78, No. 1, pp. 56-93, 1990.

مجله کنترل
جلد ۳، شماره ۱، بهار ۱۳۸۸، صفحه ۲۸-۲۳

کاهش زمان نشست در کنترل فیدبک تاخیری سیستم آشوبی روسلر با سنکرون‌کننده PI

حسن فاتحی مرج [۱]، رجب اصغریان [۲]، ناصر پریز [۳]

[۱] دانشجوی دکترای برق-کنترل ،گروه برق، دانشکدهٔ مهندسی، دانشگاه فردوسی مشهد، hassan.fatehi@gmail.com

[۲] استاد، گروه برق، دانشکدهٔ مهندسی، دانشگاه فردوسی مشهد، asghrian@ferdowsi.um.ac.ir

[۳] دانشیار، گروه برق، دانشکدهٔ مهندسی، دانشگاه فردوسی مشهد، n-pariz@um.ac.ir

چکیده: در این مقاله یک روش ساده و کارا جهت کاهش زمان نشست در روش فیدبک تاخیری کنترل آشوب ارائه گردیده است. این روش برپایه سنکرون‌سازی سیستم آشوبی با داده‌های ذخیره‌شده توسط یک سنکرون‌کننده PI می‌باشد. برخی مزایای این روش بیان گردیده و شبیه‌سازی عددی برای سیستم روسلر با و بدون پارامترهای نامعین آورده شده است.

کلمات کلیدی: کنترل آشوب فیدبک تاخیری، زمان نشست، سنکرون‌سازی، بازسازی UPO، سیستم روسلر.

Abstract: In this paper based on synchronizing the chaotic system with the stored data by a PI synchronizer, a simple and efficient technique is presented to decrease settling time in the delayed feedback chaos control method. Some advantages of the presented technique are described and numerical simulations for the Rössler system with and without uncertain parameters are presented.

Keywords: Delayed feedback chaos control, Settling time, synchronization, UPO reconstruction, Rössler system.

۱- مقدمه

پیشکسوتان کنترل آشوب آت[1]، گربوگی[2] و یورک[3] می‌باشند که روش کنترل آشوب OGY را بنا نهادند. در این روش یک مدار دوره‌ای ناپایدار[4] (UPO) محاط در جذب کننده آشوبی با خطی‌سازی نگاشت پوانکاره پایدار می‌شود[۱]. پس از آن کنترل آشوب در بسیاری از علوم موضوعی بحث برانگیز شد و روش‌های مختلفی برای آن پیشنهاد گردید[۲، ۳].

البته روش OGY در اغلب موارد جهت پایدارسازی هدف[5] به زمان زیادی نیازمند است [۴] (ما این زمان را زمان نشست[6] می‌خوانیم) که برخی تکنیک‌ها جهت بهبود این مشکل پیشنهاد شده‌اند [۴، ۵]. مرجع [۴] از تقریب سیستم آشوبی در نزدیکی مانیفولد[7] پایدار هدف استفاده کرده است و به سیستم، نیروی کمی در جهت مانیفولد ناپایدار اعمال کرده تا بدینوسیله زمان نشست را کاهش دهد.

مرجع [۵] یک روش n-مرحله‌ای پیشنهاد کرده که n بعد سیستم می‌باشد. در این روش هر شرایط اولیه در ناحیه کنترل‌پذیر به جای زیرمجموعه پایدار، مستقیما به سمت مدار هدف فرستاده می‌شود.

روش کنترل فیدبک تاخیری (DFC) ابتدا توسط پایراگس پیشنهاد گردید [۶]. این روش ساده که تحت عناوین روش پایراگس و روش تاخیر زمانی نیز شناخته می‌شود، برپایه اختلاف حالات جاری و حالات یک یا چند دوره قبل قرار دارد (شکل ۱) و به طور موفقیت‌آمیزی در بسیاری سیستم‌های آشوبی به کار گرفته شده است [۷-۹]. البته در برخی کاربردها این روش نیز نیازمند زمان زیادی است. از آنجا که روش OGY بر پایه نگاشت پوانکاره قرار دارد و اصولاً یک روش گسسته است تکنیک‌های ارائه شده برای آن مناسب روش تاخیر زمانی در سیستم‌های پیوسته که موردنظر این مقاله است نمی‌باشند. تا آنجا که ما اطلاع داریم،جهت کم کردن زمان نشست در روش فیدبک تاخیری کاری انجام نشده است. تنها مرجع [۱۰] متذکر شده است که نویز خارجی می‌تواند سرعت کنترل را در کنترل فیدبک تاخیری نگاشت

[1] Ott
[2] Grebogi
[3] York
[4] Unstable Periodic Orbit
[5] Target
[6] Settling Time
[7] Manifold

لاجیستیک[1] (که یک سیستم گسسته است) زیاد کند و در مرجع [۱۱] که از فیدبک تاخیری بر پایه مود لغزشی[2] در کنترل سیستم‌های آشوبی تاخیری استفاده کرده است نتیجه کنترل مود لغزشی یک نقطه تعادل در سیستم مکی گلاس[3] با روش تاخیر زمانی مقایسه شده است که سرعت روش ارائه شده بر پایه مود لغزشی بیشتر بوده است. البته روش ارائه شده در آن به عنوان طرحی برای کم کردن زمان نشست معرفی نشده است و همچنین سیستم مورد نظر آن و روش ارائه شده در آن کاملاً با این مقاله تفاوت دارند. لازم به ذکر است که مرجع [۱۲] شبکه‌های عصبی را برای روش‌های OGY و پایراگس آموزش داده است و در [۱۳] تخمین فازی برای روش‌های OGY و پایراگس استفاده شده است که در هر دو زمان نشست روش OGY کاهش پیدا کرده است اما زمان نشست روش تاخیر زمانی تغییر نیافته است. همچنین لازم به ذکر است که هیچکدام از مقالاتی که به نحوی به مساله زمان رسیدن به هدف کنترلی یا سرعت کنترل اشاره کرده‌اند به تعریف زمان نشست در کنترل آشوب نپرداخته‌اند و بنابراین حتی مقالاتی که جهت کم کردن این زمان در روش OGY ارائه شده اند از ارائه نتایج کمی مناسب بی‌بهره بوده‌اند.

در این مقاله یک تکنیک جهت کم کردن زمان نشست در روش DFC برپایه سنکرون‌سازی سیستم با داده‌های ذخیره شده ارائه گردیده است. این روش ساده بوده و نیازمند اضافه کردن اندازه‌گیر یا ورودی کنترلی نیست و همچنین این روش سیستم با پارامترهای نامعلوم را بهتر از روش تاخیر زمانی خطی کنترل می‌کند.

۲- تعریف مساله و مقدمات ریاضی

سیستم‌های آشوبی زمان-پیوسته nبعدی بیان‌شونده با معادلات زیر را در نظر بگیرید.

$$\dot{X}(t) = f(X(t),U(t),t) \quad (X \in R^n, U \in R^m) \qquad (۱)$$
$$Y(t) = g(X(t)) \quad (Y \in R^r)$$

که $X(t)$ حالات سیستم، $U(t)$ ورودی‌های کنترل و $Y(t)$ خروجی‌های سیستم می‌باشند.

تعریف۱: کنترل فیدبک تاخیری (DFC) [۱۴]

کنترل فیدبک تاخیری (کنترل تاخیرزمانی) برای سیستم‌های پیوسته آشوبی، پایدارسازی مدار دوره‌ای ناپایدار $\overline{X}(t)$ با پریود T که در سیستم کنترل نشده (یعنی سیستم (۱) با $U(t)=0$) ظاهر می‌شود، تنها با استفاده از فیدبک بر اساس

$$Y(\tau) - Y(\tau - T) = g(X(\tau)) - g(X(\tau - T)) \; (\tau < t)$$

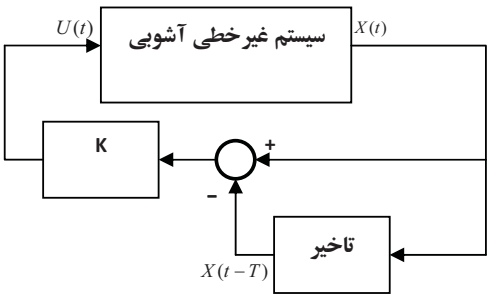
در زمان t می‌باشد.

جهت سادگی فرض می‌شود که $Y(t) = X(t)$. همچنین فرض می‌شود که رابطه (۱) می‌تواند به شکل زیر نوشته شود:

$$\dot{X}(t) = f(X(t),t) + U(t) \qquad (۲)$$

پس از فرض‌های بالا و در نظر گرفتن روش فیدبک تاخیری به صورت خطی (که بسیار معمول است و همان روش اولیه ارائه شده توسط پایراگس می‌باشد) سیستم حلقه بسته مورد نظر به شکل زیر می‌باشد (شکل ۱ را ببینید):

$$\dot{X}(t) = f(X(t),t) + K(X(t) - X(t-T)) \qquad (۳)$$



شکل ۱: دیاگرام بلوکی روش کنترل فیدبک تاخیری

تعریف۲: زمان نشست

زمان نشست t_{set} به صورت $t_{set} = t_r - t_s$ تعریف می‌شود که در آن t_s زمان شروع عمل کنترل و $t_r > t_s$ اولین زمانی است که مسیرهای سیستم به ناحیه‌ای حول UPO رسیده و پس از آن در آن ناحیه باقی می‌مانند (یا همه خطاهای بین حالات سیستم و مقادیر متناظر از UPO کمتر از حداکثر قابل قبول می‌مانند

$$\|E\| = \|X(t) - \overline{X}(t)\| < E_{\max} \quad \text{for} \quad t \geq t_r.$$

زمان نشست t_{set} طولانی یا نه به اندازه کافی کوتاه خوانده می‌شود هرگاه: $t_{set} > pT$ که در آن p یک عدد حقیقی (ضریب) و T پریود UPO می‌باشد.

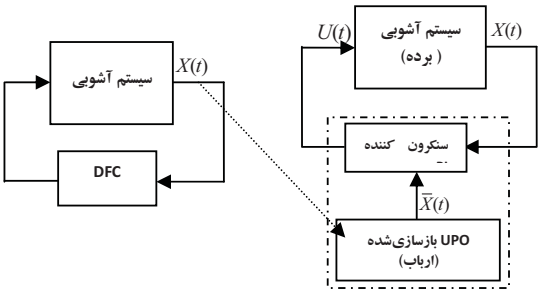
نکته۱: عمل کنترلی پس از شروع به کار سیستم کنترل‌شده و یا پس از اغتشاشی که سیستم را از نزدیکی UPO دور کند، شروع می‌شود.

نکته۲: در این مقاله فرض شده است که کنترل‌کننده تاخیرزمانی می‌تواند سیستم را کنترل کند اما زمان نشست آن طولانی است ($t_{set} > pT$) و هدف ما این است که این زمان را به اندازه کافی کوتاه کنیم ($t_{set} < pT$).

[1] Logistic Map
[2] Sliding Mode
[3] Mackey Glass

۳- تکنیک کم‌کردن زمان نشست با سنکرون‌سازی

شکل (۲) بلوک دیاگرام تکنیک کم‌کردن زمان نشست را نشان می‌دهد.



شکل ۲ : دیاگرام بلوکی تکنیک کاهش زمان نشست

جهت کم کردن زمان نشست رویه زیر انجام می‌شود:

۱.اجرای سیستم کنترل‌شده با روش تاخیر زمانی برای یک بار تا زمانی که مسیرهای سیستم به UPO مورد نظر برسند و محاسبه زمان نشست آن روش.

۲.ذخیرهٔ حداقل یک دوره از حالات (خروجی‌ها) سیستم کنترل‌شده پس از رسیدن سیستم به UPO.

۳.بازسازی[1] UPO با تکرار یا مدل‌سازی داده‌های ذخیره‌شده.

۴.سنکرون‌کردن سیستم آشوبی (برده[2]) با UPO بازسازی شده (ارباب[3]) جهت کنترل آشوب (به جای روش تاخیرزمانی).

از آنجا که فرض بر این است که روش تاخیر زمانی می‌تواند آشوب را کنترل کند و فقط زمان نشست آن زیاد است، UPO بازسازی‌شونده با استفاده از داده‌های ذخیره‌شده به خوبی بر UPO واقعی موردنظر منطبق خواهد بود.

نکته ۳: از آنجا که روش تاخیر زمانی در واقع سنکرون‌کردن حالات (خروجی‌ها) با مقادیر تاخیریافته آنها است و داده‌های ذخیره شده را می‌توان مقادیر تاخیریافته برای دوره بعد تلقی کرد (در روش تاخیر زمانی حالات یا خروجی‌ها در دوره بعد با این داده‌ها سنکرون می‌شوند)، حالات (خروجی‌های) سیستم آشوبی را می‌توان به راحتی با داده‌های ذخیره‌شده سنکرون کرد. به عبارت دیگر پایداری سیستم حلقه بسته با روش DFC به عنوان فرض مساله درنظر گرفته شده است و داده‌های ذخیره شده یک دوره از سیستم دوره‌ای شده با روش DFC می‌باشند پس در واقع در روش DFC خروجی سیستم در دوره‌های بعد فقط با استفاده از بهره k با این داده‌ها سنکرون می‌شوند (نام دیگر روش DFC روش خود سنکرون کننده است) بنابر این تنها با استفاده از کنترل‌کننده تناسبی می توان به سنکرون‌سازی رسید و از آنجا که کنترل‌کننده PI در برگیرنده کنترل‌کننده P است پس با آن نیز می‌توان سنکرون‌سازی را انجام داد. از دیدگاه دیگر پایداری سیستم حلقه بسته با تکنیک ارائه‌شده چنین بدست می آید که: از نظر تئوری اگر سیستم آشوبی دقیقا روی UPO قرار گیرد روی آن باقی می‌ماند و از آنجا که طبق فرض DFC سیستم را پایدار می‌کند آنرا روی UPO می‌برد پس می‌توان داده‌های ذخیره شده را خروجی سیستمی مشابه دانست که روی UPO است. از این دیدگاه کنترل آشوب با سنکرون کردن سیستم آشوبی با سیستمی مشابه که روی UPO قرار دارد انجام می‌شود.

بیان ریاضی این دیدگاه بدین صورت است که با توجه به رابطه (۲) سیستم آشوبی (برده) و سیستم بر اساس داده‌های ذخیره‌شده (ارباب) را به ترتیب می‌توان با معادلات (۴) و (۵) بیان کرد:

$$\dot{X}_s(t) = f(X_s(t),t) + U_s(t) \qquad (۴)$$

$$\dot{X}_m(t) = f(X_m(t),t) + U_m(t) \qquad (۵)$$

با توجه به اینکه طبق فرض سیستم ارباب توسط روش تاخیر زمانی کنترل شده است می‌توان نوشت:

$$\dot{X}_m(t) = f(X_m(t),t) + K(X_m(t) - X_m(t-T)) \qquad (۶)$$

و $X_m(t) = X_m(t-T)$ بنابراین:

$$\dot{X}_m(t) = f(X_m(t),t) \qquad (۷)$$

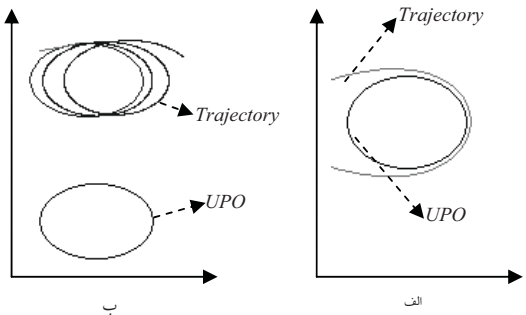
پس مساله کنترل سیستم آشوبی بیان‌شونده با رابطه (۴) به مساله سنکرون کردن آن با سیستم بیان‌شونده با رابطه (۷) که دارای ساختار مشابه است و دوره‌ای است، با استفاده از ورودی کنترلی $U_s(t)$ تبدیل می‌شود.

از آنجا که سیستم ارباب پایدار و دوره‌ای است پس از سنکرون‌سازی سیستم آشوبی نیز پایدار و دوره‌ای می شود.

نکته ۴: در روش تاخیر زمانی اختلاف بین حالات (خروجی‌های) جاری با حالات (خروجی‌های) با یک (یا چند) دوره تاخیر به عنوان خطا یا فاصله از UPO مورد نظر درنظر گرفته می‌شود در حالیکه در تکنیک ارائه شده جهت کم کردن زمان نشست، اختلاف حالات (خروجی‌ها) و مقادیر متناظر آنها در UPO بازسازی شده به عنوان خطا درنظر گرفته می‌شود که تفسیر بهتری از فاصله تا UPO مورد نظر است. شکل (۳) دو حالت را نشان می‌دهد که در آنها خطای استفاده شده در روش تاخیر زمانی بیانگر فاصله واقعی تا UPO نیست در حالیکه خطای استفاده‌شده در تکنیک کم‌کردن زمان نشست به‌خوبی بیانگر فاصله

[1] Reconstructing
[2] Slave
[3] Master

واقعی تا UPO است (زیرا UPO بازسازی‌شده منطبق بر UPO مورد نظر است). در قسمت الف شکل (۳) مسیرها به سمت UPO می‌روند و کمتر از یک دوره نزدیک آن می‌مانند، در این حالت خطا در روش تاخیر زمانی برای تمام نقاط زیاد است در حالیکه در واقع خطا برای نقاط زیادی، کم می‌باشد.



شکل ۳: دو حالت که خطا در DFC متناسب فاصله حالات تا UPO نیست

در قسمت ب شکل (۳) مسیرها حرکتی شبیه دوره‌ای اما دور از UPO موردنظر دارند، در این حالت خطا در روش تاخیر زمانی کم است در حالیکه در واقع خطا زیاد می‌باشد. لازم به ذکر است که چنین حالاتی در مسیرهای سیستم‌های آشوبی معمول است.

۴- نتایج شبیه‌سازی عددی: سیستم روسلر[1]

۴-۱-کاهش زمان نشست در سیستم نامی

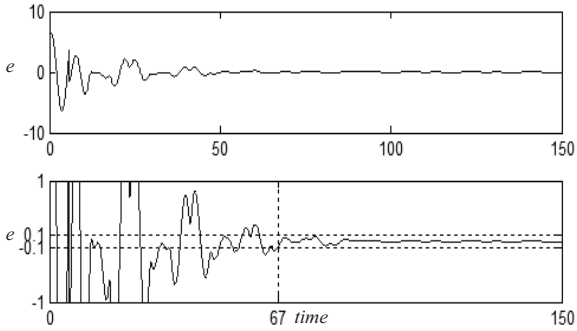
سیستم روسلر با معادلات زیر بیان می‌شود:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix}=\begin{bmatrix}0 & -1 & -1\\ 1 & a & 0\\ x_3 & 0 & -c\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}+\begin{bmatrix}0\\ 0\\ b\end{bmatrix}+\begin{bmatrix}u_1\\ u_2\\ u_3\end{bmatrix} \qquad (۸)$$

این سیستم به ازای $a=b=0.2$ و $c=5.7$ آشوبی بوده (سیستم نامی) و با کنترل‌کنندهٔ تاخیرزمانی به صورت : $u=u_2=k(x_2(t)-x_2(t-\tau))=ke$ قابل کنترل می‌باشد که در آن $\tau=T=5.881$ و تعریف خطا به صورت $e=x_2(t)-x_2(t-\tau)$ می‌باشد. در اینجا در صورتی که اندازه بهره بزرگتر از 0.7 انتخاب شود سیستم پایدار نمی‌شود همچنین مناسبترین بهره جهت کم بودن زمان نشست به صورت $k=-0.2$ می‌باشد که در شبیه‌سازی ها در نظر گرفته شده است.

ضریب طولانی بودن زمان نشست را $p=10$ انتخاب می‌کنیم. همچنین قدر مطلق حداکثر خطای قابل قبول $|e|_{\max}=0.1$ درنظر گرفته شده است.

شکل (۴) نتیجه کنترل سیستم نامی روسلر با روش تاخیر زمانی را نشان می‌دهد. چنانچه مشخص است زمان نشست $t_{set}=67\,\text{sec}$ (در تمامی شکل‌ها واحد زمان ثانیه می‌باشد) می‌باشد که طولانی $(t_{set}>pT=58.81)$ است.

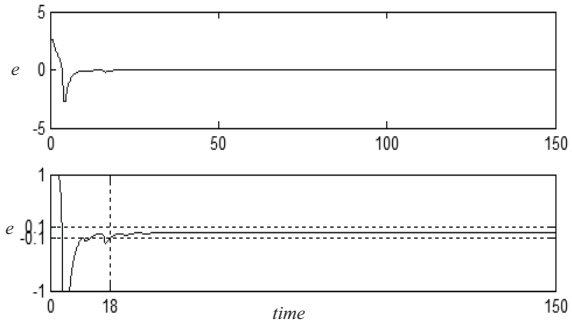


شکل ۴ : خطا در کنترل تاخیر زمانی سیستم روسلر نامی (قسمت پایین بزرگنمایی قسمت بالا می‌باشد)

جهت اعمال تکنیک کم‌کردن زمان ارائه شده، یک دوره از حالت x_2 سیستم پس از رسیدن به UPO ذخیره شد. چنانچه دیده می‌شود بازسازی کامل UPO مورد نیاز نیست بلکه جهت پرهیز از اضافه کردن اندازه‌گیر به سیستم، در بازسازی فقط حالت x_2 ذخیره‌شده با پریود $T=5.881$ تکرار می‌شود که با $\bar{x}_2$ نشان داده می‌شود. در اینجا خطا به صورت $e=x_2(t)-\bar{x}_2(t)$ تعریف می‌شود.

پس از آن (با استفاده از همان ورودی کنترلی روش تاخیرزمانی) از سنکرون‌کننده PI زیر جهت سنکرون کردن سیستم آشوبی و UPO بازسازی‌شده استفاده می‌شود:

$$\begin{bmatrix}u_1\\ u_2\\ u_3\end{bmatrix}=\begin{bmatrix}0\\ k_P(x_2-\bar{x}_2)+k_I\int_0^t(x_2-\bar{x}_2)dt\\ 0\end{bmatrix} \qquad (۹)$$



شکل ۵ : خطا در کنترل سیستم روسلر نامی با تکنیک کاهش زمان (قسمت پایین بزرگنمایی قسمت بالا می‌باشد)

رویه یافتن ضرایب تناسبی و انتگرالی چنین است که ابتدا جهت کم کردن زمان نشست تنها از کنترل‌کننده تناسبی استفاده می‌شود (یعنی $k_I=0$). سپس جهت داشتن عملکرد خوب در برابر تغیر پارامترها، قسمت انتگرالی به آن اضافه می‌شود. در اینجا انتخاب ضریب تناسبی به

[1] Rössler

صورت $-6 \le k_P \le -0.3$ زمان نشست را به خوبی کم می‌کند که از این بین $k_P = -3$ انتخاب گردید. پس از آن با کمی سعی و خطا $k_I = -0.3$ انتخاب شد.

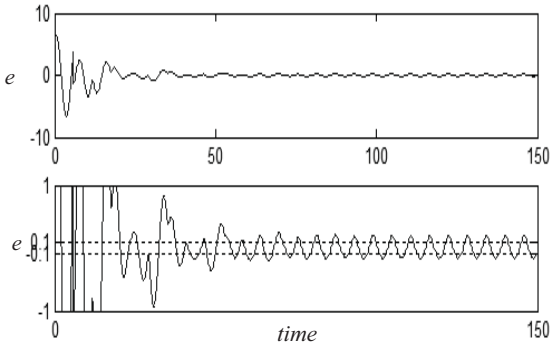
شکل (۵) نتیجه اعمال این روش را به سیستم روسلر نامی نشان می‌دهد. چنانچه در شکل (۵) مشخص است، زمان نشست $t_{set} = 18\,\text{sec}$ می‌باشد که طولانی نمی‌باشد.

۴- ۲- بررسی عملکرد تکنیک کاهش زمان نشست در سیستم با پارامترهای نامعین

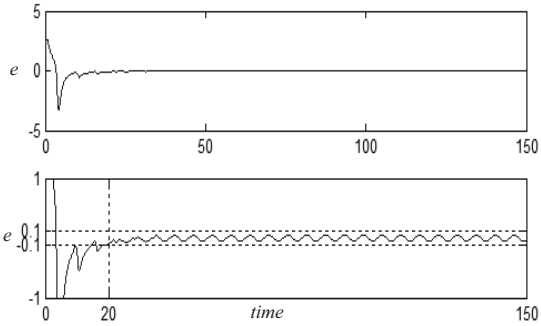
در این قسمت فرض می‌کنیم که پارامترهای سیستم روسلر نامعلوم بوده و طراحی بر اساس سیستم نامی انجام گرفته است، یعنی همان کنترل‌کننده تاخیرزمانی و تکنیک کاهش زمان قسمت (۴-۱) استفاده می‌شوند. در شبیه‌سازی پارامترهای سیستم نامعین روسلر به صورت $a = b = 0.22$ و $c = 5$ در نظر گرفته می‌شوند.

شکل (۶) نتیجه کنترل سیستم نامعین روسلر با روش تاخیر زمانی را نشان می‌دهد. چنانچه مشخص است خطا هرگز کمتر از ماکزیمم قابل قبول نمی‌شود و به عبارت دیگر زمان نشست بی‌نهایت است.

شکل (۷) نتیجه اعمال روش کاهش زمان را به سیستم نامعین روسلر نشان می‌دهند. چنانچه در این شکل مشخص است زمان نشست $t_{set} = 20\,\text{sec}$ می‌باشد که طولانی نمی‌باشد.



شکل ۶: خطا در کنترل تاخیر زمانی سیستم نامعین روسلر (قسمت پایین بزرگنمایی قسمت بالا می‌باشد)



شکل ۷: خطا در کنترل سیستم نامعین روسلر با تکنیک کاهش زمان (قسمت پایین بزرگنمایی قسمت بالا می‌باشد)

۵- نتیجه‌گیری

در این مقاله تکنیکی جهت کاهش زمان نشست در روش DFC ارائه گردید. چنانچه با مثال عددی نشان داده شد این تکنیک ساده و کارا است (البته جهت رسیدن به خواسته‌های زمان نشست کمتر و عملکرد بهتر در برابر پارامترهای نامعین، این تکنیک پیچیدگی بیشتری از روش تاخیر زمانی دارد). از آنجا که این روش به جای اختلاف حالات جاری از حالات دوره قبل، از اختلاف حالات و UPO بازسازی شده استفاده می‌کند، توانست عمل کنترل را در هر دو سیستم نامی و نامعین، بهتر و زودتر از روش تاخیرزمانی انجام دهد. مزیت دیگر این روش این است که در ساختار آن تاخیر وجود ندارد و بازه انتخاب بهره در آن بزرگ است.

به طور کلی می‌توان چنین نتیجه‌گیری کرد که هرچند روش تاخیرزمانی (بر خلاف انتظار علم کنترل) از تاخیر زمانی جهت کنترل سیستم آشوبی استفاده می‌کند اما در هر صورت وجود تاخیر در سیستم حلقه بسته امری نامطلوب است و بنابراین پیشنهاد می‌شود پس از طراحی روش تاخیر زمانی برای یک سیستم با استفاده از تکنیک ارائه‌شده در این مقاله تاخیر موجود در ساختار کنترل‌کننده حذف گردد.

اگرچه در این مقاله سیستم روسلر درنظر گرفته شد و جهت بازسازی UPO از تکرار داده‌های ذخیره شده استفاده شد، از تکنیک‌های دیگر مانند مدلسازی غیرخطی نیز می‌توان استفاده کرد و بسیاری سیستم‌های دیگر را نیز می‌توان به عنوان سیستم مورد مطالعه لحاظ کرد.

مراجع

[28] E. Ott, C. Grebogi and J. A. Yorke, “Controlling chaos”, Physical Review Letters., vol. 64, no. 11, pp. 1196-1199, March 1990.

[29] A. L. Fradkov and R.J. Evans, “Control of chaos: Methods and applications in engineering”, Annual Reviews in Control, vol. 29, pp. 33-56, April 2005.

[30] F. T. Arecchi, S. Boccalettiy, M. Ciofini and R. Meucci, “The control of chaos: theoretical schemes and experimental realizations”, International Journal of Bifurcation and. Chaos, vol. 8, no. 8, pp. 1643-1655, 1998.

[31] K. Yagasaki and T. Uozumi, “A new approach for controlling chaotic dynamical systems”, Physics Letters A, vol. 238, pp. 349-357, February 1998.

[32] J. Starrett, "Time-optimal chaos control by center manifold targeting", Physical Review E, vol. 66, pp. 046206, October 2002.
[33] K. Pyragas, "continuous controlling of chaos by self-controlling feedback", Physics. Letters A, vol. 170, pp. 421-428, 1992.
[34] K. Pyragas, "Control of chaos via extended delay feedback," Physics Letters A, vol. 206, pp. 323-330, 1995.
[35] W. Just, T. Bernard, M. Ostherier, E. Reibold and H. Benner, "Mechanism of time-delayed feedback control", Physical Review Letters, vol. 78, no. 2, pp. 203-206, January 1997.
[36] Ö. Morgül, "On the stability of delayed feedback controllers", Physics Letters A, vol. 314, pp. 278-285, 2003.
[37] J. Escalona, and P. Parmananda, "Noise-aided control of chaotic dynamics in a logistic map", Physical Review E, vol. 61, no. 4, pp. 5987-5989, May 2000.
[38] N. Vasegh, A. Khaki Sedigh, "Chaos control in delayed chaotic systems via sliding mode based delayed feedback", Chaos, Solitons and Fractals, vol. 40, no. 1, pp. 159-165, 2009.
[39] M. Ramesh and S. Narayanan, "Chaos control of Bonhoeffer–van der Pol oscillator using neural networks", Chaos, Solitons and Fractals, vol. 12, pp. 2395-2405, 2001.
[40] A. Alasty and H. Salarieh, "Controlling the chaos using fuzzy estimation of OGY and Pyragas controllers", Chaos, Solitons and Fractasl, vol. 26, pp. 379-392, 2005.
[41] H. Nakajima, "Some sufficient conditions for stabilizing periodic orbits without the odd-number property by delayed feedback control in continuous-time systems", Physics Letters A, vol. 327, pp. 44-54, 2004.

مجله کنترل جلد ۳، شماره ۱، بهار ۱۳۸۸، صفحه ۳۶-۲۹

بکارگیری روش میانگین گیری مرتب وزندار(OWA) در ترکیب داده های ربات مین یاب

محمد رضا بادلّو[۱]، بهزاد مشیری[۲]، بابک نجار اعرابی[۳]

[۱] دانشجوی دکتری برق- کنترل، دانشگاه آزاد اسلامی واحد علوم و تحقیقات،badello@myway.com

[۲] استاد، قطب علمی کنترل و پردازش هوشمند دانشکده مهندسی برق و کامپیوتر دانشگاه تهران،moshiri@ut.ac.ir

[۳] دانشیار، قطب علمی کنترل و پردازش هوشمند دانشکده مهندسی برق و کامپیوتر دانشگاه تهران،arabi@ut.ac.ir

چکیده: امروزه از ربات ها در کاربردهای بسیاری بهره برداری می شود و استفاده از تجهیزات ابزار دقیق گسترده و متنوع جزو مشخصه های چنین رباتهایی محسوب می گردد. بنابراین در برنامه ریزی ربات مقوله ترکیب داده های ربات از اهمیت خاصی برخوردار است چرا که عملکرد بهینه ربات هنگامی تضمین می شود که بتوان از حجم زیاد اطلاعات دریافتی توسط سنسورهای مختلف ربات به بهترین نحو استفاده نموده و آنها را به روشی مناسب با یکدیگر ترکیب نمود و بهترین نتیجه را دریافت کرد. در این مقاله ابتدا با یک بررسی اجمالی برروی روشهای مختلف ترکیب داده ها و مقایسه کمی و کیفی آنها، روش میانگین گیری مرتب وزندار را برروی ربات مین یاب پیاده سازی نموده و نتایج عملی و شبیه سازی شده آنرا بیان می کنیم. جهت تعیین ضریب وزن سنسورها از یک روش منحصر بفردی استفاده نموده و در مرحله اجرایی جهت عملکرد سریع و بهینه ربات ضریب وزنها را شناور در نظر گرفته که با توجه به موقعیت ربات و اطلاعات کسب شده توسط سنسورها ضریب وزنها تغییر میکند، در نهایت به کاربرد های عملی این روش می پردازیم.

کلمات کلیدی: ترکیب داده ها , ربات مین یاب , میانگین گیری مرتب وزندار.

Abstract: Robots have different applications and using various precision tools is considered among features of such robots. Thus combination of robot data is of considerable importance in planning robots since optimal performance of robots ensures when a great volume of data received by different sensors of robot can be used best and can be combined by a suitable manner to achieve the best result. By a brief study of different methods of data combination and their quality and quantity comparison, this article deals with ordinary weighting averaging method on mine detector robot and describes the practical assimilated results. A unique method is used to determine weight coefficient of sensors and weight coefficients are considered floating in executive stage for the purpose of optimal performance. Weight coefficients change based on robot state and data obtained by sensors. Practical applications of this method are finally described.

Keywords: data fusion, mine detector robot, ordinary weighting averaging.

۱- مقدمه

ترکیب داده ها در سیستمهای هوشمند و خودمختار ضروری می باشد، چرا که با توسعه تجهیزات ابزار دقیق و بکار گیری سنسورهای متعدد در سیستمها جهت استفاده بهینه از تکمیل داده ها و جلوگیری از بروز اشکال بدلیل افزونگی اطلاعات باید ترکیب داده صورت گیرد تا عملکرد مطلوب سیستم تضمین شود. بیشتر مقالات جدید این مبحث در کنفرانس بین المللی ترکیب اطلاعات[۱] ارائه می شوند که در [1] مرور مختصری بر آنها وجود دارد. به دلیل داشتن بیشترین بودجه تحقیقاتی، بیشترین کاربرد ترکیب اطلاعات در پژوهشهای نظامی صورت می گیرد . در کاربردهای نظامی، از ترکیب اطلاعات به منظور شناسایی[۲] دقیق تر اهداف و نیز در هدایت[۳] انواع وسایل قابل کنترل از راه دور مانند موشکها و هواپیما های بدون سرنشین استفاده می شود. هدایت از راه دور در علم رباتیک نیز کاربرد وسیعی دارد، اگرچه از ترکیب اطلاعات در شناسایی محیط اطراف ربات نیز می توان استفاده کرد. می توان با استفاده از سنسورهایی ارزان قیمت ولی متنوع، اطلاعات مختلفی با ویژگی های متفاوت از محیط اطراف به دست آورده و با ترکیب آنها ربات را قادر به شناختی بهتر از فضای ناشناخته اطراف خود نمود. متداولترین کاربردهای

[1] -International Conference on Data Fusion

[2] - Identification

[3] - Navigation

ترکیب اطلاعات در رباتیک، استفاده از آن در تشکیل جدول اشغال[1] و نیز استفاده در ترکیب دستور[5] می باشند[2]. در مقابل عمل هدایت، عمل تعقیب اجسام[6] را می توان عنوان کرد که علاوه بر استفاده در علم رباتیک یکی از ارکان اصلی در صنایع نظامی می باشد. به عنوان مثال می توان از [3] نام برد، که در آن با استفاده از منطق فازی در ترکیب اطلاعات، راه حلی ارائه داده شده است که کشتی های مختلف را بتوان در آبهای آزاد تعقیب و شناسایی کرد. در علوم نقشه برداری و زمین شناسی از ترکیب اطلاعات، در دستیابی به تصاویری مطلوب از عکسهای تهیه شده از ماهواره های مختلف استفاده می شود، تا بتوان محدوده مناطق مختلفی از جمله دشت ها، جاده ها، دریا ها، جنگل ها و مناطق مسکونی مثل شهرها را به خوبی از یکدیگر تفکیک نمود.

موارد بسیار متنوعی وجود دارد که از ترکیب اطلاعات استفاده شده است، از جمله می توان به کاربرد نظریه ترکیب اطلاعات در بازشناخت الگو، پردازش تصاویر پزشکی، بیوانفورماتیک، اینترنت و تجارت الکترونیک، و کاربرد در سیستم های حمل و نقل هوشمند اشاره کرد . در این میان روش میانگین گیری مرتب وزندار (OWA) بعنوان یک روش هوشمند در ترکیب اطلاعات شناخته شده [4] که ضمن سادگی در مقایسه با سایر روشهای هوشمند قابلیت های خاص و ویژه ای دارد. ربات بعنوان یک سیستم هوشمند نیز کاربردهای فراوانی دارد که در این میان ربات مین یاب بدلیل وجود حسگرهای متعدد بستر آزمایش مناسبی می باشد که می توان روشهای مختلف کلاسیک و هوشمند ترکیب داده ها را بر روی آن اجرا نمود. پیش از این همین روش میانگین گیری مرتب وزندار در ربات تعقیب خط ساده و مغناطیسی بکار گرفته شده است[18] .

۲- ترکیب اطلاعات

به دلایل مختلف، در اندازه گیریهای انجام شده توسط حسگرهای یک سیستم، مقداری عدم قطعیت[7] وجود دارد و در واقع می توان گفت نتایج اندازه گیری ها تخمینی از کمیت های اصلی می باشد. بر اساس تعریف ارائه شده در[5] ترکیب اطلاعات عبارت است از ترکیب توامان[8] اطلاعات اخذ شده از منابع مختلف به نحوی که نتیجه حاصل، جامع و مانع بوده و قابل استفاده برای انجام وظیفه ای از پیش تعیین شده به صورت خودمختار باشد. اطلاعات تحت ترکیب، ممکن است همزمان از چندین منبع مختلف حاصل شده باشند و یا از یک منبع در مقاطع زمانی مختلف، گردآوری شده باشند .ترکیبی از هر دو حالت نیز می تواند انجام شود، یعنی ممکن است در مرحله ای، اطلاعات اخذ شده، از دو منبع مختلف در دو زمان مختلف با هم ترکیب شوند و محصول این ترکیب، در مرحله ای دیگر، با اطلاعات یک زیر مجموعه دیگر از سیستم ترکیب شود.

۲-۱- چرا ترکیب اطلاعات

مهم ترین مزیت ترکیب اطلاعات در کم کردن میزان عدم قطعیت است. این امر توسط مفهومی با عنوان افزونگی[9] اطلاعات میسر می باشد. اطلاعاتی که از چند منبع یا از یک منبع در چند لحظه مختلف جمع آوری شده اغلب شامل افزونگی می باشند. افزونگی به معنای وجود همپوشانی در اطلاعات است. ترکیب اطلاعاتی که دارای همپوشانی هستند سبب می شود که نتیجه، دارای قطعیت و قابلیت اعتماد[10] بیشتری باشد. مزیت دیگر استفاده از روش های ترکیب اطلاعات در تکمیل شدن است. در بسیاری از موارد، منابع مختلف، هر یک ویژگی خاصی از ورودی را تشخیص می دهند و معمولا هیچ منبعی نمی تواند همه ویژگیهای موجود را تشخیص دهد. به کارگیری ترکیب اطلاعات، موجب تکمیل اطلاعات منابع مختلف و دستیابی به اطلاعات جامعی از موضوع می شود. استفاده از روشهای ترکیب اطلاعات مزایای دیگر نیز دارد که از آن بین می توان به تسریع در پردازش و اقتصادی بودن اشاره کرد سرعت پردازش اطلاعات چند منبع، که همزمان اطلاعات خود را به پردازشگر می دهند تا آ نها را با یکدیگر ترکیب کند، نسبت به حالت وجود یک منبع، جهت دریافت اطلاعات افزایش می یابد. در واقع ترکیب اطلاعات، موجب تحقق نوعی پردازش موازی در سیستم می شود که این امر به نوبه خود موجب تسریع در پردازش خواهد بود. همچنین به علت دریافت اطلاعات از منابع مختلف، ترکیب داده ها با یکدیگر و متمرکز شدن بر اطلاعات مفید دریافتی، می توان از منابع ارزان تری در سیستم استفاده کرد و دیگر احتیاجی به صرف هزینه های بالا برای استفاده از حسگر های دقیق نمی باشد.

در نهایت چهار دلیل بارز جهت توجیه ترکیب اطلاعات عبارت است از: افزونگی اطلاعات، تکمیل شدن داده ها، درستی و دقت داده ها و کاهش هزینه اطلاعات.

۲-۲- سطوح مختلف ترکیب اطلاعات

می توان تقسیم بندی ساده ای از سطوح مختلف ترکیب اطلاعات ارائه کرد [6]، در این تقسیم بندی، ترکیب اطلاعات میتواند در سطح پایین (حسگر) سطح میانی (ویژگی) ویا در سطح بالا (تصمیم گیری) رخ دهد. ترکیب در سطح حسگر به این معنی است که در این سطح اطلاعات خام و پردازش نشده ای که از حسگرها به دست آمده اند با یکدیگر ترکیب می شوند. روشهای ترکیبی که شامل پیش بینی اطلاعات می شوند، عمدتاً در این سطح از ترکیب به کار می روند. در سطح بعدی، یعنی ترکیب در سطح ویژگی، ویژگیهای استخراج شده از اطلاعات حسی، با یکدیگر ترکیب می شوند. روشهای ترکیب در این سطح، عموماً روشهای مبتنی بر ترکیبات و تبدیلات هندسی و ترکیب نسبت های متناظر با یک ویژگی خاص هستند. در بالاترین سطح، ترکیب در سطح نماد صورت می گیرد. در این

[1] - Occupancy Grid
[5] - Command Fusion
[6] - Object Tracking
[7] - Uncertainty
[8] - Synergistic
[9] - Redundancy
[10] - Reliability

سطح، آنچه که ترکیب می شود نوعی نماد یا تصمیم است که پس از قدری پردازش و تبدیل بر روی اطلاعات ورودی حاصل شده است .

۲-۳- پیکر بندی حسگرها در ترکیب اطلاعات

توزیع منابع مختلف اطلاعاتی در یک شبکه ی پردازش اطلاعات ارائه شده است [7]، که نشان دهنده ترتیب پردازش اطلاعات منابع مختلف می باشد , توزیع مذکور می تواند بصورت سری، موازی، موازی سری و سری موازی باشد.

۲-۴- روشهای مختلف ترکیب اطلاعات

در یک تقسیم بندی سه روش کلی برای ترکیب اطلاعات وجود دارد که عبارت است از: روشهای کلاسیک – روشهای هوشمند – روشهای الهام گرفته شده از علوم مختلف.

روشهای کلاسیک درترکیب اطلاعات

روشهای کلاسیک درترکیب اطلاعات با استفاده از دو نظریه اساسی امکان پذیر است : استفاده ازنظریه احتمالات – استفاده از نظریه شواهد [8,9].

روشهای هوشمند در ترکیب اطلاعات

روش فازی – روش میانگین گیری مرتب وزندار– روش شبکه های عصبی از مهمترین روشهای هوشمند در ترکیب اطلاعات محسوب می شود [10,11,12].

انسان ها معمولا با مفاهیمی سروکار دارند که غیر دقیق است و این عدم دقت از مرزهای نادقیق تعریف این مفاهیم ناشی می شود[13]. مجموعه فازی در حقیقت بیان ریاضی این عدم دقت به صورت تعمیمی از نظریه مجموعه ها می باشد. در نظریه مجموعه های فازی، درجه عضویت می تواند مقداری بین صفر و یک اختیار کند و به این صورت می تواند عدم قطعیت را مدل کند. در این منطق مساله ای به نام تجمیع وجود دارد که در مبحث ترکیب اطلاعات بسیار کاربرد دارد یک عملگر تجمیع یک زیر مجموعه از اشیاء مربوط به یک مجموعه خاص را به عضوی از آن مجموعه نسبت می دهد.به بیان ریاضی یک عملگر تجمیع تابعی است که تعدادی از اعداد حقیقی را به یک عدد حقیقی دیگر نسبت می دهد. عملگرهای تجمیع خواص ریاضی و رفتاری زیادی می توانند داشته باشند که این خواص لزوماً با یکدیگر سازگار نیستند. در [14] این خواص به طور کامل شرح داده شده اند. از معروف ترین عملگرهای تجمیع می توان عمگرهای انتگرال گیری[5] ، مانند انتگرال فازی سوگنو[۱۱][15] و انتگرال فازی چوکوی[۱۲] را نام برد که هر دو به نوعی عمل میانگین گیری در هم ریخته[۱۳] را انجام می دهند.

عملگر میانگین گیری مرتب وزن دار[۱۴] (OWA) از جمله عملگرهای مطرح در ادبیات مجموعه های فازی می باشد. می توان ثابت کرد که این عملگر مانند حالتی خاص از انتگرال فازی عمل می کند. یکی از مسائل مهم در زمینه تصمیم گیری، ترکیب معیارها و تشکیل تابع تصمیم است. [16]از یک سو، گاهی برآورده شدن تمام معیارها مهم بوده، و از سوی دیگر در برخی موارد اینکه حداقل یکی از آن ها برآورده شود اهمیت پیدا می کند. این دو مسیر به استفاده از عملگرهای AND و OR برای ترکیب توابع معیار، بر می گردند. عملگرد OWA نوع تجمیع انجام می دهد که در میان دو حالت ذکر شده، قرار می گیرد و از این رو می توان آن را عملگری با خاصیت AND-OR نامید. نخستین بار یاگر این عملگر را معرفی کرد و تاکنون از سوی خود او و محققین بسیار دیگری،کاربردها و نسخه های جدیدی برای آن، ارائه شده است[6] .

شبکه های عصبی مصنوعی[۱۵] یا سیستم های عصبی تطبیقی، سیستمهایی سخت افزاری یا نرم افزاری هستند که سعی می کنند عملکرد سیستمهای عصبی بیولوژیکی را تقلید کنند. از آنجا که شبکه های عصبی دارای توانایی شایان توجهی در یادگیری مدلها و رفتارها هستند، در بسیاری از کاربرد های هوش مصنوعی به کار گرفته شده اند. شبکه عصبی یک تبدیل غیرخطی روی اطلاعات ورودی انجام می دهد که منجر به یک بردار خروجی می شود .چنین تبدیلی می تواند مشابه تکنیک های خوشه بندی[۱۶] نگاشتی از اطلاعات به دسته های مختلف هویت باشد. بنابراین شبکه های عصبی می توانند برای تبدیل اطلاعات چند حسگر به یک اعلام مشترک در مورد هویت شیء استفاده شوند. به همین صورت از آن ها در کاربر دهای رباتیک نیز می توان کمک گرفت. به عنوان مثال در [8] از یک شبکهٔ عصبی بازگشتی برای ترکیب اطلاعات ورودی با تخمینی از حالت بعدی استفاده می شود که هدف آن تعقیب یک ربات متحرک شبیه سازی شده است.

روشهای الهام گرفته شده از علوم مختلف

علوم بکار گرفته شده در ترکیب اطلاعات عبارت است از :

۱- علوم زیست شناسی که خود بر دو نوع است : (علم روانشناسی و علم عصب شناسی) ۲- علوم شناختی

۳- ربات مین یاب

بر اساس گزارشات موجود، ایران سومین کشور دنیا از لحاظ مین های دفن شده است. در طول ۸ سال جنگ تحمیلی عراق بر علیه ایران بالغ بر ۱۶ میلیون مین در نقاط مرزی غرب و جنوب کشور کاشته شده است که بعضا بعلت دور افتاده و صعب العبور بودن مناطق , خنثی نشده باقی مانده و همه ساله افراد زیادی را از بین مردم عادی و نظامیان قربانی میکند.کشف و خنثی

[۱۱] Sugeno
[۱۲] Choquet
[۱۳] Distorted Averaging
[۱۴] Ordered Weighted Averaging
[۱۵] Artificial Neural Networks(ANN)
[۱۶] Clustering

سازی مین ها عملی بسیار خطرناک و پرهزینه است. این امید وجود دارد که رباتهای مین یاب بتوانند این خطر را رفع نموده وتلفات انسانی و هزینه مین یابی را کاهش دهد.

جستجو، یافتن و خنثی کردن انواع مین های مختلف را میتوان از اهداف یک ربات مین یاب دانست. با توجه به نوین بودن این شاخه از رباتیک تقریبا می توان گفت دو هدف اول یعنی جستجو و یافتن مین محقق شده است اما در مورد سوم یعنی خنثی کردن مین، ربات های حال حاضر قادر به انجام این کار آن هم در مورد انواع مختلف مین های موجود نمی باشد.

۳-۱- ربات مین یاب ونوس

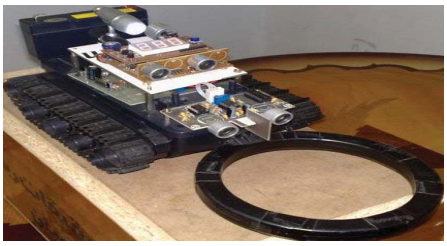
ربات مین یاب ونوس را می توان یکی از کامل ترین ربات های حال حاضر دانست، این ربات با بهره گیری از یک پردازنده ی پر سرعت قادر به جستجوی مین و کشف آن است در این ربات از یک سیستم تشخیص مین پیشرفته استفاده شده که عمق جستجو و تفکیک مین های پلاستیکی از فلزی جزو قابلیت های بالای این سیستم تشخیص مین است. برای تشخیص موانع از ماژول های ما فوق صوت (ultra sonic) که با سیستم هدایت کننده ی مرکزی ارتباط دارد استفاده شده است.

بخش مکانیکی این ربات به لحاظ حرکت در مناطق نا هموار بسیار حائز اهمیت بود لذا برای این منظور از سیستم تانکی با مکانیزم حرکتی شنی استفاده شد که به لحاظ استفاده از این مکانیزم در ربات مین یاب ونوس اقدامی جدید و بدیع بوده است.

شکل ۱- نمای جانبی ربات ونوس سیستم هدایت مرکزی

از آنجا که این ربات به صورت کاملا هوشمند به جستجو پرداخته و فعالیت می کند اهمیت این بخش بسیار زیاد بود لذا جهت هر چه بالاتر بردن سرعت این بخش از نسل جدید ریز کنترل کننده های خانواده atmel یعنی میکرو کنترلر های AVR استفاده شد وظیفه ی این بخش را به اختصار می توان چنین بیان کرد ;ارتباط با کلیه ی بخشهای جانبی از جمله موتورها، سنسورها، سیستم تشخیص مین و...، دریافت داده های محیطی، پردازش داده ها، تصمیم گیری بر اساس نتایج پردازش و صدور فرمان برای سیستم محرک یا سیستم هشدار یا سایر بخش ها.

ارتباط بخشهای جانبی با هسته مرکزی از طریق پروتکل سریال usart برقرار می شود استفاده از سخت افزارکمتر مزیت بزرگ این پروتکل ارتباطی است که این ویژگی با توجه به فضای محدود، در داخل بدنه ربات برای ما از اهمیت فوق العاده ای برخوردار است.

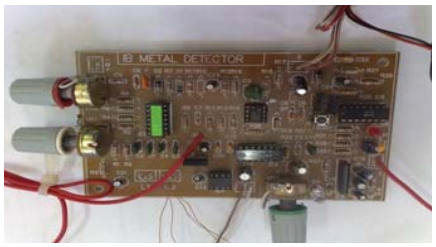


شکل ۲- نمایی از ربات ونوس بهمراه فلز یاب ربات سیستم تشخیص مین

این بخش متشکل از دو قسمت مهم است ۱- مدار فلزیاب برای دریافت داده ها ۲- پردازنده جهت تحلیل داده ها

A – مدار فلز یاب:

از میان فلز یاب های موجود که هر کدام بر اساس تکنولوژی خاصی عمل می کند، مدار فلزیاب IB را که برای ما دارای مزیت های بیشتری بود انتخاب شد و در ساخت از آن استفاده کردیم.



شکل ۳- نمایی از ربات ونوس بهمراه فلز یاب ربات

B - پردازش داده ها:

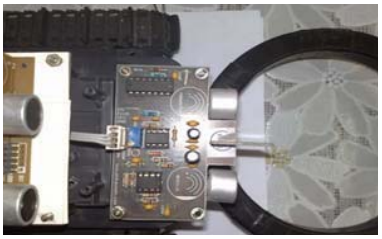
داده های خام بدست آمده از مدار فلزیاب توسط بخش پردازنده دریافت و پردازش می شود. این پردازش شامل تشخیص مین، نوع و اندازه آن و به موازات تشخیص نویز و دفع آن می باشد. برای این منظور از یک ریز کنترل کننده و مدارات جانبی آن استفاده شده که در نهایت این سیستم را به یکی از کامل ترین سیستم های موجود تشخیص مین مبدل کرده است.

سیستم فاصله سنج و تشخیص موانع

از آنجایی که این ربات به صورت کاملا خودکار و بدون دخالت انسان فعالیت می کند، آگاهی از محیط پیرامون لازمه ی اصلی برای حرکت و جستجو است. لذا برای تشخیص موانع پیرامون و همچنین اندازه گیری فاصله ربات تا موانع موجود از سنسور های ما فوق صوت استفاده شده، اصول کار این سنسور ها بر اساس بازتاب صوت و اندازه گیری زمان رفت و برگشت صوت و در نتیجه اندازه گیری فاصله است، حداکثر برد این سنسورها ۴ متر و حداقل آن ۴ سانتی متر است که البته مقادیر فوق با تغییر حالت فیزیکی فرستنده و گیرنده قابل تغییر است.

ربات مین یاب ونوس از دو ماژول آلتراسونیک بهره می برد که یکی از آنها در قسمت پایین ربات قرار دارد به صورت دیجیتال عمل می کند و ماژول بعدی که در قسمت بالای ربات قرار دارد به صورت آنالوگ عمل می کند.

قسمت آنالوگ فاصله تا مانع را با دقت ۱ سانتی متر و تا فاصله حداکثر ۴ متر به ما می دهد و قسمت دیجیتال جهت جلوگیری از برخورد ربات با دیواره ها و همچنین مانع های کوچک که با سنسور آنالوگ قابل تشخیص نیست تعبیه شده است.



شکل۴- نمایی از مدار فاصله سنج و تشخیص موانع ربات ونوس

در کل استفاده از دو جفت سنسور ما فوق صوت در بهبود عملکرد ربات و بالا بردن ضریب امنیت ربات به جهت برخورد با موانع موجود در محیط تاثیر بسزایی دارد.

سیستم محرک و مکانیزم حرکتی ربات

همانطور که قبلا نیز اشاره شد مکانیزم حرکت شنی با توجه به محیط فعالیت ربات تعبیه شده است برای نیروی محرکه ربات از موتورهای DC استفاده شده که برای بالا بردن قدرت موتورها , گیربکس ها بکار گرفته شدند و در کل اتصال موتورها، گیربکس و در نهایت شنی ها یک سیستم کامل قابل انعطاف را تشکیل داده است. درایور موتور ها از ترانزیستور های TIP بهره می برد که استفاده از ترانزیستور علاوه بر پایین آوردن هزینه ها قابلیت کنترل سرعت توسط روشPWM رانیز امکان پذیر میکند.[17]

۳-۲- نحوه برنامه ریزی ربات

آنچه که در برنامه ریزی ربات حائز اهمیت است کشف تعداد زیادی مین در حداقل زمان می باشد , (بدون برخورد ربات با موانع) با سیستم معرفی شده تا حدی به این هدف نائل شده ایم اما تغییراتی باید صورت پذیرد تا عملکرد اصلاح گردد. آنچه که به تازگی مشاهده می شود استفاده از تکنیک پردازش تصویر در تشخیص مانع است که عملکرد ربات را نسبت به سایر سنسورهای تشخیص مانع بهبود می بخشد، ولی اشکال وارده در مقدار و حجم اطلاعاتی است که ربات باید در هر لحظه پردازش نماید.

موضوع دیگر تغییر در سنسور تشخیص مین است که افزایش ابعاد این سنسور کاهش دقت عمل و افزایش ابعاد ربات را بدنبال دارد که چندان عملی و منطقی نمی باشد چرا که افزایش ابعاد ربات مشکل برخورد با موانع را ایجاد می کند، در نتیجه راه حل پیشنهادی استفاده چندین سنسور کوچکتر بجای استفاده از یک سنسور بزرگ می باشد که می تواند مزیتهایی را بدنبال داشته باشد. از آن جمله:

- بدون تغییر در ابعاد ربات می توان موقعیت سنسورها را متناسب با محیط مین گذاری شده تنظیم نمود.

- موقعیت مین را می توان با دقت بیشتری تعیین کرد.

- در مقایسه با روش پردازش تصویر پردازش اطلاعات سریعتر صورت می گیرد.

- می توان ربات را بگونه ای برنامه ریزی کرد که عملکرد ربات تحت تاثیر آخرین مین پیدا شده قرار گیرد، بعبارت دیگر محیط جزئی از سیستم (برنامه ریزی عملکردی ربات) خواهد شد. در اینجا به محیط اصطلاح هوش استاتیک و به عملکرد ربات هوش (حافظه) دینامیک گوییم. در واقع برنامه ریزی عملکرد ربات به گونه ای است که میتوان بین هوش دینامیکی و هوش استاتیکی یک تعامل زیبا و منطقی برقرار ساخت.

۴- روش پیشنهادی

استفاده از سه سیم پیچ بجای یک سیم پیچ محور اصلی تغییرات قرار گرفت، اما مسئله موجود نحوه استفاده از مجموعه اطلاعات حاصل از سنسورهای مغناطیسی و صوتی در ایجاد فرامین حرکتی لازم به ربات است تا اهداف مورد نظر بخوبی دنبال شود. اگر تا پیش از این حرکت ربات متاثر از سنسورهای تشخیص مانع و تک سنسور مین یاب بود حال باید ربات را به گونه ای برنامه ریزی کرد که ربات تحت تاثیر مجموعه سنسورهای تشخیص مین قرار گرفته و از خود واکنش نشان دهد تا عملکرد ربات بهینه گردد، برای این منظور ابتدا ربات را فقط با توجه به اطلاعات سه سنسور مین یاب برنامه ریزی کرده سپس اطلاعات تشخیص موانع را وارد می کنیم.

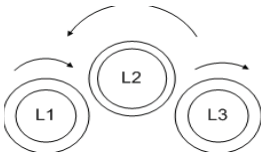
۴-۱- روابط حاکم

چگالی میدان مغناطیسی ایجاد شده در اطراف یک سیم پیچ طبق رابطه زیر محاسبه می شود :

$$B = K \frac{I}{R} \qquad (۱)$$

در این رابطه K ضریب ثابت , I جریان سیم پیچ و R شعاع سیم پیچ میباشد. در انتخاب جهت جریان سیم پیچها باید دقت کرد و تاثیر القاء متقابل سیم پیچها را باید در نظر گرفت تا جریان در هر سیم پیچ موجب تقویت فوران مغناطیسی سیم پیچ مجاورش شود.

شکل زیر این موضوع را بهتر نشان می دهد , همچنین روابط حاکم بر ولتاژ و جریان هر سیم پیچ بر حسب خود القایی و ضریب القای متقابل آورده شده است.



شکل۵- جهت جریان انتخاب شده در هر یک از سیم پیچها

$$V1 = L1\frac{di1}{dt} + M12\frac{di2}{dt}$$
$$V2 = L2\frac{di2}{dt} + M12\frac{di1}{dt} + M23\frac{di3}{dt} \qquad (۲)$$
$$V3 = L3\frac{di3}{dt} + M23\frac{di2}{dt}$$

ضریب القاء متقابل با فاصله سیم پیچ ها از یکدیگر نسبت عکس دارد، و میزان تاثیر گذاری هر سیم پیچ بر سیم پیچ مجاورش به شدت فوران مغناطیسی ایجاد شده و این فاصله بستگی دارد. که می توان با انتخاب مناسب جریان هر سیم پیچ و فواصل موجود بهترین سیستم تشخیص مین را طراحی کرد.

شکل۶- افزایش تعداد سنسورها بر روی ربات و نحوه قرار گرفتن آنها

۴-۲- استفاده از روش میانگین گیری مرتب وزندار (OWA) در ترکیب اطلاعات سنسورها

حال که سنسورهای تشخیص مین را افزایش دادیم قصد داریم عملکرد ربات بگونه ای کنترل نمائیم تا به صورت یک برنامه از پیش تعیین شده نباشد بلکه حرکت ربات بطور هوشمندانه بر اساس آخرین مین گشف شده شکل گیرد برای این منظور از روش OWA در ترکیب اطلاعات سنسورها استفاده کرده و سه عملگر تعریف می کنیم تا ولتاژهای لازم جهت اعمال به موتورها ایجاد شود.

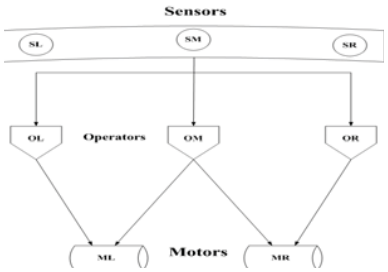
در این روش برای هر سنسوربا توجه به موقعیت و اهمیت آن ضریب وزنی بین صفر و یک در نظرگرفته می شود به نحوی که مجموع ضریب وزنها برابر یک گردد آنگاه عملگر را تعریف میکنیم.

روابط حاکم در روش OWA چنین است:

$$W_j \in [0,1] \rightarrow \sum_j W_j = 1 \qquad (۳)$$
$$OWA(a_1, a_2, \ldots, a_n) = \sum_{j=1}^{n} w_j b_j$$

حال ما از سه عملگر OWA جهت کنترل اتوماتیک ربات استفاده می کنیم. ۱- عملگر با ضریب وزن غالب برای سنسور وسط جهت ایجاد ولتاژ بایاس هر دو موتور در حرکت مستقیم ۲- عملگر با ضریب وزن غالب برای سنسورسمت راست جهت ایجاد فرمان به موتور سمت چپ ۳- عملگر با ضریب وزن غالب برای سنسورسمت چپ جهت ایجاد فرمان به موتور سمت راست

چگونگی الگوریتم در دیاگرام بلوکی زیر مشخص شده است.



شکل۷- دیاگرام بلوکی نحوه اجرای الگوریتم OWA

عملگر OM با ایجاد ولتاژ بایاس برای هر دو موتور هدایت ربات در مسیر مستقیم را بر عهده دارد، عملگر OR با دریافت اطلاعات از سنسور SL به موتور سمت راست فرمان می دهد تا ربات را بطرف چپ بچرخاند و عملگر OL بر عکس OR عمل میکند و بدین ترتیب ربات تمامی مین های مسیر حرکتش را شناسایی و دنبال می کند.

۴-۳- استفاده از تابع توزیع گوسی در تعیین ضریب وزن سنسورها

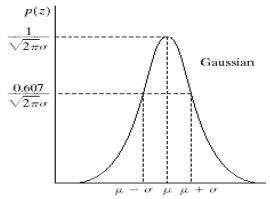
توزیع گوسی یا نرمال یکی از کاربردی ترین توزیعها در میان متغیرهای تصادفی است.

رابطه تابع توزیع گوسی چنین است:

$$P(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-\mu)^2/2\sigma^2} \qquad (۴)$$

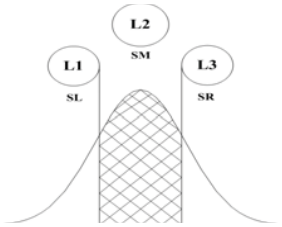
- میانگین : μ
- انحراف استاندارد : σ

مهمترین عاملی که باعث شد ازتابع توزیع نرمال جهت تعیین ضریب وزن سنسورها استفاده کنیم این است که سطح زیر منحنی تابع توزیع برابریک است.



شکل۸- تابع توزیع گوسی

لذا از همین اصل استفاده نموده و ضریب وزن هر سنسور را برابر سطح زیر منحنی ایجاد شده با سنسور مجاورش در نظر میگیریم و بدین صورت یک رابطه منطقی و زیبا بین تعداد و فاصله سنسورها از یکدیگر با ضریب وزنشان ایجاد می شود[18].

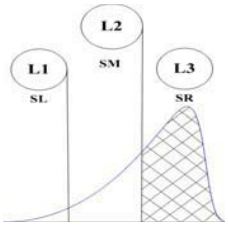


شکل۹- استفاده از سطح زیر منحنی تابع توزیع نرمال در محاسبه ضریب وزنها

همانگونه که مشاهده می شود مساحت بخش هاشور زده را جهت تعیین ضریب وزن سنسور وسط در عملگر اول یعنی ایجاد ولتاژ بایاس برای هر دو موتور بکار می بریم. برای سایر سنسورها نیز به همین نحو عمل می کنیم , نتیجه حاصل چنین است:

$$OM=0.18SL+0.64SM+0.18SR \qquad (۵)$$

برای عملگرهای دوم و سوم میتوان به دو شکل عمل کرد , میتوان از تابع توزیع برش زده استفاده کرد و یا با تغییر مرکز ثقل تابع توزیع نرمال به ضرایب وزن مناسبی رسید.که ما با استفاده از روش دوم به نتایج عملی بهتری رسیدیم لذا همین روش را توضیح می دهیم.



شکل۱۰- استفاده از سطح زیر منحنی تابع توزیع نامتقارن در محاسبه ضریب وزنها

همانگونه که درشکل مشاهده می شود از تابع توزیع نامتقارن جهت محاسبه ضریب وزن سنسورها برای عملگر دوم و سوم که فرمان حرکت و کنترل ربات را به موتور سمت راست و چپ صادر می کند استفاده می کنیم. نتایج حاصله چنین است:

$$OL= 0.58\ SR+0.32\ SM+0.1\ SL$$
$$OR= 0.1\ SR+0.32\ SM+0.58\ SL \qquad (۶)$$

بدین ترتیب ما از این سه تابع توزیع جهت محاسبه عملگرها استفاده می کنیم و همان گونه که مشاهده شد ضریب وزن سنسورها با توجه به تعداد و موقعیت سنسورها محاسبه می شود.

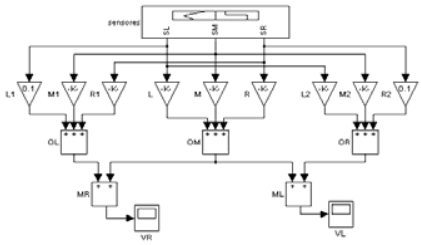
۴-۴- مزایا و نتایج حاصل از این روش

مزایای حاصل از این روش را می توان به دو بخش تقسیم کرد.

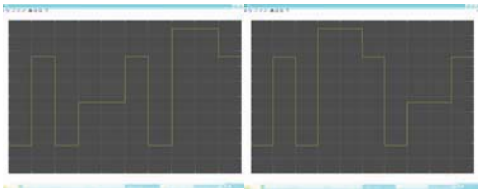
۱- مزایای ناشی از استفاده از سه سنسور که عبارت است از : موقعیت مین ها با دقت بیشتری پیدا شود، با جریان الکتریکی یکسان سطح و عمق بیشتری جستجو می شود، برای شرایط مختلف می توان موقعیت سنسورها را تغییر داد.

۲- مزایای ناشی از ترکیب داده ها به روش OWA عبارت است از: سادگی این روش در مقایسه با سایر روشهای هوشمند مثل شبکه های عصبی، عملکرد بهینه این روش در پیدا کردن مین، کاهش خطا و سرعت عمل این روش از آن جمله است.

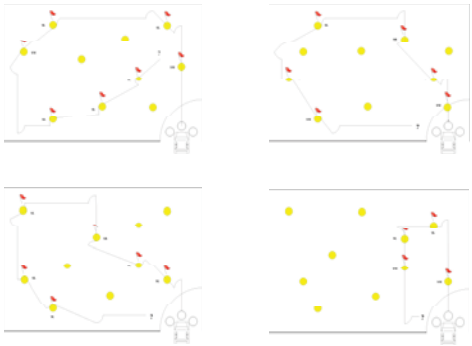
حال نتایج عملی و شبیه سازی شده این روش ارائه می شود که برای این منظور از محیط matlab استفاده شده و اطلاعات سنسورها را توسط سیکنال ژنراتور ایجاد کرده و نتیجه نهایی به صورت سطح ولتاژ های اعمال شده به هر یک موتورها جهت حرکت ربات ظاهر می شود.در ادامه نتایج این شبیه سازی مشاهده می شود.



شکل۱۱- دیاگرام بلوکی شبیه سازی شده توسط نرم افزار Matlab



شکل۱۲- ولتاژهای رسیده به هر یک از موتورها در برخورد با سه مین در موقعیت های مختلف وسط,چپ و راست



شکل۱۳- نتایج عملکرد ربات در زمینی با ده مین موجود

شکل ها سطح ولتاژ هر یک از موتورها را در رسیدن ربات به یک مین در مسیر مستقیم سپس تشخیص مین توسط سنسور سمت چپ و در نهایت تحریک سنسور سمت راست،که مشاهده می شود پس از رسیدن ربات به هر یک از مین های چپ و راست ربات تغییر مسیر حرکت داده و در راستای همان مین ادامه مسیر می دهد، حال میتوان اطلاعات هر یک از سنسورهای

تشخیص مانع را نیز وارد عمل کرده و در هنگام رسیدن ربات به مانع تغییر مسیر حرکت صورت گیرد تا از برخورد ربات با موانع جلوگیری شود.

همانگونه که در شکل مشاهده می شود ربات قادر است در بهترین شرایط هشت و در بدترین شرایط چهار مین را تشخیص دهد.

جدول ۱- نتایج عملی حاصل از آزمایش ربات

نتایج عملی حاصل از یک دور چرخش ربات در پیست های مختلف با تعداد ده مین	تعداد مینهای کشف شده در پیست ۱	تعداد مینهای کشف شده در پیست ۲	تعداد مینهای کشف شده در پیست ۳	تعداد مینهای کشف شده در پیست ۴
ربات با یک سنسور	۵	۴	۴	۴
ربات با سه سنسور روش شرطی	۵	۶	۵	۵
ربات با سه سنسور روش OWA	۷	۶	۶	۴

جهت جلوگیری از قرار گرفتن ربات در مسیر تکراری(لوپ) می توان اصلاحات نرم افزاری و سخت افزاری مختلفی انجام داد، بعنوان مثال با علامت گذاری مین های شناسایی شده توسط ربات و تعبیه سنسورهای مربوطه، ربات قادر است با قرار گرفتن در موقعیت مینی که علامت گذاری شده تغییر مسیر داده و سایر مین ها را جستجو نماید.

۵- نتیجه گیری

بکار گیری روش OWA در ربات مین یاب نتایج جالب توجهی را بدنبال دارد، همچنین استفاده از تابع توزیع گوسی در انتخاب ضریب وزنها محدودیت و مشکلات افزایش تعداد سنسورها و موقعیت آنها را منتفی می سازد. این روش در مقایسه با سایر روشهای هوشمند دارای هزینه کمتر و سرعت عمل بیشتری است. الگوریتم فوق می تواند کاربردهای دیگری نیز داشته باشد در واقع هر گاه بخواهیم از مجموعه ای از سنسورها جهت نمایش یا کنترل استفاده نماییم به نحوی که بتوان توسط تابع توزیع گوسی درجه اهمیت سنسورها را بیان کرد می توان از روش فوق استفاده کرد.

مراجع

[1] D. Koks and S. Challa, "An introduction to Bayesian and Dempster-Shafer data fusion, "Department of Defence, Edinburgh, Australia, Tech Rep. DSTO-TR-1436, 2005.

[2] J. K. Rosenblatt, "Optimal Selection of Uncertain Actions by Maximizing Expected Utility," *Autonomous Robots*, vol. 9, no. 1, pp. 17-25, 2000.

[3] M. A. Simard, E. Lefebvre and C. Helleur, "Multi-source information fusion applied to ship identification for the recognized maritime picture," *Sensor Fusion: Architectures, Algorithms and Applications*, vol. 4, pp. 67-78, 2000.

[4] R. Yager, "On ordered weighted averaging aggregation operators in multi criteria decision making," *IEEE Transaction on Systems, Man and Cybernetics*, vol. 18, no. 1, 1988.

[5] A. Abidi and R. C. Gonzalez ,"*Data Fusion in Robotic and Machine Intelligence*" Academic, Press,1999.

[6] V. Steinmeyz , F. Sevila , and V. Bellon - Maurel ,"A methodology for sensor fusion design: application to Fruit quality assessment," *Journal of Agriculture Research*,vol.74, pp.21-31- 1999.

[7] B. Dasaraty , *Decision Fusion*. IEEE Computer Society Press , 1994.

[8] K. Sentz and S. Ferson, "Combination of evidence in Dempster - Shafer theory ,"SANDIA National Laboratory ,Springfield ,USA, Tech Rep. SAND 2002 – 0835, 2002.

[9] M. Grabish , "Fuzzy integrals as a flexible and interpretable of aggregation," *in Aggregation and Fusion of Imperfect Information*, pp.51-72, 1998

[10] R. R. Yager, "Families of OWA operators," *Fuzzy Sets and Systems*, vol.59, no.2, pp. 125-148, 1993.

[11] H. Wu, M, Siegel, R. Stiefelhagen, and J. Yang, "Sensor fusion using Dempster-Shsfer theory", in *proceeding of IEEE Conference on Instrumentation and Measurement Technology*, Anchorage, USA, 2002.

[12] F. Kobayashi, F. Arai and F. Fukuda, "Sensor selection by reliability based on possibility measure," *Robotics and Automation*, pp. 2614-2619, 1999.

[13] R. R. Yager, "Intelligent decision making and information fusion," *Intelligent Systems*, vol. 4, pp 1-4, 2004.

[14] A. H. Keyhanipour, "Design and implementation of a new intelligenct meta-search engine based on information fusion theory," M.Sc. Thesis, School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran, 2006.

[15] M. Sugeno, "Fuzzy measures and Fuzzy integrals-A survey," *Fuzzy Automata and Decision Processes*, North-Holland, Amsterdam, pp. 89–102, 1977.

[16] B. Araabi, "Fuzzy statistical systems and their identification," M.Sc. Thesis, School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran, 1996.

[17] R. Osorio, C. Jose, A. Romero, M. Penaco , I. Lopez J, "Intelligent Line Follower Mini-Robot System" *International Journal of Computers, Communications &Control*. Vol, 1 (2006) no, 2, pp, 73-83, 2006.

[18] بادلو، مشیری، اعرابی، "ترکیب اطلاعات هوشمند توسط عملگر میانگین گیری مرتب وزندار در ربات تعقیب خط ساده و مغناطیسی" دومین کنگره مشترک سیستم های فازی و سیستم های هوشمند {هشتمین کنفرانس سیستم های فازی و نهمین کنفرانس سیستم های هوشمند} ص ۱۸۳ آبان ۱۳۸۷.

مجله کنترل جلد ۳، شماره ۱، بهار ۱۳۸۸، صفحه ۴۶-۳۷

ارائه یک سیستم آشوبناک بعد بالای جدید همراه با یک نقطه تعادل و پایدارسازی آن با استفاده از کنترل‌کننده فیدبک حالت خطی

علی ابویی [۱]، محمدرضا جاهدمطلق [۲]، زهرا رحمانی چراتی [۳]

[۱] دانشجوی کارشناسی ارشد، برق- کنترل، آزمایشگاه سیستم‌های پیچیده، دانشگاه علم‌و صنعت ایران، aliabooee@elec.iust.ac.ir

[۲] دانشیار، آزمایشگاه سیستم‌های پیچیده، دانشگاه علم و صنعت ایران ، jahedmr@iust.ac.ir

[۳] استادیار، دانشکده مهندسی برق، دانشگاه صنعتی نوشیروانی بابل ، Rahmaniz@nit.ac.ir

چکیده: در این مقاله یک سیستم آشوبناک بعد بالای جدید ارائه شده است. این سیستم، دارای یک نقطه تعادل در مبدأ بوده و ویژگی شاخص آن، وجود دو نمای لیاپانوف مثبت بزرگ در مقایسه با سیستم‌های آشوبناک بعد بالای دیگر می‌باشد. در ادامه، معیارهایی جهت اثبات وجود آشوب بعد بالا در این سیستم، مورد بررسی قرار گرفته‌اند. که از آن جمله معیارها می‌توان به بررسی اتلافی بودن سیستم، ناپایداری تنها نقطه تعادل سیستم، جاذب عجیب، نماهای لیاپانوف، بعد کسری، نگاشت پوانکاره، طیف فرکانسی گسترده و حساسیت شدید پاسخ‌های زمانی متغیرهای حالت سیستم به شرایط اولیه اشاره کرد. بررسی تمام این معیارها، نشان از وجود آشوب بعد بالا در این سیستم داشت. با تغییر یکی از پارامترهای سیستم، رفتارهای متفاوت دینامیکی را برای این سیستم آشوبناک نشان دادیم، که از آن جمله می‌توان به آشوب بعد پایین، سیکل حدی، شبه پریودیک و آشوب بعد بالا اشاره کرد. در انتها با استفاده از کنترل‌کننده فیدبک حالت خطی، سیستم آشوبناک حول نقطه تعادل خود، پایدار شده است.

کلمات کلیدی: سیستم آشوبناک بعد بالا- نمای لیاپانوف- بعد کسری- اتلافی بودن و جاذب عجیب

Abstract: A new hyperchaotic system which is represented in this paper has an equilibrium point located at origin. Having two large positive Lyapunov Exponents, is the prominent feature of this system in comparison to other hyperchaotic systems. Some basic dynamical properties are studied in order to prove existence of hyper chaos in this system including Dissipativeness of System, Instability of Unique Equilibrium Point, Strange Attractor, Lyapunov Exponents, Fractal Dimension, Poincare Mapping and Sensitivity of Time Response related to state variables to initial condition. All of the properties studied show that the system is hyperchaotic. By changing a parameter of the system, various dynamical characteristics is obtained such as Chaos, limit Cycle, Quasi-Periodic and Hyperchaos. At last, the chaotic system is stabilized around its equilibrium point by using linear state feedback controller

Keywords: Hyperchaotic System, Lyapunov Exponent, Fractal Dimension, Dissipativeness and Strange Attractor

۱- مقدمه

در دو دهه‌ی اخیر تحقیقات زیادی جهت معرفی سیستم‌های آشوبناک جدید و تحلیل رفتار آشوبگونه در این نوع سیستم‌ها صورت گرفته است[۱-۱۲]. با توجه به کاربرد گسترده آشوب در سیستم‌های مهندسی، دو زمینه تحقیقاتی جدید در ارتباط با آشوب باز شده است که این دو زمینه، یکی کنترل آشوب [۲و۱] و دیگری سنکرون‌سازی[۱] آشوب [۴و۳] می‌باشد. سیگنال آشوبناک با یک نمای لیاپانوف[۲] مثبت برای انتقال امن داده‌ها در مخابرات مورد استفاده قرار می‌گرفت. تا اینکه در سال ۱۹۹۵ پرز[۳] و سردیرا[۴] نشان دادند که انتقال داده‌ها با سیگنال

[1] Synchronization
[2] Lyapunov Exponent
[3] Perez
[4] Cerdeira

آشوبناک امن نیست و می‌توان داده‌ها را از آن استخراج کرد[۵]. برای غلبه بر این مشکل، سیستم‌های آشوبناک بعد بالا[1] معرفی شدند. سیستم های آشوبناک بعد بالا به علت افزایش تصادفی بودن[2] و بالا بودن عدم قابلیت پیش بینی[3] در آنها، جایگزین سیگنال‌های آشوب بعد پایین در مخابرات امن شدند[۶]. سیستم‌های آشوبناک بعد بالای پیوسته زمان[4]، حداقل دارای ۴ متغیر حالت بوده و ویژگی شاخص این سیستم‌ها وجود دو نمای لیاپانوف مثبت می‌باشد. این ویژگی باعث می‌شود که دینامیک این نوع سیستم‌ها در بیش از یک جهت به طور همزمان گسترش یابد. به علت دینامیک پیچیده سیستم‌های آشوبناک بعد بالا، تحقیقات زیادی در علوم مهندسی بر روی این نوع از سیستم‌ها انجام شده است. به عنوان نمونه، آشوب بعد بالا در مواردی مانند اسیلاتورهای کلپیتس[5] [۷]، مدارهای غیرخطی [۸]، لیزرها [۹] و مخابرات امن [۱۰] دارای کاربرد فراوان است.

آشوب بعد بالا اولین بار توسط راسلر[6] در [۱۱] ارائه شد. در سال ۲۰۰۳ ، کافاگنا[7] و گراسی[8] یک روش برای تولید جاذب عجیب[9] آشوبناک بعد بالا، با استفاده از زنجیره‌ای از مدارهای چوآی[10] کوپل شده معرفی کردند [۱۲]. در این ساختار از سه مدار چوآی بهم پیوسته همراه با یک تحریک سینوسی برای تولید آشوب بعد بالا استفاده شده است. تاکنون هیچ روش کلی برای تولید و طراحی سیستم‌های آشوبناک بعد بالا ارایه نشده است. یکی از روش‌های متداول برای طراحی سیستم آشوبناک بعد بالا که اخیراً در مقالات [۲۱-۱۳] به آن اشاره شده است، بدین صورت است که یک سیستم آشوبی بعد پایین با ۳ متغیر حالت را در نظر گرفته و با اضافه کردن کنترل‌کننده فیدبک حالت و تنظیم دوباره ضرایب سیستم، آشوب بعد بالا را در سیستم ایجاد می‌کنند. به عنوان نمونه لی [11] دو سیستم آشوبناک با عناوین " سیستم آشوبناک بعد بالای لورنز تعمیم یافته" [۱۳] و " سیستم آشوبناک بعد بالای چن[12] اصلاح شده" [۱۴] را معرفی کرده است. این دو سیستم آشوبناک بعد بالا، با اضافه شدن یک متغیر حالت جدید به سیستم‌های آشوبناک بعد پایین لورنز[13] و چن ساخته شده‌اند. در [۱۵] با اضافه شدن یک کنترل‌کننده به معادله دوم سیستم آشوبناک چن، سیستم آشوبناک بعد بالای جدیدی معرفی شده است. مقاله [۱۶] با اضافه کردن یک کنترل‌کننده مربعی غیرخطی[14] به معادله دوم سیستم آشوبناک لورنز، سیستم آشوبناک بعد بالای دیگری را ارائه داده است. اخیراً در [۱۷] یک سیستم آشوبناک بعد بالای جدید ارائه شده است که دارای نقطه تعادل یکتا بوده و می-تواند رفتار پریودیک[15]، شبه پریودیک[16]، آشوب بعد پایین و آشوب بعد بالا از خود نشان دهد. این سیستم آشوبناک جدید با اضافه شدن یک کنترل‌کننده فیدبک حالت به سیستم آشوبناک بعد پایین کی[17] [۱۸] ساخته می‌شود. در سال ۲۰۰۶، چن و وانگ[18] دو سیستم آشوبناک بعد بالا، با اضافه کردن یک متغیر حالت چهارم به سیستم آشوبناک لو[19] ارایه دادند [۲۰و۱۹]. در سال ۲۰۰۷ ونجوآن وو[20] یک سیستم آشوبناک بعد بالای جدید ارایه داد که تمام رفتارهای دینامیکی ممکنه را برای یک سیستم آشوبناک شامل می‌شد [۲۱].

یکی دیگر از روش‌های تولید سیستم آشوبناک بعد بالا، ایجاد اغتشاش‌های کوچک در پارامترهای متغیر با زمان سیستم‌های آشوب با بعد پایین می‌باشد که به عنوان نمونه در [۲۲] با اضافه شدن یک ورودی کنترلی سینوسی به سیستم آشوبناک یکپارچه[21] بعد پایین، سیستم آشوبناک بعد بالای جدیدی ساخته شده است. رفتار آشوبناک بعد بالا در شبکه‌های عصبی مصنوعی[22] نیز دیده شده است که می‌توان به آشوب بعد بالا در شبکه‌های عصبی مصنوعی از نوع هوپفیلد[23] اشاره کرد. در [۲۳] یک شبکه عصبی از نوع هوپفیلد با ۴ نرون در نظر گرفته شده است و با تنظیم ضرایب وزنی میان نرون‌ها، آشوب بعد بالا در این شبکه ایجاد شده است. تمام سیستم‌های آشوبناک بعد بالا که بدان اشاره شد، دارای تعبیر فیزیکی خاصی نبوده و تنها معادلات ریاضی صرف می‌باشند. اما آشوب بعد بالا در سیستم‌های فیزیکی نیز رخ می‌دهد. مقاله‌ی [۲۴] - وجود آشوب بعد بالا را در معادلات اسیلاتور شبکه عصبی سلولی کوانتومی[24] نشان می‌دهد. ویژگی شاخص این سیستم آشوبناک بعد بالا، وجود سه نمای لیاپانوف مثبت می‌باشد.

در این مقاله، ویژگی شاخص سیستم آشوبناک بعد بالا، بزرگ بودن دو نمای لیاپانوف مثبت آن در مقایسه با اغلب سیستم‌های

[1] Hyperchaotic Systems
[2] Increasing Randomness
[3] Higher Unpredictability
[4] Hyperchaotic System In Continuous Time
[5] Colpitts Oscillators
[6] Rössler
[7] Cafagna
[8] Grassi
[9] Strange Attractor
[10] Chua's Circuits
[11] Li
[12] Chen
[13] Lorenz
[14] Nonlinear Quadratic Controller
[15] Periodic
[16] Quasi Periodic
[17] Qi
[18] Wang
[19] Lü
[20] Wenjuan Wu
[21] Unified Chaotic System
[22] Artificial Neural Networks
[23] Hopfield
[24] Quantum Cellular Neural Network Oscillator

آشوبناک بعد بالای دیگر می‌باشد. این ویژگی، در یک بخش جداگانه از مقاله مورد بررسی قرار خواهد گرفت.

ساختار کلی این مقاله بدین صورت است که در بخش ۲، پس از معرفی دینامیک سیستم آشوبناک بعد بالا، جهت اثبات وجود آشوب در این سیستم، نشان می‌دهیم که سیستم اتلافی و نقطه تعادل آن ناپایدار است. در ادامه این بخش به بررسی جاذب عجیب، پاسخ‌های زمانی، نماهای لیاپانوف، حساسیت به شرایط اولیه، بعد کسری[1] ، نگاشت پوانکاره[2] و طیف فرکانسی گسترده[3] سیستم می‌پردازیم. در بخش ۳ رفتارهای مختلف دینامیکی را با تغییر یکی از پارامترهای سیستم نشان می‌دهیم. بخش ۴ به مقایسه میان نماهای لیاپانوف این سیستم با سیستم‌های آشوبناک دیگر اختصاص یافته است. در بخش ۵ با طراحی کنترل‌کننده فیدبک حالت خطی[4] سیستم حول نقطه تعادل خود پایدار می‌شود و نتیجه‌گیری کلی از مقاله در بخش ۶ ارایه می‌گردد.

۲- معادلات دینامیکی سیستم آشوبناک بعد بالای جدید

معادلات دینامیکی این سیستم آشوبناک بعد بالا با الهام از سیستم آشوبناک بعد پایین لیو [۲۷] ساخته شده است. معادلات این سیستم آشوبناک جدید به صورت رابطه (۱) قابل بیان است. این معادلات در واقع با اضافه کردن متغیر حالت چهارم w و افزودن چند ترم غیرخطی به معادلات سیستم آشوبناک لیو، تشکیل شده‌اند.

$$\begin{cases} \dot{x} = a(y-x) + byz^2 = f_1(x,y,z,w) \\ \dot{y} = cx + dxz^2 + ew = f_2(x,y,z,w) \\ \dot{z} = fz + gy^2 + hxw = f_3(x,y,z,w) \\ \dot{w} = ky = f_4(x,y,z,w) \end{cases} \quad (۱)$$

در رابطه (۱) x, y, z, w بیانگر متغیرهای حالت سیستم آشوبناک می‌باشند. با انتخاب پارامترها به صورت رابطه (۲)، سیستم رفتار آشوب بعد بالا از خود نشان می‌دهد.

$$\begin{aligned} &a = 7.7,\ b = -1,\ c = 8,\ d = 4,\ e = 8 \\ &f = -4,\ g = 1,\ h = 1,\ k = -2 \end{aligned} \quad (۲)$$

برای وجود آشوب در یک سیستم شرایط زیر الزامی است:

الف- سیستم باید اتلافی[5] باشد. اتلافی بودن به مفهوم این است که انرژی سیستم در حال کاهش بوده و سیستم پایدار کلی[6] است.

ب- سیستم باید نقاط تعادل ناپایدار داشته باشد. ماتریس ژاکوبین[7] محاسبه شده در نقاط تعادل، باید دارای مقادیر ویژه ناپایدار باشد. این نکته در واقع بیانگر ناپایداری محلی[8] سیستم است.

ج- مدارهای[9] سیستم باید محدود و کراندار باشند.

در ادامه، این شرایط بر روی سیستم مورد بررسی قرار خواهند گرفت.

۲-۱- بررسی اتلافی بودن سیستم

خاصیت حفظ سطح و یا حجم در فضاهای بالاتر، مشخصه‌ی کلی سیستم‌های پایستار[10] می‌باشد. این ویژگی، سیستم‌های دینامیکی را بر حسب اینکه حجم‌های فضای فاز ثابت بماند و یا کاهش یابند به ترتیب به دو گروه پایستار و اتلافی تقسیم می‌کند.

همان‌طوری که در بخش دوم مقاله گفته شد، یکی از شرایط لازم برای وجود آشوب در یک سیستم این است که سیستم اتلافی باشد. چنانچه معادلات دینامیکی یک سیستم به صورت رابطه (۳) باشد، برای بررسی اتلافی بودن، ترم $\nabla F = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}$ را محاسبه کرد، چنانچه این مقدار صفر باشد سیستم پایستار و اگر این مقدار منفی باشد سیستم اتلافی است.

$$\begin{cases} \dot{x}_1 = f_1(x_1, x_2, \ldots, x_n) \\ \dot{x}_2 = f_2(x_1, x_2, \ldots, x_n) \\ \quad \ldots \\ \dot{x}_n = f_n(x_1, x_2, \ldots, x_n) \end{cases} \quad (۳)$$

رابطه (۴) شرط اتلافی بودن سیستم را بررسی می‌کند و نشان می‌دهد که سیستم اتلافی است.

$$\begin{aligned} &\nabla F = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} + \frac{\partial f_4}{\partial w} \Rightarrow \\ &\nabla F = -7.7 - 4 = -11.4 \Rightarrow \nabla F < 0 \end{aligned} \quad (۴)$$

با توجه به این که $\nabla F < 0$ می‌باشد، بنابراین سیستم، اتلافی و پایدار کلی است.

[1] Fraction Dimension
[2] Poincare Map
[3] Continuous Spectrum
[4] Linear Feedback Controller

[5] Dissipative
[6] Globally Stable
[7] Jacobian Matrix
[8] Local Instability
[9] Orbits
[10] Conservative

۲-۲- بررسی ناپایداری نقطه تعادل سیستم

با توجه $f_i = 0,\ i = 1,2,3,4$، سیستم تنها دارای یک نقطه تعادل در $E = (0,0,0,0)$ می‌باشد. ماتریس ژاکوبین سیستم در این نقطه تعادل، به صورت رابطه (۵) محاسبه می‌شود.

$$J_{(0,0,0,0)} = \begin{bmatrix} -7.7 & 7.7 & 0 & 0 \\ 8 & 0 & 0 & 8 \\ 0 & 0 & -4 & 0 \\ 0 & -2 & 0 & 0 \end{bmatrix} \tag{۵}$$

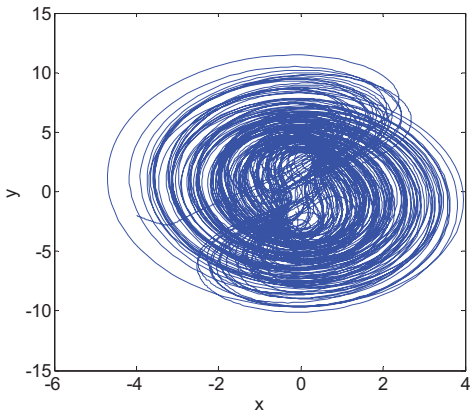
مقادیر ویژه ماتریس ژاکوبین بالا، در رابطه (۶) آورده شده‌اند.

$$\begin{cases} \lambda_1 = -12.2454 \\ \lambda_2 = 2.2727 + 2.2126i \\ \lambda_3 = 2.2727 + 2.2126i \\ \lambda_4 = -4 \end{cases} \tag{۶}$$

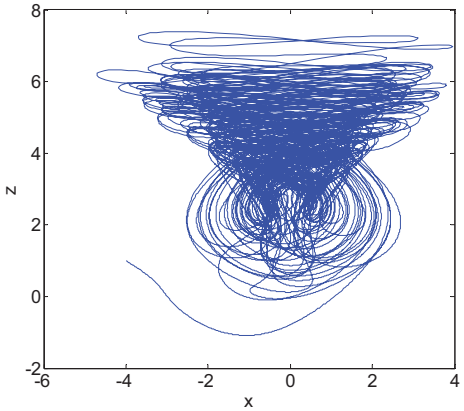
با توجه به علامت مقادیر ویژه ماتریس ژاکوبین سیستم، نقطه تعادل سیستم از نوع زینی شکل و ناپایدار می‌باشد. بنابراین می‌توان نتیجه گرفت که سیستم به طور محلی ناپایدار است.

۲-۳- جاذب عجیب سیستم آشوبناک بعد بالا

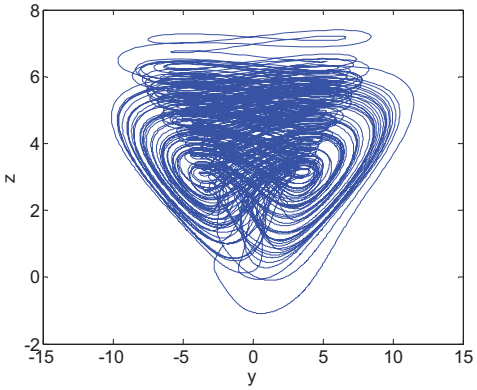
این سیستم آشوبناک با استفاده از نرم افزار MATLAB، شبیه‌سازی شده است و تعدادی از تصویرهای جاذب عجیب این سیستم بر روی فضای دو بعدی و سه بعدی به صورت شکل‌های (۱) تا (۶) نتیجه شده‌اند. شرایط اولیه برای شبیه‌سازی سیستم آشوبناک به صورت $(x_0, y_0, z_0, w_0) = (-4,-2,1,3)$ در نظر گرفته شده است.



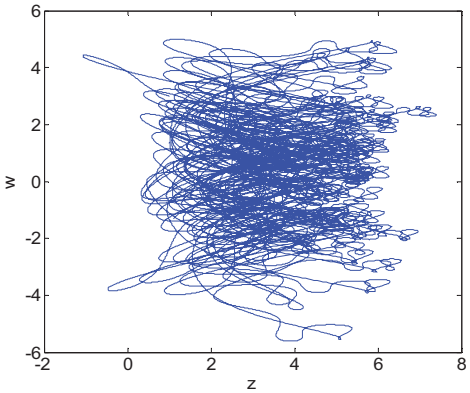
شکل(۱): تصویر جاذب عجیب سیستم در فضای فاز $(x-y)$



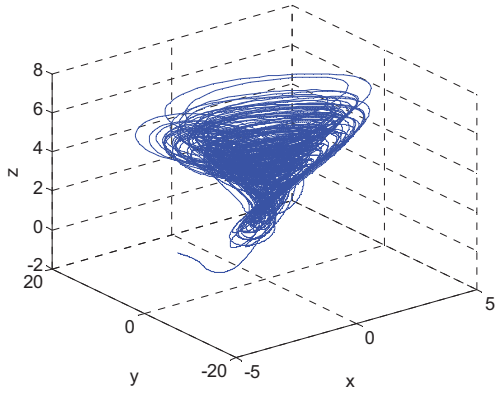
شکل(۲): تصویر جاذب عجیب سیستم در فضای فاز $(x-z)$



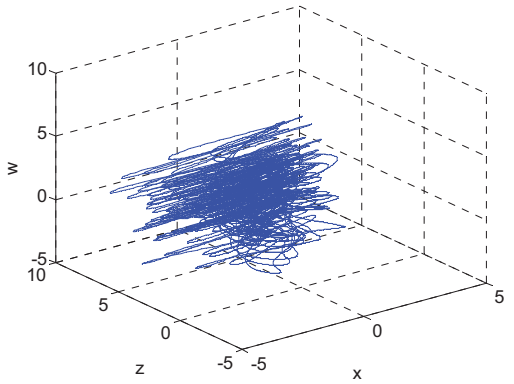
شکل(۳): تصویر جاذب عجیب سیستم در فضای فاز $(y-z)$



شکل(۴): تصویر جاذب عجیب سیستم در فضای فاز $(z-w)$



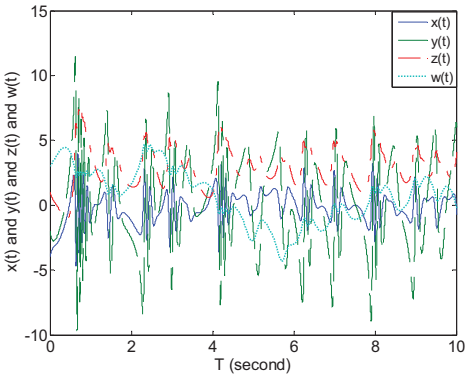
شکل(۵): تصویرجاذب عجیب سیستم در مختصات ($x-y-z$)



شکل(۶): تصویرجاذب عجیب سیستم در مختصات ($x-z-w$)

۲-۴- پاسخ‌های زمانی متغیرهای حالت سیستم آشوبناک بعد بالا

شکل (۷) پاسخ‌های زمانی متغیرهای حالت سیستم آشوبناک بعد بالا را نشان می‌دهد. شرایط اولیه برای شبیه‌سازی سیستم آشوبناک به صورت- $(x_0, y_0, z_0, w_0) = (-4, -2, 1, 3)$ درنظر گرفته شده است.



شکل(۷): پاسخ زمانی متغیرهای حالت سیستم آشوبناک

۲-۵- بررسی نماهای لیاپانوف سیستم آشوبناک بعد بالا

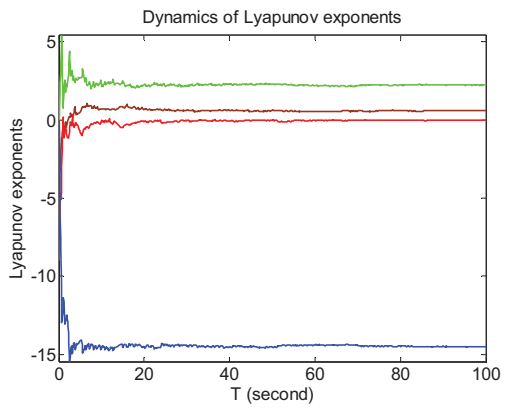
نمای لیاپانوف، یک کمیت اندازه‌گیری است که میزان حساسیت دینامیکی سیستم را به شرایط اولیه مشخص می‌کند. این کمیت در واقع نرخ متوسط همگرایی و واگرایی دو مسیر نزدیک به هم را در فضای فاز مشخص می‌سازد و یک کمیت استاندارد جهت تعیین آشوبگونه بودن یا نبودن یک سیستم است. مسیرهای یک سیستم آشوبناک در فضای حالت، دارای طول بینهایت هستند که در یک فضای محدود محصور شده‌اند لذا باید مسیرهای یک سیستم آشوبناک در بعضی جهات واگرا و در بعضی جهات همگرا شوند، نماهای لیاپانوف برای بررسی کمی واگرایی و همگرایی مسیرهای حالت سیستم استفاده می‌شوند.

جدول (۱) حالت‌های دینامیکی را برای یک سیستم آشوبناک با ۴ متغیر حالت بر حسب علامت نماهای لیاپانوف نشان می‌دهد. در جدول(۱) L_i , $i = 1,2,3,4$، بیانگر i امین نمای لیاپانوف سیستم می‌باشد [۲۱].

جدول (۱) : حالت‌های دینامیکی یک سیستم آشوبناک با ۴ متغیر حالت بر حسب علامت نماهای لیاپانوف[۲۱]

نوع رفتار دینامیکی	L_1	L_2	L_3	L_4
نقطه تعادل	–	–	–	–
سیکل‌حدی	0	–	–	–
شبه‌پریودیک	0	0	–	–
3 torus	0	0	0	–
رفتار آشوبی	+	0	–	–
آشوب بعد بالا	+	+	0	–

نماهای لیاپانوف این سیستم با استفاده از جعبه ابزار MATDS که در محیط نرم افزاری MATLAB قابل اجرا است، محاسبه شده‌اند. شکل (۸) نماهای لیاپانوف را برای این سیستم آشوبناک، نشان می‌دهد.



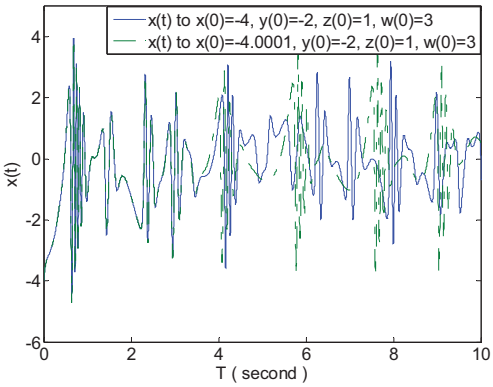
شکل(۸): نماهای لیاپانوف سیستم آشوبناک بعد بالا

همانطوری که در شکل (۸) دیده می‌شود، سیستم دارای دو نمای لیاپانوف مثبت، یک نمای لیاپانوف صفر و یک نمای لیاپانوف منفی است و با توجه به جدول (۱)، این حالت بیانگر آشوب بعد بالا در این سیستم می‌باشد. رابطه (۷) مقادیر ۴ نمای لیاپانوف این سیستم را نشان می‌دهد.

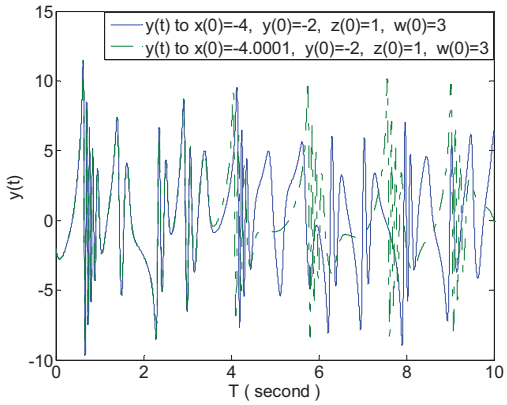
$$\begin{cases} L_1 = 2.2316 \\ L_2 = 0.59014 \\ L_3 = 0.022414 \\ L_4 = -14.4994 \end{cases} \qquad (۷)$$

۲-۶- حساسیت شدید سیستم آشوبناک بعد بالا به شرایط اولیه

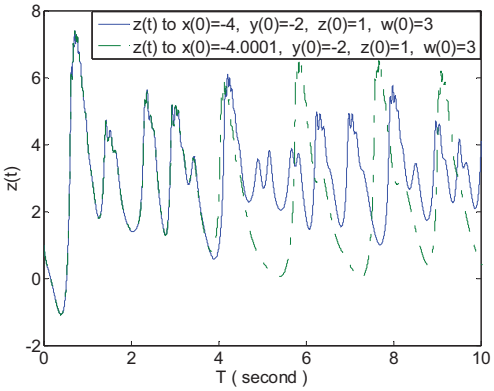
سیستم آشوبناک حساسیت زیادی به شرایط اولیه دارد و تغییر کوچکی در شرایط اولیه سیستم، باعث می‌شود که پاسخ‌های زمانی متغیرهای حالت سیستم آشوبناک متفاوت شوند. این مطلب در شکل (۹)، (۱۰)، (۱۱) و (۱۲) نشان داده شده است. در این حالت شرایط اولیه سیستم از $(x_0, y_0, z_0, w_0) = (-4, -2, 1, 3)$ به $(x_0, y_0, z_0, w_0) = (-4.0001, -2, 1, 3)$ تغییر کرده‌اند.



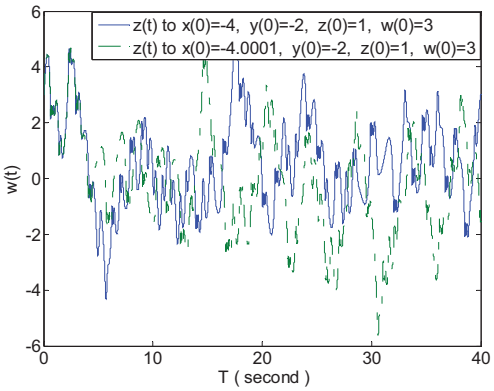
شکل (۹) : پاسخ‌های زمانی متغیر $x(t)$ سیستم آشوبناک با توجه به تغییر شرایط اولیه



شکل (۱۰) : پاسخ‌های زمانی متغیر $y(t)$ سیستم آشوبناک با توجه به تغییر شرایط اولیه



شکل (۱۱) : پاسخ‌های زمانی متغیر $z(t)$ سیستم آشوبناک با توجه به تغییر شرایط اولیه



شکل (۱۲): پاسخ‌های زمانی متغیر $w(t)$ سیستم آشوبناک با توجه به تغییر شرایط اولیه

۲-۷- بعد کسری سیستم آشوبناک بعد بالا

با در نظر گرفتن مقادیر نماهای لیاپانوف سیستم آشوبناک، بعد کاپلان- یورکه[1] سیستم با استفاده از رابطه (۸) محاسبه می‌شود [۲۵]. بعد کاپلان- یورکه همراه با نماهای لیاپانوف ارتباط مهمی بین هندسه فرکتالی جاذب عجیب و ویژگی حساسیت شدید به شرایط اولیه برقرار می‌سازند. در واقع از نماهای لیاپانوف همراه با بعد کاپلان-یورکه می‌توان برای مطالعه شرایط ایجاد آشوب در یک سیستم با تغییر یکی از پارامتر-های آن استفاده کرد. همان طوری که می‌بینیم بعد سیستم با استفاده از رابطه (۸) به صورت یک عدد غیرصحیح نتیجه شده است و این یکی از ویژگی‌های سیستم آشوبناک می‌باشد.

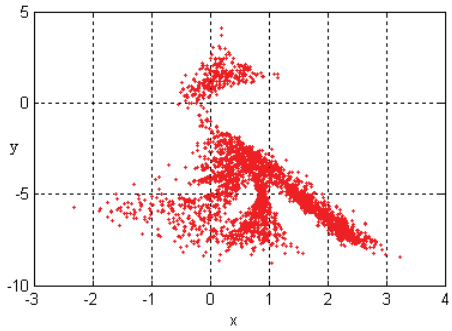
$$\begin{aligned} D_{KY} &= j + \frac{1}{|L_{j+1}|}\sum_{i=1}^{j} L_i \\ &= 3 + \frac{(L_1 + L_2 + L_3)}{|-14.4994|} \\ &= 3 + \frac{(2.2316 + 0.59014 + 0)}{14.4994} = 3.1946 \end{aligned} \qquad (۸)$$

[1] Kaplan-Yorke Dimension

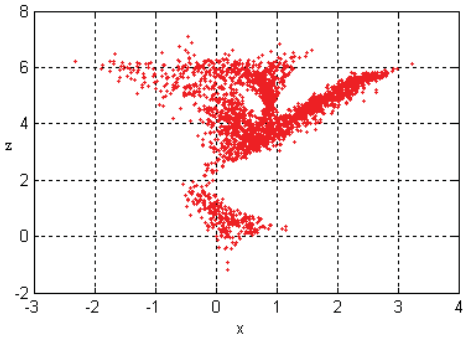
در رابطه (۸) L_i ها $(L_1 > L_2 > L_3 > L_4)$ بیانگر مقادیر نماهای لیاپانوف سیستم می‌باشند که در رابطه (۷) آورده شده‌اند و j اندیس کوچکترین نمای لیاپانوف نامنفی سیستم بوده که در اینجا با توجه به رابطه (۷)، $j = 3$ در نظر گرفته شده است.

۲-۸- نگاشت پوانکاره سیستم آشوبناک بعد بالا

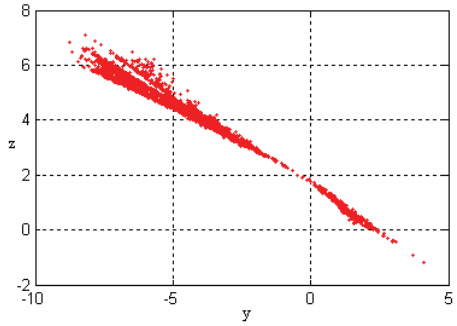
صفحه پوانکاره مورد نظر به صورت $x + 2y + 3z = 5$ در نظر گرفته شده است و به عنوان نمونه، تصویر نگاشت پوانکاره سیستم بر روی صفحات $x - y$، $x - z$ و $y - z$ در شکل‌های (۱۳)، (۱۴) و (۱۵) آورده شده است.



شکل(۱۳): تصویر نگاشت پوانکاره سیستم آشوبناک بعد بالا برروی صفحه $x - y$



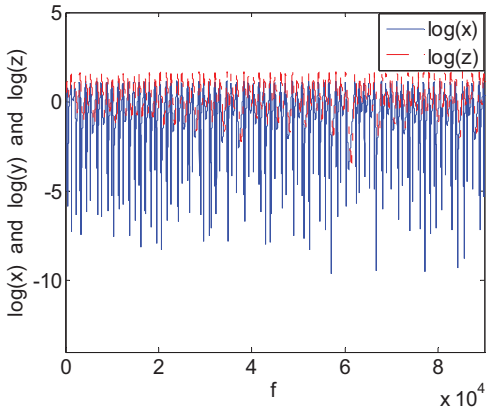
شکل(۱۴): تصویر نگاشت پوانکاره سیستم آشوبناک بعد بالا برروی صفحه $x - z$



شکل(۱۵): تصویر نگاشت پوانکاره سیستم آشوبناک بعد بالا برروی صفحه $y - z$

۲-۹- طیف فرکانسی گسترده سیستم آشوبناک

شکل (۱۶) نمودار طیف فرکانسی متغیرهای x و z سیستم آشوبناک را به عنوان نمونه نشان می‌دهد. می‌دانیم که سیستم‌های آشوبناک دارای طیف فرکانسی گسترده هستند. در این شکل نیز این مطلب به وضوح دیده می‌شود.

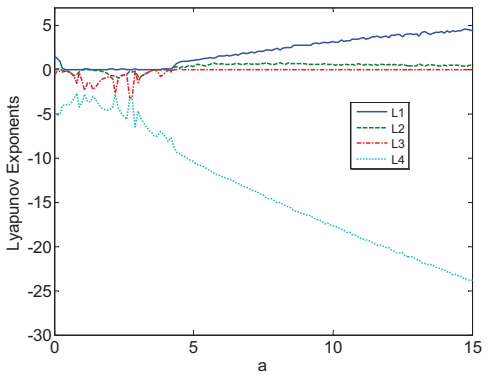


شکل(۱۸): نمودار طیف فرکانسی متغیرهای x و z سیستم آشوبناک بعدبالا

۳- ایجاد رفتارهای دینامیکی گوناگون برای سیستم آشوبناک بعد بالا با استفاده از تغییر پارامتر a

نوع رفتار دینامیکی سیستم آشوبناک از طریق مقادیر نمای لیاپانوف سیستم مشخص می‌شود. در این بخش از مقاله، با تغییر پارامتر a در بازه $[0,15]$ و ثابت نگه داشتن دیگر پارامترهای سیستم، مقادیر نماهای لیاپانوف سیستم، محاسبه شده‌اند. از جعبه ابزار Lab432 1.3 که در محیط نرم‌افزاری MATLAB قابل اجراست برای محاسبه مقادیر نماهای لیاپانوف سیستم بر حسب تغییرات پارامتر a استفاده شده است. این جعبه ابزار از الگوریتم ولف[1] [۲۸] برای محاسبه مقادیر نماهای لیاپانوف استفاده می‌کند و همگرایی مقادیر نماهای لیاپانوف سیستم به ازای هر مقدار پارامتر a تضمین شده است. شکل (۱۷) مقادیر نماهای لیاپانوف را بر حسب پارامتر a نشان می‌دهد.

[1] Wolf



شکل(۱۷): مقادیر نماهای لیاپانوف سیستم بر حسب پارامتر a

با توجه به شکل (۱۷) و جدول (۱)، نوع رفتار دینامیکی این سیستم آشوبناک به صورت جدول (۲) نتیجه می‌شود.

جدول(۲): نوع رفتار سیستم بر حسب تغییرات پارامتر a

نوع رفتار دینامیکی سیستم	محدوده تغییرات پارامتر a
آشوب بعد پایین	$0 \le a \le 0.3$
سیکل حدی (پریودیک)	$0.4 \le a \le 0.9$
شبه پریودیک	$1 \le a \le 1.4$
سیکل حدی (پریودیک)	$1.5 \le a \le 3.5$
شبه پریودیک	$3.7 \le a \le 3.9$
آشوب بعد بالا	$4.3 \le a \le 15$

۴- مقایسه نماهای لیاپانوف مثبت سیستم آشوبناک این مقاله با دیگر سیستم‌های آشوبناک بعد بالا

همان طوری که در ابتدای مقاله نیز بیان شد ویژگی شاخص این سیستم آشوبناک، بزرگ بودن نماهای لیاپانوف مثبت آن در مقایسه با سیستم‌های آشوبناک دیگر می‌باشد. برای نشان دادن این مطلب دو جدول در ادامه آورده شده است. جدول(۳) بیانگر نماهای لیاپانوف مثبت چندین سیستم آشوبناک بعد بالای معروف بوده و جدول (۴) نماهای لیاپانوف مثبت سیستم آشوبناک این مقاله را برای چند پارامتر مختلف a نشان می‌دهد. مقایسه میان این دو جدول درستی ادعای ما را تایید می‌کند.

جدول(۳): چند سیستم آشوبناک بعد بالا همراه با مقادیر دو نمای لیاپانوف مثبت آنها

سیستم آشوبی	L_1	L_2
راسلر[۱۱]	0.11	0.02
کافاگنا [۱۲]	0.774	0.3120
لی [۱۳]	0.6317	0.0175
چن [۱۷]	4.4090	0.1310
وانگ [۲۶]	1.0181	0.4180
جا [۱۶]	0.969	0.042

جدول(۴): مقادیردو نمای لیاپانوف مثبت سیستم آشوبناک این مقاله با در نظر گرفتن چند مقدار متفاوت برای پارامتر a

پارامتر a	L_1	L_2
17.7	5.0099	0.3284
17.9	5.1291	0.3918
18.8	5.2044	0.3379

۵- کنترل آشوب بعد بالا برای سیستم آشوبناک

فرض می‌کنیم که معادلات یک دسته از سیستم‌های آشوبناک بعد بالای غیرخطی پیوسته زمان به صورت رابطه (۹) می‌باشد.

$$\begin{cases} \dot{X} = f(X,t) \\ X(0) = X_0 \in R^n \end{cases} \qquad (۹)$$

در رابطه (۹) بردار X به صورت $X = (x, y, z, w)$ در نظر گرفته شده است. در اینجا قصد داریم تا کنترل‌کننده فیدبک خطی $v = v(X, k)$، را چنان تعیین کنیم که سیستم کنترل شده‌ی $\dot{X} = f(X,t) + v$، به سمت نقطه تعادل ناپایدار یا مدارهای پریودیک[1] خود، همگرا شود. چنانچه $v = 0$ فرض شود، سیستم کنترل شده به سیستم آشوبناک اولیه تبدیل خواهد شد [۲۵]. کنترل‌کننده فیدبک خطی را به صورت $v = k(x + z - w)$ در نظر گرفته و این کنترل‌کننده را به معادله دوم سیستم آشوبناک اضافه می‌کنیم. این کنترل‌کننده با روش سعی و خطا و جستجو میان توابع مختلفی از متغیر-های حالت، استخراج شده است. در رابطه کنترل‌کننده، k بهره فیدبک بوده و باید به گونه‌ای تعیین شود که سیستم به سمت نقطه تعادل خود

[1] Period Orbits

همگرا شود. رابطه (۱۰)، سیستم آشوبناک را همراه با کنترل‌کننده فیدبک خطی نشان می‌دهد.

$$\begin{cases} \dot{x} = 7.7(y-x) - yz^2 \\ \dot{y} = 8x + 4xz^2 + 8w + k(x+z-w) \\ \dot{z} = -4z + y^2 + xw \\ \dot{w} = -2y \end{cases} \quad (۱۰)$$

معادله مشخصه[1] ماتریس ژاکوبین سیستم کنترل شده در نقطه تعادل مبدأ، به صورت رابطه (۱۱) نتیجه می‌شود. در این رابطه مقادیر c_1, c_2, c_3 به صورت $c_1 = 7.7$، $c_2 = -9.7k - 45.6$، $c_3 = -15.4k + 123.2$ می‌باشند.

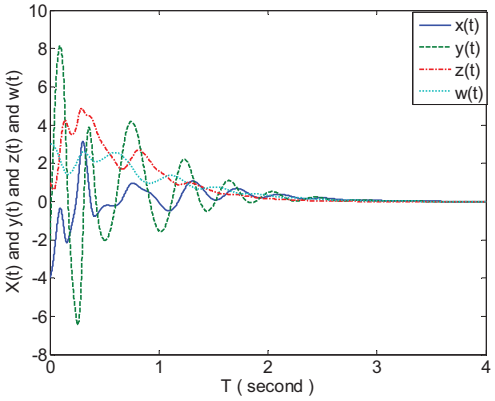
$$(\lambda+4)(\lambda^3 + c_1\lambda^2 + c_2\lambda + c_3) = 0 \quad (۱۱)$$

برای اینکه سیستم کنترل شده در نقطه تعادل مبدأ پایدار مجانبی باشد، باید تمام ریشه‌های چندجمله‌ای درجه سوم رابطه (۱۱) در سمت چپ محور موهومی قرار داشته باشند. بنابراین با توجه به معیار روث-هرویتس[2]، باید شرایط رابطه (۱۲) برای ضرایب این چند جمله‌ای برقرار باشد[۲۵].

$$\begin{aligned} &(1)\ c_i > 0\ (i = 1,3) \\ &(2)\ c_1 c_2 > 0 \end{aligned} \quad (۱۲)$$

با توجه به مقادیر c_1, c_2, c_3 و رابطه (۱۲)، محدوده‌ی مورد نظر برای بهره فیدبک k، به گونه‌ای که سیستم کنترل شده به مبدا همگرا شود، به صورت $k < -7.6$ نتیجه می‌شود. شکل (۱۸) پاسخ-های زمانی چهار متغیر حالت سیستم کنترل شده را برای بهره فیدبک $k = -30$ نشان می‌دهد. شرایط اولیه سیستم برای شبیه سازی به صورت $(x_0, y_0, z_0, w_0) = (-4, -2, 1, 3)$ اعمال شده است. نماهای لیاپانوف سیستم کنترل شده با بهره فیدبک $k = -30$، به صورت $L_1 = -2.8$ ، $L_2 = -2.57$، $L_3 = -2.6$ ، $L_4 = -4$ نتیجه می‌شوند که همگی منفی بوده و تایید کننده این مطلب هستند که سیستم کنترل شده، به سمت مبدا همگرا می‌شود.

[1] Characteristic Equation
[2] Routh-Hurwitz



شکل(۱۸): پاسخ‌های زمانی ۴ متغیر حالت سیستم کنترل شده با بهره فیدبک

$k = -30$

۶- نتیجه‌گیری

در این مقاله ضمن ارایه یک سیستم آشوبناک بعد بالای جدید، به صورت تحلیلی و شبیه‌سازی وجود آشوب بعد بالا در این سیستم نشان داده شد. مقایسه دو جدول ارایه شده در بخش ۴ نشان داد که نماهای لیاپانوف مثبت این سیستم در مقایسه با سیستم‌های آشوبناک بعد بالای دیگر بزرگتر می‌باشند. همچنین با محاسبه نقاط تعادل نشان دادیم که این سیستم دارای یک نقطه تعادل یکتا است. این دو ویژگی شاخص باعث می‌شوند که این سیستم آشوبناک به عنوان یک شاخص ارزیابی[3] خوب برای پیاده‌سازی روش‌های کنترل و سنکرون‌سازی آشوب بعد بالا مورد استفاده قرار گیرد. به عنوان نمونه، کنترل‌کننده فیدبک حالت خطی برای پایدارسازی این سیستم حول نقطه تعادل، مورد استفاده قرار گرفته است. در ادامه برای کارهای بعدی می‌توان انواع روش‌های کنترلی مانند کنترل لغزشی،کنترل‌کننده فیدبک سرعت، کنترل‌کننده فیدبک به صورت توابع غیرخطی، روش‌های کنترل تطبیقی و کنترل مقاوم و انواع سنکرون‌سازی-ها را بر روی این سیستم پیاده‌سازی کرد.

مراجع

[1] G. Qi, Z. Chen and Z. Yuan, "Model-free control of affine chaotic systems," *Physics Letters A* , vol. 344, pp. 189-202, 2005.
[2] B. R. Nana and P. Woafo, "Active control with delay of horseshoes chaos using piezoelectric absorber on a buckled beam under parametric excitation," *Chaos, Solitons and Fractals*, vol. 32, pp. 73-79, 2007.
[3] JH. Park, "Chaos synchronization between two different chaotic dynamical systems," *Chaos, Solitons and Fractals,* vol. 27, pp. 549-554, 2006.
[4] C. Wu, T. Fang and H. Rony, "Chaos synchronization of two stochastic Duffing oscillators by feedback control,"

[3] Benchmark

[5] G. Perez and HA. Cerdeira, "Extracting messages masked by chaos," *Physical Review Letters*, vol. 74, pp.1970-1973, 1995.
[6] L. Pecora, " Hyperchaos harnessed," *Physics World*, vol.9, pp. 17-18, 1996.
[7] A. Genys, A. Tamasevicius and A. Bazailiauskas, "Hyperchaos in coupled colpitts oscillators," *Chaos, Solitons and Fractals,* vol. 17 , pp. 349-353, 2003.
[8] S. Cincotti and SD. Stefano, "Complex dynamical behaviors in two non-linearly coupled chua's circuits," *Chaos, Solitons and Fractals,* vol. 21, pp. 633-641, 2004.
[9] JP. Goedgebuer, L. larger and H. Porle, "Optical cryptosystem based on synchronization of hyperchaos generated by a delayed feedback tunable laser diode," *Physical Review Letters*, vol. 80, pp. 2249-2252, 1998.
[10] C. Li, X. Liao and K. Wang, "Lag synchronization of hyperchaos with application to secure communication," *Chaos, Solitons and Fractals,* vol.23, pp. 183-193, 2005.

[11] O. E. Rossler, "An equation for hyperchaos," *Physics Letters A* , vol. 71, pp. 155-157, 1979.
[12] D. Cafagna and G. Grassi, "New 3D scroll attractors in hyperchaotic chua's circuit forming a ring," *International Journal of Bifurcation and Chaos,* vol.13, pp. 2889-2903, 2003.
[13] Y. Li, WKS. Tang and G. Chen, "Hyperchaos evolved from the generalized lorenz equation," *International Journal of Circuit Theory and Applications,* vol. 33, pp. 235-251, 2005.
[14] Y. Li, WKS. Tang and G. Chen, "Generating hyperchaos via state feedback control," *International Journal of Bifurcation and Chaos,* vol. 15, pp. 3367-3375, 2005.
[15] T. Gao, Z. Chen, Z. Yuan and G. Chen, "A hyperchaos generated from chen's system," , *International Journal of Bifurcation and Chaos*, vol.17, pp.471-478, 2006.
[16] T. Gao, G. Chen, Z. Chen and S. Chen, "The generation and circuit implementation of a new hyperchaos upon Lorenz system," *Physics Letters A*, vol.361, pp. 78-86, 2007.
[17] Z. Chen, Y. Yuang, G. Qi and Z. Yuan, "Anovel hyperchaos system only with one equilibrium," *Physics Letters A*, vol.36 , pp. 696-701, 2007.
[18] G. Qi, G. Chen, S. Du, Z. Chen and Z. Yuan,"Analysis of a new chaotic system," *Physica A : Statistical Mechanics and its Applications,* vol.352 , pp.295-308, 2005.
[19] A. Chen, J. Lu and S. Yu, "Generating hyperchaotic Lu attractor via state feedback control," *Physica A: Statistical Mechanics and its Applications*, vol. 346, pp. 103-110, 2006.
[20] C. Wang, X. Zhang, Y. Zheng and Y. Li, "A new modifid hyperchoatic Lu system," *Physica A: Statistical Mechanics and its Applications*, vol. 371, pp. 260-272, 2006.
[21] W. Wu, Z. Chen and Z. Yuan, "The evolution of a novel four-dimensional autonomous system: Among 3-torus, limit cycle , 2-torus, chaos and hyperchaos," *Chaos, Solitons and Fractals,* Accepted 2 July 2007.
[22] Y. Li, G. Chen and WKS. Tang, "Controlling a unified chaotic system to hyperchaotic," *IEEE Transaction Circuit System II* , vol. 52, pp. 204-207, 2005.
[23] Q. Li, X. S. Yang and F. Yang, "Hyperchaos in Hopfield-type neural networks," *Neurocomput-ing, vol.* 67, pp. 275-280, 2005.
[24] Z. M. Ge and G. H. Yang, "Hyperchaos of four state autonomous system with three positive lyapunov exponents," *Physics Letters A*, vol.373 , pp. 349-353, 2009.
[25] Q. Jia, " Hyperchaos generated from the Lorenz chaotic system and its control," *Physics Letters A*, vol.366 , pp 217-222, 2007.
[26] J. Wang, Z. Chen and Z. Yuan, " The generation of a hyperchaotic system based on a three dimensional autonomous chaotic system, " *Chynese Physics* , vol. 15, pp. 1216-1225, 2006.
[27] C. Liu, T. Liu, L. Liu, K. Liu, " A novel chaotic attractor," *Chaos, Solitons and Fractals,* vol.22, pp. 1031-10383, 2004.
[28] A. Wolf, J. B. Swift, H. L. Swinney, J. A. Vastano, " Determining Lyapunov exponents from a time series," *Physica D*, vol. D16, pp. 285-317, 1985.

Journal of Control Vol. 3, No. 1, pp. 1-11, Spring 2009

A Combined DC–Filter and Optimized Modulation to Absorb DC–Link Oscillations of Cascaded H-Bridge Converters

M. Tavakoli Bina[1], B. Eskandari[2]

[1] Associate Professor, Faculty of Electrical Engineering, K. N. Toosi University of Technology, tavakoli@kntu.ac.ir

[2] M.Sc. Student, Electrical Engineering, Faculty of Electrical Engineering, K. N. Toosi University of Technology, b_eskandari@kntu.ac.ir

Abstract: Cascaded H–bridge converters may well be used for applications requiring higher voltages, without needing series connection of semiconductor switches. Two problems are necessary to be addressed; first, the AC voltages of the sub-modules of the cascaded converter which are unequal due to the deployed modulation technique, thus introducing different fundamental components and harmonic contents. Second, is the DC-links of sub-modules that are subjected to low-frequency oscillations when operating under three-phase unbalance condition and/or single-phase active power exchange. This paper begins with the first problem, suggesting a generalized swapping technique for a cascaded H-bridge converter. This is applied to the SPWM and the optimal PWM, while the optimal PWM is suggested to be subjected to a complementary constraint. A five-level H-bridge cascaded converter is developed to implement the suggested modulations. Practical results confirm that AC voltages of the sub-modules are equalized. Furthermore, three various external DC active filter circuits are presented to tackle the second problem, including the independent DC source, the auxiliary S-bridge and the buck-boost design. These circuits are simulated, and their performances are compared. Moreover, the buck-boost design is implemented, and applied to the DC output of a single-phase full-bridge rectifier. Then, three control strategies are further tested on the buck-boost compensating circuit. Experimental results show that the effective value of the DC ripples is considerably lowered down in comparison with the original DC oscillation.

Keywords: Cascaded converter, DC-link oscillation, balanced AC voltages, optimal-PWM, SPWM.

چکیده: مبدلهای چند سطحی کاسکاد در کاربردهایی که ولتاژ بالا نیاز است مورد توجه قرار گرفته اند. در حالیکه این امر از سری کردن سوییچهای نیمه هادی جلوگیری می نماید، دو مساله مهم نیزدر اینگونه مبدلها باید مورد مطالعه واقع گردند. اول، مساله توزیع متعادل ولتاژ AC در خروجی هریک از مبدلهای سری شده که میتوانند در اثر مدولاسیون بکارگرفته شده بوجود آیند. دوم، مساله تعادل ولتاژهای لینک DC که در اثر عدم تعادل لینک AC و یا تبادل توان اکتیو حاصل میگردد. این مقاله با مساله اول شروع کرده و روش جابجایی عمومی برای مبدلهای کاسکاد نوع H ارایه مینماید. ضمن در نظر گرفتن مدولاسیون SPWM و PWM بهینه، یک میدل پنج سطحی ساخته شده و روش ارایه شده توسط مبکزوکنترولر برنامه ریزی گردیده است. نتایج عملی بدست آمده یکسان سازی خروجی AC مبدلهای نوع H را تایید مینماید. بعلاوه، سه روش اکتیو جهت فیلتر کردن توسانات سمتDC ارایه شده است. همچنین، یک مبدل باک – بوست نیز بهمراه سه روش کنترلی پیشنهاد و پیاده سازی شده است که نتایج عملی موید کاهش چشمگیر ریپل سمت DC نسبت به نوسانات اولیه آن می یاشد.

کلمات کلیدی: مبدل کاسکاد، نوسانات سمت DC ، متعادلسازی خروجی هایAC، PWM بهینه، SPWM

Nomenclatures:

N : total number of H-bridges

K : number of switching instants

α_{ij} : the jth switching instant of the ith H-bridge converter

V_{a-n} : the nth harmonic component

d : practical duration between two consecutive switching instants

V_{dc} : the DC voltage of the compensating H-bridge converter

t_{on} : the on-time duration of each switch

V_f : oscillations on top of the DC-link

C_{f1} and C_{f2} : DC-link active filter capacitances

C : DC-link capacitance

ΔV : hysteresis-like positive band

L : AC-side commutation inductance

V_L : the voltage drop on the inductance L

i_L : current through the inductance L

1- Introduction

Multilevel converters can potentially overcome the practical impediments associated with the series connection of semiconductor switches to increase the system voltage [1]-[4]. Figure 1(a) shows a five-level cascaded H-bridge converter with which the harmonic performance is expected to be improved, while the supply voltage is distributed across the sub-modules. However, two problems should be considered. First, the applied voltage can be distributed unevenly across the H-bridge converters depending on the applied modulation technique. Second, the AC voltages of the sub-module may also add different harmonic contents, both in magnitude and phase. Since the series current is identical for all converters, sub-modules could be subjected to different power levels. Thus, it is also expected to observe unbalance of the DC-link voltages as well as oscillation on each DC-link voltage.

Second, each H-bridge sub-module exchanges *instantaneous* active power with the electrical network and thereby leading to DC-link voltage oscillations. (Note that the *average* absorbed active power by the converters in a power system period could be still as low as the power losses of the converters.) This exchanged power will highly affect the DC–link voltage of an H-bridge converter, causing low frequency oscillations (e.g. 100/120 Hz oscillations for synchronous frequency 50/60 Hz like that shown in Fig. 1(c)). The bigger the exchanged power within the sub-module, the larger will be the observed oscillations on top of the DC-link voltage [5]-[6].

Conventional PWM techniques have been introduced in related literatures. These techniques are simple to implement, while the two raised problems will, more or less, persist in the H-bridge sub-modules. Since harmonic magnitudes of the H-bridge converters may vary, different oscillating components with different magnitudes can be added to the DC-link voltages. Some other techniques objectively eliminate certain harmonics. E.g., an optimized pulse width modulation technique (OPWM) is presented in [7] wherein the modulation is capable of eliminating any given number of harmonics. The problem, however, is that some optimized switch-on or switch-off durations might be very short. In practice, the switching frequency of power semiconductor switches is limited. Hence, implementation of the resulting optimized modulation cannot be fully achieved.

This paper begins with sinusoidal PWM (SPWM) and the OPWM techniques for cascaded H-bridge converters in order to address the first problem of uneven AC voltage distribution. One suggestion is introduced to improve the conventional SPWM for five-level cascaded H-bridge converters. This will make AC voltages of the two sub-modules similar, while the DC-link oscillations are still there. A five-level cascaded H-bridge converter is implemented to verify above suggestion. Furthermore, the constraints of the OPWM are so improved that the switching durations can actually be programmed. Then, the improved OPWM is also verified by the five-level converter, and practical results are compared with those of the improved SPWM. To deal with the DC-link oscillations, three different DC-filters are proposed to be connected across the DC–link capacitor, including the independent DC source, auxiliary S-bridge converter and the buck-boost compensating converter. The suggested methods are then examined and simulated by MATLAB to assess their performances and their suitability for the cascaded H-bridge converters. Among these converters, the buck-boost design is implemented, and applied to a single-phase full-bridge rectifier in which the DC oscillations are strongly imposed by the DC source. Practical results confirm that this latter proposal introduces marked improvement in lowering effective values of the DC–link oscillations.

2- Generalized switching pulse transition

Consider the cascaded five-level converter in Fig. 1(a) which is modulated through a conventional SPWM. MATLAB simulations are so arranged that four triangular carriers of 2100 Hz (twice as the number of H-bridges) are compared with a sinusoidal reference (50 Hz). Resultant comparisons lead to the output waveforms shown by Fig. 1(b) including two AC voltages of the H-bridges together with the total

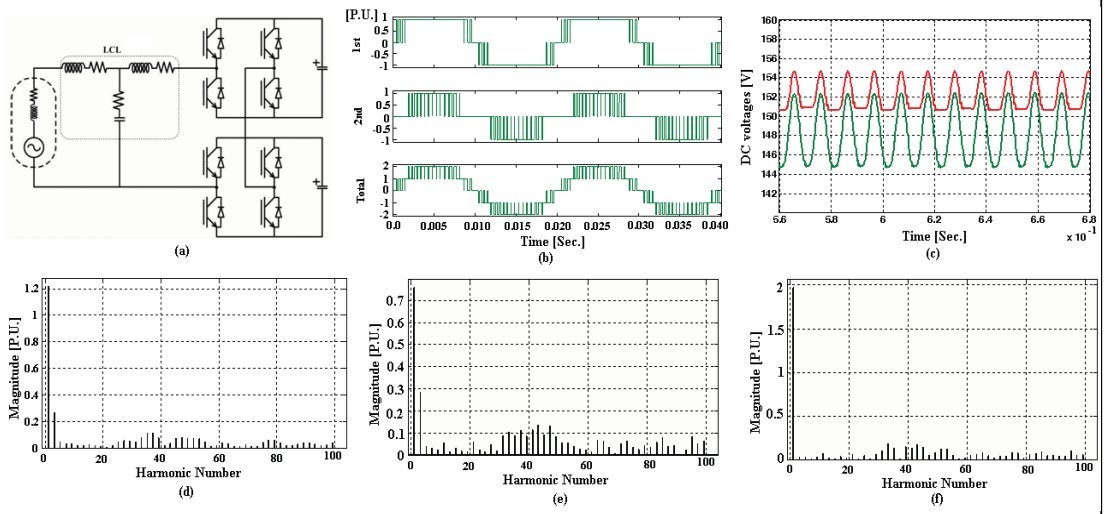


Figure 1: (a) Cascaded five-level H-bridge converter, (b) AC voltages of the two H-bridges along with the total AC voltages resulting from the SPWM switching technique, (c)unbalance and low-frequency oscillations of the DC-link voltages, and (d)-(f) Fourier analysis of the three AC voltages introduced by (c).

cascaded voltage. Fourier analysis of these waveforms is introduced in Figs. 1(d)-(f). The analysis shows that the first H-bridge introduces a fundamental component of 1.215 P.U., the second H-bridge 0.755 P.U., and the total cascaded voltage 1.972 P.U. This clearly implies that the first H-bridge has a bigger modulation index, and higher duty ratios in comparison with the second converter. Since the series current is identical for both converters, the DC-link voltage of the first H-bridge is different from the second one. Hence, the first H-bridge converter introduces larger oscillations compared to the second one and this is due to their different power levels. Simulations shown by Figs. 1 confirm the presence of the two stated problems (uneven AC voltage distribution and DC-link oscillation).

2.1 *Switching pulse transposition (SPT)*

The conventional SPWM not only produces unequal fundamentals, but it also produces different harmonic levels for the two H-bridges. One suggestion for symmetrical operation of the sub-modules could be a *regular transposition* of the switching pulses that are generated by the conventional SPWM. In practice, the SPT should be performed within every synchronous period such that a DC component *will not* appear on the AC voltage of a sub-module. For instance, when the switching pulses of the two sub-modules (generated by the conventional SPWM) are swapped over every half of a synchronous period, simulations show nonzero DC values on AC voltages of both H-bridge converters.

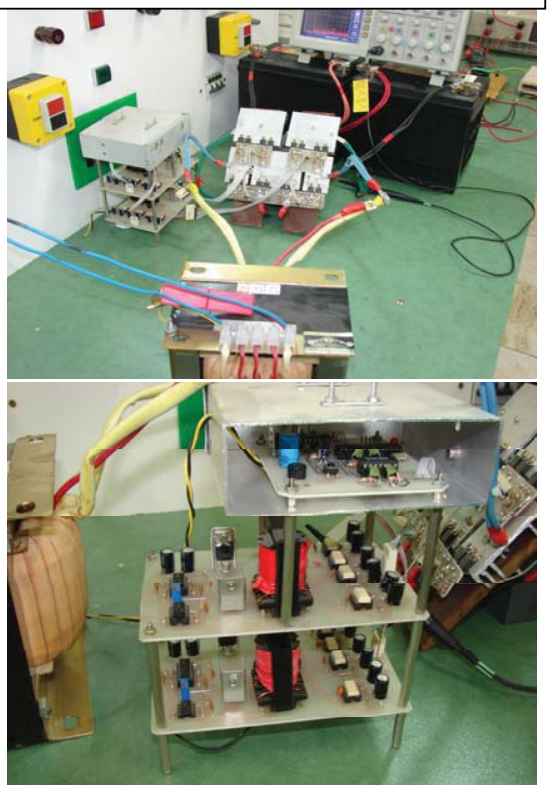


Figure 2: Experimental arrangement for balancing the capacitor voltages using the improved SPWM and the OPWM techniques.

As the number of H-bridges increases, application of the SPT among the sub-modules needs to be regulated.

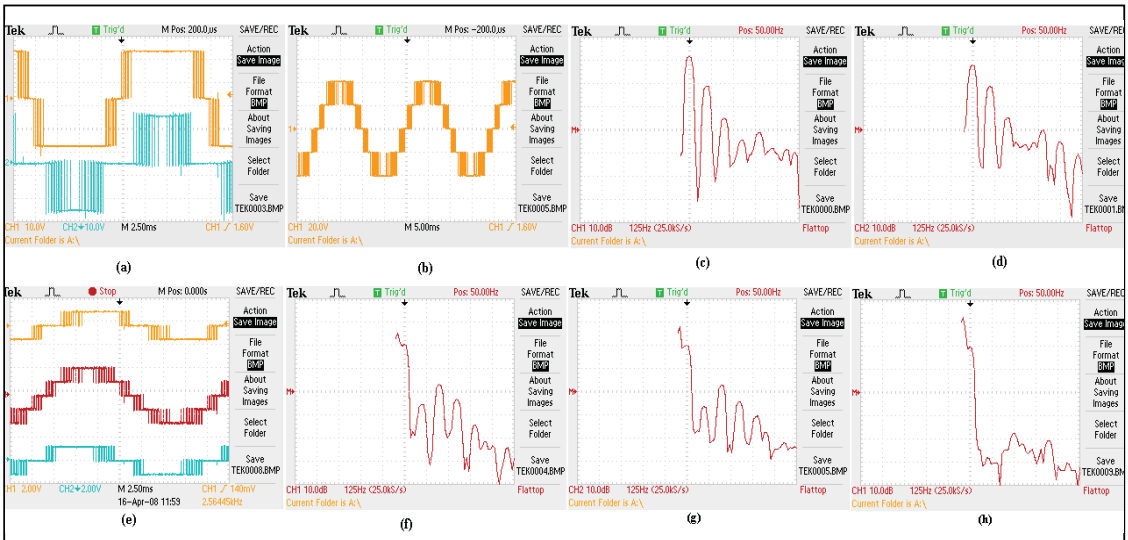


Figure 3: (a)-(d) Application of the conventional SPWM by the AVR: (a) AC voltages of the two sub-modules, (b) total AC voltage of the cascaded converter, (c)-(d) Fourier analysis of the waveforms introduced by (a), and (e)-(h) when the improved SPWM is applied by the microcontroller: (e) AC voltages of the sub-modules along with the total AC voltage of the cascaded converter, (f)-(h) Fourier analysis of the waveforms introduced by (e).

To propose a rule for objective SPT technique, huge number of simulations has been arranged starting from two H-bridges up to six sub-modules. Table I lists resultant SPT angles that improves both fundamentals and harmonic contents of all sub-modules nearly equally. According to these recorded results, *the number of swapping* of the switching pulses within *every half-cycle* equals to *six times of that of the number of the sub-modules* (e.g. 12 SPT for two sub-modules and 30 SPT for five sub-modules). The only exception is that of the two sub-module in which *the swapping* can also take place *every quarter of a synchronous period* in addition to the suggested general rule (transposition of the SPWM outcome every 15°). Simulations confirm that the harmonic behavior of both sub-modules are quite similar, in particular, both fundamentals and low order harmonics become identical.

Figure 4 illustrates how 30 SPT occur during a half-cycle among five H-modules (H_1~H_5). Every SPWM outcome is applied first to H_1 during [0°, 6°], the second is transmitted to H_2 during [6°, 12°], third to H_3 during [12°, 18°], fourth to H_4 during [18°, 24°] and fifth to H_5 during [24°, 30°]. Then, this procedure is *repeated six times* until the half cycle is completed. It is noticeable that various repeating procedures may be deployed for the six SPT amongst the sub-modules. Nevertheless, the best solution is achieved when the repeating procedure always starts from the first sub-module down to the last one as is shown in Fig. 4.

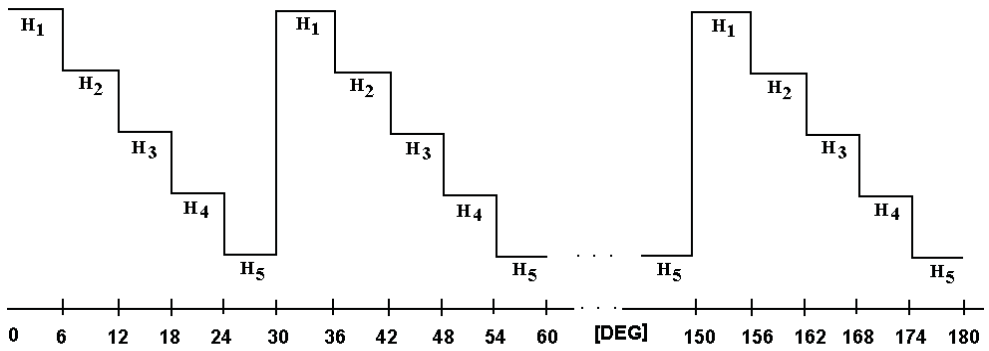


Figure 4: Illustration of the SPT suggested procedure for a cascaded H-bridge converter that includes five sub-modules.

2.2 *Experimental validation*

A five-level cascaded H-bridge converter, shown by Fig. 1(a), is implemented to examine the equalization of the AC voltages of the sub-modules. Figures 2(a)-(b) show the power circuit together with

the microcontroller and switching driver circuits. Sixteen switches (MOSFET 2N06L05) are used in all, eight for each sub-module, where every two parallel switches make one pack. The IC number TLP250 is used for switching drive circuits that transmit the switching pulse train produced by the AVR microcontroller. The microcontroller consists of two ATMEGA8-16PU units. A 15 Amp fast fuse is used to protect the device. Two 12 V batteries supply the DC-link voltage, and a 1200 VA transformer is optionally available to be connected either to 220 V loads or to a possible grid connection (synchronization techniques have to be applied).

TABLE I: THE SUGGESTED PST FOR THE SPWM IN ORDER TO MAKE FUNDAMENTALS AND HARMONIC CONTENTS NEARLY THE SAME.

Number of cascaded H-bridges	Duration of one transposition in Degrees	Number of repeating of transpositions during each half-cycle	Total number of transpositions each half cycle
2	15°	6	12
3	10°	6	18
4	7.5°	6	24
5	6°	6	30
6	5°	6	36

Both the conventional SPWM and the suggested complementary improvement to the conventional SPWM have been programmed using the AVR microcontroller, and applied to the cascaded five-level converter. Four triangular carrier frequencies are all 2100 Hz. Experimental results are shown in Figs. 3(a)-(d) for the conventional SPWM, and Figs. 3(e)-(h) for the improved SPWM. Figures 3(a) and 3(e) demonstrate the AC voltages of the two sub-modules, and Figs. 3(b) and 3(f) show the total AC voltage of the five-level cascaded converter. To find out the effect of switching pulse swapping in every quarter, Fourier analysis of the two modulating programs are compared. Figures 3(c)-(d) zoom in the fundamentals as well as the low order harmonics of the two H-bridge AC voltages for the conventional SPWM. While the difference between the fundamentals is about 4dB (their ratio is about 1.58 like the simulations of Fig. 1), the harmonics also show different magnitudes. Considering the improved SPWM, Figs. 3(g)-(h) verify that both fundamentals are nearly identical, and the harmonic behavior of sub-modules are quite similar. It should be noted that low-frequency oscillations of the DC-links still remain as an issue for both conventional SPWM and the improved one. This is even worse in the case of the improved SPWM compared to the conventional SPWM simulated by Fig. 1(c).

3- Simulation Results

Conventional multilevel SPWM schemes produce uneven fundamental voltages for the H-bridge sub-modules [14]-[18]. As an alternate technique, optimized algorithms are introduced in [6]-[10], to eliminate certain harmonic components by seeking the best switching instants. Assume N H-bridge converters are cascaded wherein every sub-module's AC voltage has both half-wave and quarter-wave symmetry. Then, determination of switching instants during each quarter-period is enough to recognize the whole period, and the Fourier series include only odd sinusoidal terms ($\sin(n\omega t)$). Hence, general description of the nth harmonic voltage using the Fourier series for N cascaded H-bridges and K switching instants within each quarter-wave is calculated as below:

$$V_{a-n} = \frac{4}{n\pi}\sum_{i=1}^{N}\sum_{j=1}^{K}(-1)^{k+1}\cos(n\alpha_{ij}) \quad (1)$$

Where the angle α_{ij} is the jth switching instant of the ith H-bridge converter, and V_{a-n} is the nth harmonic component of the total cascaded converter voltage. It is shown in [7] that harmonics up to the order ($2NK$-1) should be eliminated for N sub-modules and K switching instants. An optimization problem can now be arranged having an objective function that minimizes the remaining odd harmonics starting from the third up to ($2NK$-1). This is also subjected to several constraints. One necessary condition has to be satisfied by all sub-modules AC voltages on having identical fundamental magnitudes equal to $\frac{|V_{a-1}|}{N}$. This constraint introduces N equations, each having K unknown switching instants. Also, obvious conditions have to be set on K ascending angles α_{ij} for each sub-module within the period of $\left[0, \frac{\pi}{2}\right]$. In practice, a certain value should be considered between every two consecutive switching instants (d) to guarantee proper switching transitions. Thus, the third group of limitations is proposed here (added in (2)) in which the *pulse widths are bigger than a certain value d* (e.g. 10μs) for implementation purposes. Choosing this value depends on the switch

technical specifications, including the turn-on and turn-off durations that affect the switching frequency. The combinatorial optimization problem can be formed as below:

$$\text{Minimize} \sum_{\substack{n=3 \\ (n \text{ is odd})}}^{2NK-1} V_{a-n}^{2}$$

$$\text{Subject to:} \begin{cases} \dfrac{4}{\pi}\sum_{j=1}^{K}(-1)^{k+1}\cos(\alpha_{ij}) = \dfrac{|V_{a-1}|}{N} \\ 0 < a_{i1} < a_{i2} < \cdots < a_{iK} < \dfrac{\pi}{2} \\ \begin{cases} a_{i2} - a_{i1} \geq d \\ a_{i3} - a_{i2} \geq d \\ \vdots \\ a_{iK} - a_{i(K-1)} \geq d \end{cases} \end{cases} \quad \text{for } i = 1,2,\ldots,N \tag{2}$$

3.1 *Implementation of the complementary optimal modulation*

The complementary optimal modulation of (2) can be solved using genetic algorithm. Since different initial conditions lead to different local minimums, ninety different initial conditions within [0°, 90°] are applied to the optimization problem. Solutions are sorted according to their resultant THD%. The solution with the lowest THD% is selected and stored in a lookup table by the microcontroller. This process is repeated for various fundamental magnitudes, which eventually provides all the needed lookup tables for the microcontroller.

Table II lists a typical lookup table containing the obtained optimal switching instants for N=2 and K=21 according to the stated procedure when d=10μs and $|V_{a-1}| = 36$ dB (about 63 V). Switching angles, related to other three quarters of the waveform, can simply be generated according to the assumed quarter-wave and half-wave symmetry for the waveform.

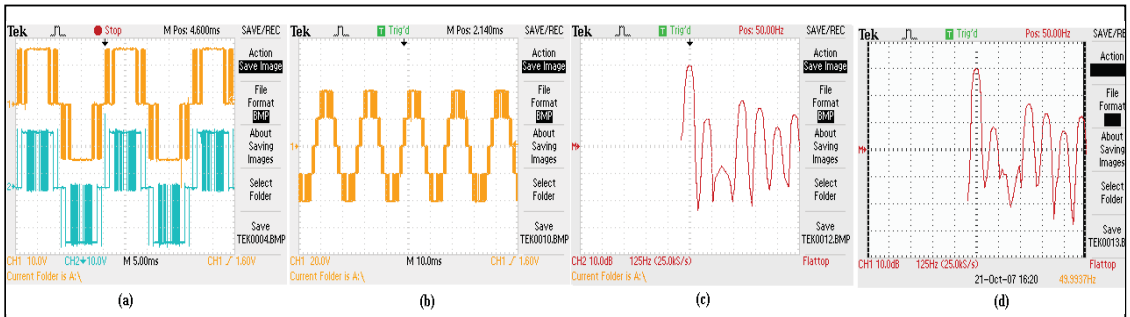


Figure 5: Application of the complementary optimal switching modulation by the AVR: (a) AC voltages of the two sub-modules, (b) total AC voltage of the cascaded converter, (c)-(d) Fourier analysis of the waveforms introduced by (a).

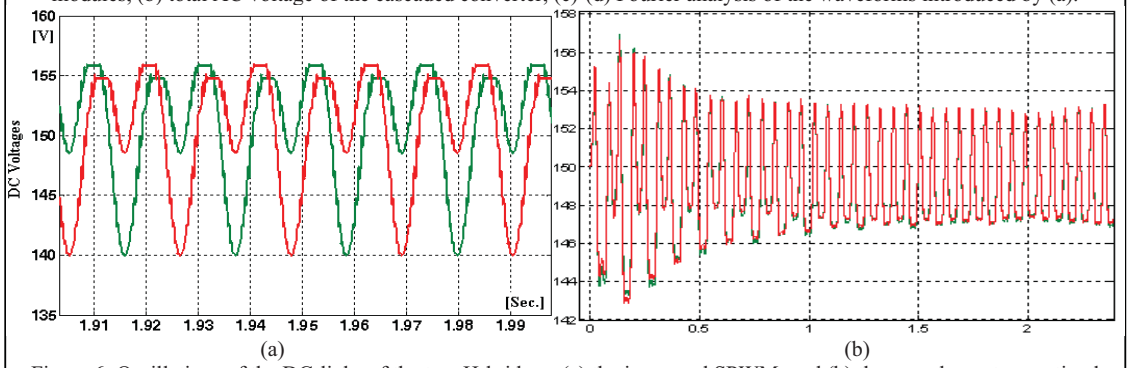


Figure 6: Oscillations of the DC-links of the two H-bridges, (a) the improved SPWM, and (b) the complementary optimal switching modulation.

TABLE II: THE COMPLEMENTARY OPTIMAL SOLUTION LISTING THE SWITCHING INSTANTS OF THE FIRST QUARTER OF THE WAVEFORM THAT IS STORED AS A LOOKUP TABLE.

Sub-module Number	Optimized switching instants in the first quarter-period (angles in degrees)									
1	12	13	15	17	25	27	28	30	31	32
Cont. 1	36	41	43	45	46	48	50	58	59	64
2	8	9	19	21	22	25	28	29	31	36
Cont. 2	40	45	46	52	53	55	56	61	62	76

The microcontroller is then uploaded with all the obtained optimal solutions, and the modulation program is applied to the five-level cascaded H-bridge converter. Figure 4 shows the implementation outcomes. Both AC voltages of the two H-bridges are shown in Fig. 4(a), and the total cascaded voltage in Fig. 4(b). Comparing the fundamentals of the two sub-modules, it can be seen that they are slightly

different (29.8 dB and 30.14 dB). Hence, the switching pulses are *swapped* every *quarter-period*, and applied to the sub-modules. Figures 4(c) and 4(d) illustrate the Fourier analysis of the two AC voltages of the sub-modules, showing identical fundamentals for them with similar low-order harmonic behaviours.

Practical results indicate that both the complementary optimal switching modulation and the improved SPWM are capable of distributing identical fundamental AC voltages on the two sub-modules. However, the complementary optimal modulation with quarter-period swapping provides *better* AC voltage distribution and harmonic behaviour compared to the improved SPWM. But, both methods are still unable to attenuate the DC-links oscillation sufficiently as shown by simulations in Figs. 5(a)-(b). While the steady state peak-to-peak oscillation magnitude is below 16 V for the improved SPWM in Fig. 5(a), it is less than 7 V for the complementary optimal modulation in Fig. 5(b).

4- Suppression of DC-link oscillations

Connection of various DC passive/active filtering circuits across the DC capacitor is discussed in [11]. Predominant frequency of these oscillations is related to the second harmonic when the power system operates at 50/60 Hz. Other higher harmonic orders can also be present if the applied voltage includes any components in addition to the synchronous frequency. Typical uncompensated DC-link oscillations are shown in Figs. 5(a)-(b) when the improved SPWM and the OPWM modulation techniques are used, respectively. A PI controller is used to control the phase angle of the H-bridges output voltages such that the average DC-voltage remains fixed at 150 V.

One approach to filter out the low-frequency oscillations of the DC-links may be a passive LC-filter tuned at 100/120 Hz. This would effectively reduce oscillations. Nevertheless, the disadvantages of the passive LC-filters are their high cost, big parameters and size. Hence, the following sub-sections suggest and examine new designs that make use of active filters to damp the oscillations effectively.

4.1 *Independent DC source (IDC): Proposition 1*

The proposition of Fig. 6(a) uses the idea of passive LC-filter to absorb the oscillations in which the inductance of the passive filter is replaced by the primary winding of a transformer. The secondary of the transformer is connected to a low-power low-voltage H-bridge converter through a passive low-pass LCL-filter. The suggested design provides much more satisfactory attenuation of DC-link oscillations compared to the passive LC-filter. Thus, the control and operation of this proposal will be examined in detail.

First, the oscillations of DC-link voltage (V_f) is extracted by subtracting the average value from the exact value of the DC-link voltage. Using the zero-order hold (ZOH) function of SIMULINK, the sampled signals (rated at 10 kHz) are converted to continuous signals. Then, the volt-second balance law is applied to each switching period (100 μs) to find the duty ratio of each switch as follows:

$$V_{dc} \times t_{on} = V_f \times 100(\mu s) \tag{3}$$

Where V_{dc} is the DC voltage of the compensating H-bridge converter, and t_{on} is the on-time duration of each switch. Computing t_{on} of the switches, they are applied to the compensating H-bridge converter. The whole design is simulated with MATLAB, where the transformer has a turn ratio slightly bigger than one. Figure 6(c) shows the compensated DC-link oscillations, which is largely attenuated down to about 2.5 V peak-to-peaks (or 1.67%).

It is noticeable that the compensating elements including the DC source, transformer, low-pass filter and switches operate under low-voltage low-power conditions. But in practice, according to (3) the lower the magnitudes of oscillations (V_f), the lower the duty ratios of switches will be. This implies a limit on reduction of magnitudes of oscillations as a drawback of the method. To remedy this, the voltage V_{dc} can be decreased when the oscillations are needed to be attenuated significantly. This also will add extra-cost to the converter for developing regulated controllable V_{dc}.

4.2 *Auxiliary S-bridge compensation (ASC): proposition 2*

This proposal uses two identical capacitors (C_{f1} and C_{f2}) along with three switches for each DC-link like that which it is illustrated in Fig. 6(b). Capacitances C_{f1} and C_{f2} have identical average DC voltages equal to three quarters of the DC-link voltage of capacitor C. Three switches are operated in a way that when the DC-link voltage is increased beyond a positive band (ΔV), the two vertical switches are turned on and the horizontal switch turns off. This puts both capacitances C_{f1} and C_{f2} in parallel, absorbing charging currents through the inductance L from C by the following slope:

$$\frac{di_L}{dt} = \frac{V_L}{L} = \frac{V_{DC} - \frac{3}{4}V_{DC}}{L} = \frac{V_{DC}}{4L} \qquad (4)$$

Where V_L is the voltage drop on inductance L and i_L is the absorbed current. This relatively big slope forces the main DC-link voltage to come down more rapidly. Similarly, when the DC-link voltage is dropped below a negative band ($-\triangle V$), the two vertical switches are turned off and the horizontal switch turns on. The two capacitors C_{f1} and C_{f2} operate in series, injecting current to the DC-link capacitor C through the inductance L as follows:

$$\frac{di_L}{dt} = \frac{V_L}{L} = \frac{V_{DC} - \frac{6}{4}V_{DC}}{L} = -\frac{V_{DC}}{2L} \qquad (5)$$

Here the negative slope reverses the current from C_{f1} and C_{f2} to charge the DC-link capacitor C. Simulations show that peak-to-peak magnitude of the oscillations is smaller than 1 V (or 0.6%) for a very high switching frequency case as shown in Fig. 6(d). In practice, the hysteresis-band could limit the performance of the S-bridge considerably such that when the selected ΔV is small, the switching frequency is very high.

4.3 An equivalent buck-boost circuit: proposition 3

The two foregoing propositions have the advantage of lowering oscillations, while being forced to operate in high switching frequencies. In practise, this will highly depend on how fast and how advanced the semiconductor technology is. Also, other drawbacks are the design costs, large size and occupying space. Hence, another suggestion is raised in which a buck-boost circuit is responsible for transferring energy between the DC-link capacitor and a DC battery. Figure 7(a) demonstrates this proposal that mimics the oscillations of the DC-link using a rectifier. A full-wave diode rectifier supplies a DC voltage (V_{DC}) containing an average voltage (V_{mean}) plus a dominant 100 Hz oscillation ($\triangle V$) to a load. A buck-boost circuit is used to compensate these emulated low-frequency oscillations.

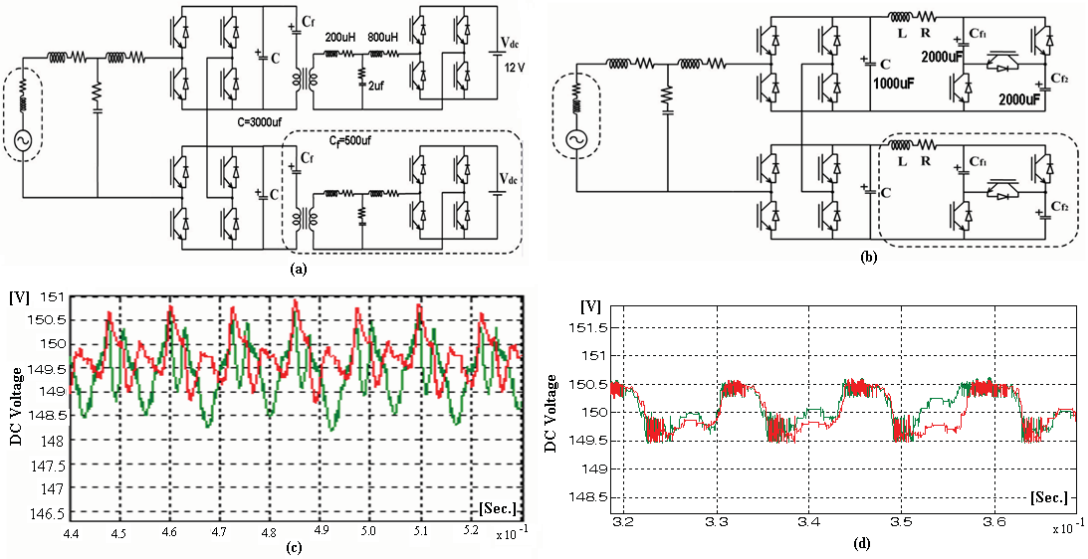


Figure 7: (a)-(b) Circuit diagrams of the IDC and the ASC, respectively, and (c)-(d) simulations illustrating the effect of applying the IDC and the ASC to the DC-link oscillations.

Buck and boost operation modes take place when using two switches (one controllable switch and a diode allocated in each branch). Duty ratios of the switches are decided on the basis of pulse width modulation provided by Fig. 7(b), where V_{dc} is the battery voltage. When V_{DC} is dropping, the DC battery injects power to the DC-link (boost-mode); when V_{DC} is increasing, the DC battery absorbs power from the DC-link (buck-mode). This rule is emulated in Fig. 7(b), wherein a 50 kHz ramp is compared with V_{dc}/V_{mean} for the buck-mode, and 1-V_{dc}/V_{mean} for the boost-mode. These are actually obtained from the well-known theory of DC-DC converters [12]. Simulations are shown in Figs. 7(c)-(d) wherein the last picture zooms on a small region of Fig. 7(c). It can be seen that the effective value of uncompensated oscillations efficiency drops down drastically. Practically speaking, the buck-boost compensation method is simple to implement, and that the oscillations can be traced properly. In the mean time, the cost of this design is much less than

the other two suggestions that use the concept of passive LC-filter compensation.

frequency estimation techniques. Further, the source equivalent impedance is needed to be included in the

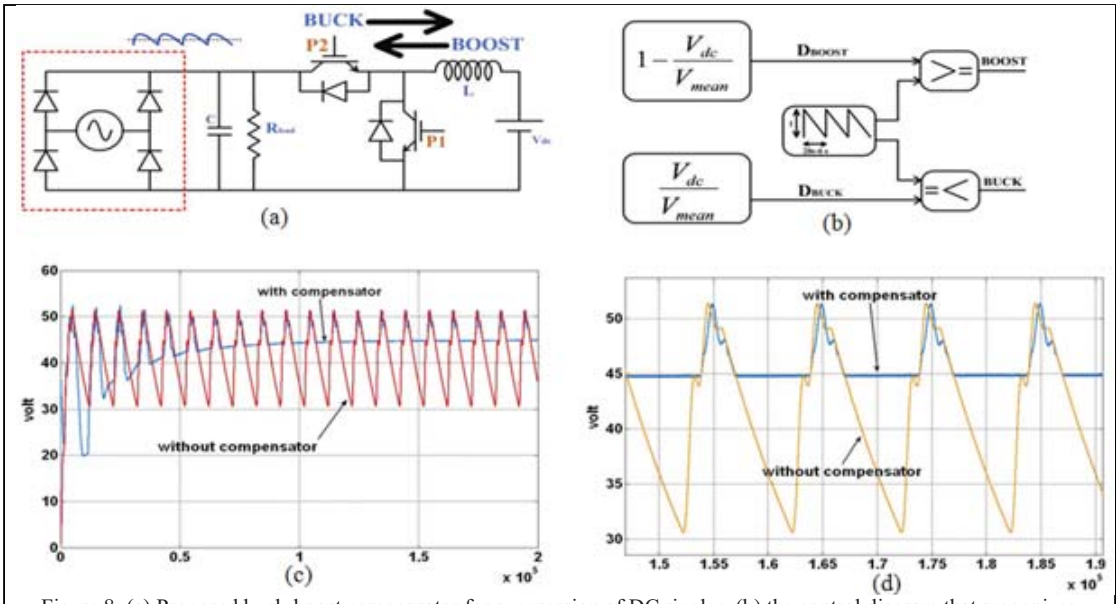


Figure 8: (a) Proposed buck-boost compensator for suppression of DC ripples, (b) the control diagram that recognizes buck and boost operating modes, (c)-(d) compensated oscillations by applying the buck-boost suggestion simulated by MATLAB.

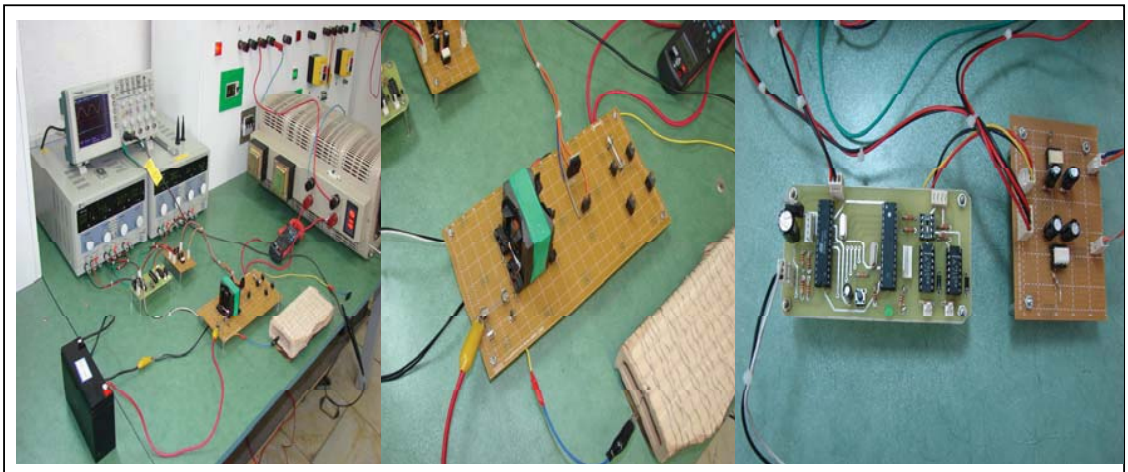


Figure 9: Experimental arrangement for suppression of low-frequency DC-link oscillations using the buck-boost proposition.

4.3.1 Implementation of the buck-boost proposal

The third proposition is implemented as the developed device is shown in Figs. 8(a)-(c). The power circuit along with control design of Fig. 7 is used to generate low-frequency oscillations by a full-bridge diode-rectifier. Supplied DC voltage by this full-wave rectifier basically provides an uneasy situation in which suppression of the produced oscillations need considerable attenuations compared to the DC-link oscillations of cascaded converter. It is noticeable that an exact cascaded converter can also be connected across the grid system, which will need additional control algorithms both for compensation approach and the synchronizing control algorithm for a reliable response. This would be an extensive practical field, which needs the well-known FACTS controller techniques for shunt grid-connected devices. Hence, this made us to think about producing oscillations with the same frequency as that of the cascaded converter by establishing the test circuit of a full-bridge rectifier. The rectifier is composed of four 30 A diodes that are connected to a 133 μF capacitor in parallel with an 18 Ω resistor. Two power MOSFETs (P50NE1) are used in the buck-boost circuit, which are driven by the IC driver TLP250.

Considering the compensating circuit in Fig. 7(a), the full-bridge rectifier generates a DC voltage that

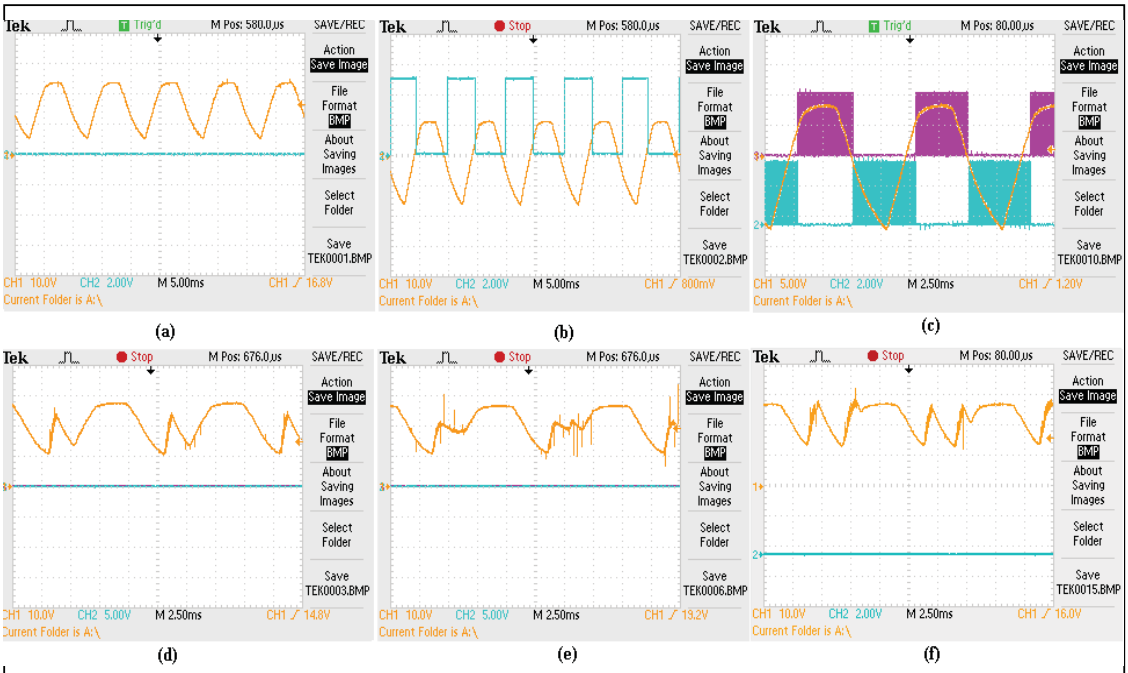


Figure 10: (a) Generated 100 Hz oscillations by the full-bridge rectifier, (b) recognition of buck/boost operation modes, (c) the switching patterns generated for both buck and boost modes, (d)-(f) suppression of oscillations using the FD, VC and WD, respectively.

contains significant 100 Hz ripples. The peak of these ripples is imposed by the rectifier such that compensation circuits have to suppress the oscillations with respect to the peak (not the average). Thus, this situation would be rather a difficult task compared to the exact cascaded converter application. The microcontroller is programmed in three controlled-cases for suppression of oscillations; fixed duty-cycle modulation (FD), voltage-controlled modulation (VC) and wider-band boost duty-cycle modulation (WD).

Experimental results are shown in Fig. 9. Figure 9(a) introduces the generated 100 Hz oscillations by the full-bridge rectifier. The DC value of this voltage is 15.00 V, while the RMS value of ripples is 20.90 V. Figure 9(b) illustrates the recognition of buck and boost operating modes according to the control rule shown in Fig. 7 (zero for the buck mode, and 5 V for the boost mode). The switching pulse train for each of the two operating modes are shown in Fig. 9(c). The upper switching signals relate to the activation of boost converter, and the lower ones to the activation of buck converter.

Considering the FD modulation for the buck-boost, Fig. 9(d) demonstrates the resultant compensation of DC oscillations (ΔV). It can be seen that the buck-boost effectively suppresses the ripples such that the RMS of the oscillations is 4.689 V. Comparing this RMS value with that of the initial rectifier DC ripples (20.90 V), buck-boost compensator along with the FD modulation reduces the RMS value of the oscillations down by 77.52%. Meanwhile, the DC value of the rectifier output is raised up to 20. 89 V. Application of the VC modulation to the buck-boost is recorded in Fig. 9(e) in which the duty ratio is regulated proportionally with respect to the mean value of the DC voltage. The RMS value of the oscillation is lowered down to 4.103 V, leading to the DC voltage rise up to 20.62 V. This shows slightly better performance than that of the FD modulation that reduces the effective value of the initial DC ripples down by 80.37%. Finally, the WD modulation is applied to the buck-boost compensator in which a wider-band is assigned to the boost converter. The RMS value of the oscillations and the DC value are 4.459 V and 21.94 V, respectively. Comparing with other two methods, the WD modulation introduces 78.67% reduction in the RMS magnitude of the oscillations, presenting the highest DC value.

5- Conclusion

Two problems are discussed in this paper in conjunction with cascaded H-bridge converters. The first issue deals with uneven distribution of the applied AC voltage on the H-bridge sub-modules due

to the engaged modulation technique. Second problem is related to a dominant 100/120 Hz oscillation on the DC-links because of active power exchange between AC and DC sides. Oscillations are then modulated by the converter, injecting undesirable uncharacteristic harmonics to the AC system. Furthermore, the DC voltages of the cascaded converters will *not* oscillate together, and this produces further sub-modules' DC voltage unbalance. To overcome the first problem, a switching pulse transposition procedure is suggested for conventional SPWM along with the optimal PWM modulation techniques. Suggestions are then verified using a practical five-level converter that is implemented by cascading two H-bridge converters, showing a significant improvement in AC voltage distribution. Further, to suppress the DC-link oscillations, three compensating circuits are proposed, which are the independent DC source, the auxiliary S-bridge and the buck-boost converter design. While all suggested circuits effectively reduce the oscillations, the low-cost buck-boost converter is the most inexpensive design. Hence, a buck-boost design is arranged and tested under certain oscillations. Duty ratios of the switches are regulated using three different methods. Experimental results confirm that the effective values of the DC-link oscillations are lowered significantly compared to the uncompensated case.

References

[1] J. S. Lai, F. Z. Peng "Multilevel Inverters: A survey of topologies, controls, and applications", IEEE Transactions on Industrial Electronics, vol. 49, August 2002, pp. 724-738.

[2] J. Rodriguez, J. S. Lai, F. Z. Peng "Multilevel Converters- A new breed of power converters", IEEE Transactions on Industry Applications, vol. 32, no. 3, May/June 1996, pp. 509-517.

[3] J. Rodriguez, J. S. Lai, F. Z. Peng, "Multilevel Converters- A new breed of power converters", *IEEE Transactions on Industry Applications*, vol. 32, No. 3, pp. 509–517, May/June 1996.

[4] J. Aziz, Z. Salam, S. I. Safie, "Analytical Approach to Obtain the Harmonics Spectra on a Five Level Cascaded Inverter Subjected to a New Modulation Scheme", IEEE PEDS 2003, Singapore.

[5] S. Ali Khajehoddin, *Student Member, IEEE*, Alireza Bakhshai, *Member, IEEE*, and Praveen K. Jain, *Fellow, IEEE ,"* A Simple Voltage Balancing Scheme for m-Level Diode-Clamped Multilevel Converters Based on a Generalized Current Flow Model*"* IEEE TRANSACTIONS ON POWER ELECTRONICS, VOL. 23, NO. 5, SEPTEMBER 2008.

[6] Dae-Wook Kang, *Member, IEEE*, Byoung-Kuk Lee, *Senior Member, IEEE*, Jae-Hyun Jeon, *Student Member, IEEE*,Tae-Jin Kim, *Student Member, IEEE*, and Dong-Seok Hyun, *Fellow, IEEE,"* A Symmetric Carrier Technique of CRPWM for Voltage Balance Method of Flying-Capacitor Multilevel Inverter*"* IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, VOL. 52, NO. 3, JUNE 2005.

[7] N. A. Azli and W. S. Ning, "Application of Fuzzy Logic in an Optimal PWM Based Control Scheme for a Multilevel Inverter", IEEE PEDS'03, November 2003, Singapore.

[8] J. Aziz, Z. Salam, S. I. Safie, "Analytical Approach to Obtain the Harmonics Spectra on a Five Level Cascaded Inverter Subjected to a New Modulation Scheme", IEEE PEDS 2003, Singapore.

[9] Zhong Du, Leon M. Tolbert, John N. Chiasson, "Active Harmonic Elimination for Multilevel Converters", IEEE TRANSACTIONS ON POWER ELECTRONICS, vol. 21, no. 2, March 2006.

[10] Mohamed S.A. Dahidah, Vassilios G. Agelidis, Machavaram V. Rao, "Hybrid genetic algorithm approach for selective harmonic control", Electric Power System Research ELSEVIER, in Press 2007.

[11] Toshihisa shimizu, "Dc Active Filter Concept and Single-phase p-q Theory for Single-phase Power Converters", IEEE PEDS2003, Tutorial A, November 2003, Singapore.

[12] Robert W. Erickson , Dragan Maksimovich "Fundamentals of Power Electronics Second edition"

[13] Bai Hua, Zhao Zhengming, Meng Shuo,Liu Jianzheng, Sun Xiaoying, "Comparison of Three PWM Strategies, SPWM,SVPWM & One-cycle Control", IEEE PEDS 2003 Singapore.

[14] Rajesh Gupta, Arindam Ghoshb, Avinash Joshi, "Control of cascaded transformer multilevel inverter based DSTATCOM", ELSEVIER- Electric Power System Research, October 2006.

[15] Feel-soon Kang, Su Eog Cho, Sung-Jun Park, Cheul-U Kim, Toshifumi Ise, "A new control scheme of a cascaded transformer type multilevel PWM inverter for a residential photovoltaic power conditioning system", ELSEVIER-2004.

[16]] Wang Yi, Li Heming, Shi Xinchun, Zhu, "Harmonic Analysis and Filter Design for Medium-Voltage Multilevel PWM Inverters", IEEE PEDS 2003, Singapore.

[17] D. Soto, T. C. Green, "A Comparison of High Power Converter Topologies for the implementation of FACTS Controllers", IEEE Transactions on Industrial Electronics, vol. 49, No. 5, pp. 1072–1080, October 2002.

[18] A. Lesnicar, and R. Marquardt "An Innovative Modular Multilevel Converter Topology Suitable for a Wide Power Range", ETG-Fachtagung, Bad Nauheim, Germany, 2003.

[19] R. Marquardt and A. Lesnicar "New Concept for High Voltage – Modular Multilevel Converter" IEEE PESC 2004.

Journal of Control Vol. 3, No. 1, pp. 12-22, Spring 2009

Utilizing Multi-Parametric Programming for Optimal Control of Linear Constrained Systems

Farhad Bayat[1], Ali A. Jalali[2], Mohamad R. Jahed Motlagh[3]

[1] Ph.D. Student, Control Engineering, fbayat@iust.ac.ir
[2] Associate Professor, ajalali@iust.ac.ir
[3] Associate Professor, jahedmr@iust.ac.ir

Department of Electrical Engineering, Iran University of Science and Technology, Tehran 16846, Iran,

Abstract: Multi-Parametric Programming (mpp) is a widely acknowledged mathematical method for its ability to handle parameter depended optimization problems explicitly. Using this powerful tool, the traditional model predictive control (MPC), which is highly time consuming, can be solved efficiently, so-called explicit MPC (EMPC). Its main advantage is that it obtains the control actions as explicit function of the plant measurements. This allows for the control actions to be obtained on-line via simple function evaluations instead of solving repetitively a computationally demanding on-line optimization. This significant characteristic allows the MPC strategy to be applicable for the fast dynamics, which may be prohibitive for traditional MPC that relies on optimization methods. This paper aims to study this method from theoretical and practical point of view. The complete unified framework for developing EMPC and its mathematical foundations are studied. Also, in this paper a simple closed form for some matrices well established to achieve low computational complexity in design procedure. Finally the problem of servo system control has been dealt with and thereby the performance and effectiveness of the presented method verified.

Keywords: Explicit Predictive Control, Optimal Control, Constrained Systems, Servo System.

چکیده: روش برنامه‌ریزی چندپارامتری بدلیل توانایی بالای آن در برخورد صریح با مسائل بهینه‌سازی پارامتری بسیار مورد توجه قرار گرفته است. به کمک این ابزار قدرتمند می‌توان کنترل پیش‌بین سنتی که حل آن بسیار زمانبر می‌باشد را بصورت کارآمدی حل نمود (کنترل پیش‌بین صریح). مهمترین مزیت این روش آن است که قانون کنترل استخراج شده بوسیلهٔ این متد بصورت تابعی صریح از حالتهای سیستم می‌باشد. این ویژگی باعث می‌شود که محاسبهٔ قانون کنترل در حالت برخط تنها با محاسبهٔ یک تابع ساده بجای حل یک دسته مسائل بهینه‌سازی زمانبر انجام پذیرد. این ویژگی برجسته موجب می‌شود که بتوان روش کنترل پیش‌بین صریح را برای سیستمهای دینامیکی سریع که روش سنتی برای آنها کارایی ندارد نیز بکار برد. در این مقاله ضمن بررسی روش کنترل پیش‌بین صریح از نقطه نظر تئوری یک ساختار یکپارچه برای پیاده‌سازی و توسعهٔ این روش همراه با مبانی تئوریک آن ارائه شده است. همچنین در این مقاله بمنظور سهولت پیاده‌سازی و کاهش محاسبات در مرحلهٔ طراحی، فرم بسته‌ای برای ماتریسهای طراحی استخراج شده است. نهایتاً مسألهٔ کنترل سرو سیستم بصورت کامل با استفاده از کنترل کنندهٔ پیش‌بین صریح مورد ارزیابی قرار گرفته و نتایج حاصل از آن مؤید عملکرد مناسب و راندمان قابل قبول روش مورد نظر می‌باشند.

کلمات کلیدی: کنترل پیش‌بین صریح، کنترل بهینه، سیستمهای مقید و سرو سیستم.

1- Introduction

The first key question in control theory is the computation of suitable stabilizing controller for dynamical systems. From the practical point of view this question should be corrected, because in practice we have not only a dynamical system but also several constraints must be satisfied. The pore over the theoretical and practical literatures shows that the most important class of dynamical systems in practice is the "linear systems with constraints". Handling constraints besides the performance objectives makes the controller design procedure more complicated. It is well accepted that for this class of systems, in general, stability and good

performance can only be achieved with a nonlinear control law.

The most popular approaches for designing nonlinear controller for linear systems with constraints fall into two categories [1], Anti-windup shames and model predictive control. In first approach it is assumed that a well functioning linear controller is available for small excursions from the nominal operating point. This controller is augmented by the anti-windup shame in somewhat ad hoc fashion, to take care of situations when constraints are met. In [2] numerous apparently different anti-windup shames were reviewed. The main drawback of anti-windup shames is that, in general, the systematic and automatic synthesis tool which guarantees closed loop stability and achieve some kind of optimal performance has remained an open problem. Recently some promising works were done to introduce a systematic approach for designing and stability analyzing of the anti-windup shames [3, 4 and 5]. Despite these drawbacks, the anti-windup shames are widely used in practice in most SISO cases; they are simple to design and work adequately. For complex constrained multivariable control problems, model predictive control has become the accepted standard, especially in process industries [6]. From the theoretical and practical point of view, the main advantages of MPC are (i) stability of the closed-loop system can be guaranteed, (ii) It is possible to handle hard constraints in multivariable systems in a systematic manner. This allows the designers to guarantee the satisfaction of all constraints during the operation, and (iii) the control system performance is optimal with respect to the objectives. Unfortunately, the MPC has a prohibitive drawback rooted in its demanding computational effort. In [7] Lee and Markus were written an interesting paragraph that described a hypothetical method for obtaining a closed loop controller from open loop trajectories [8]: "One technique for obtaining a feedback controller synthesis from knowledge of open-loop controllers is to measure the current control process state and then compute very rapidly for the open loop control function. The first portion of this function is then used during a short time interval, after which a new measurement of the process state is made and a new open loop control function is computed for this new measurement. The procedure is then repeated". This remarkable proposition was simply forgotten probably due to the high computational cost of the algorithm. More than ten years later this idea recalled by Richalet namely model predictive heuristic control (MPHC) [9]. After few years, in the 1980s model predictive control under its various guises (DMC [10] and GPC [11]) took the process industries by storm [12].

Here, the model predictive control refers to a class of control algorithms that compute a manipulated variable trajectory from a linear process model to minimize an objective function subject to linear constraints on a prediction horizon. At each sampling time, starting at current state, the above open loop optimal control problem is solved over a finite prescribed horizon. At the next sample, measurements are used to update the optimization problem, and the computation is repeated over a shifted horizon, leading to moving horizon policy.

Over the last 25 years a solid theoretical foundation for MPC has emerged and it has undergone many developments regarding the on-line implementation and type of systems and applications that can be handled. Unfortunately, all mentioned nice properties of MPC are accompanied with several sever drawbacks that prevent a broad use of this method in the practice. These obstacles can be summarized in two parts [13, 14], (i) computational complexity connected with the optimization problem, and (ii) the question of the control system robustness to deviations between actual process and its model used for prediction.

This paper substantially deals with the first problem, aims to show the possible way to overcome the computational complexity. Generally speaking, this is done by moving all the computations necessary for the implementation of MPC off-line, while preserving all its other characteristics. This should increase the range of applicability of MPC to problems where anti-windup shames and other ad hoc techniques dominated up to now [15, 16].

Optimal control problems for constrained discrete-time systems based on linear programming were formulated in the early 1960s by Zadeh [17]. Explicit solutions for different types of MPC have been available for unconstraint linear systems minimizing a quadratic objective for decades, namely linear quadratic regulator (LQR), but for no other problem class including constrained systems until recently.

In the pioneering work [18], the authors show how to formulate an optimal control problem for constrained linear systems as a multi-parametric program, where the states of the system are treated as parameters, and control inputs as optimization variables. The solution to this problem has a closed form expression as a function of the plant states. In [1] and [19], it is shown that for linear (norms $1/\infty$) and quadratic cost

functions (i.e. mp-LP and mp-QP), the optimal control input is a piecewise affine (PWA) state feedback law over a polyhedral partition of the feasible state space [19]. The computation of optimal control input in practice is twofold, in the first step, according to the current state the active polyhedral partition is recognized, and then a corresponding affine function is evaluated very quickly. In the following sections, the complete framework of the explicit MPC has been studied from theoretical and practical point of view. Also, in this paper a simple closed form representation of six matrices (*Y, H, F, G, W* and *E*) which are presented in the following are well established. Finally, the illustrative example describes the approach through the paper.

2- Mathematical Background: Multi Parametric Programming

The operations research community addressed parameter variations in mathematical programs at two levels [21]: sensitivity analysis, which characterizes the change of the solution with respect to small perturbations of the parameters; and parametric programming, where the characterization of the solution for a full range of parameter values is sought. One can distinguish between parametric programs, in which only one parameter is considered, and multi-parametric programs depending on a vector of parameters. Solving parametric linear programs was proposed by Gass and Satty [22]. Also extensive research has been devoted to sensitivity and parametric analysis by Gal [23]. Multi-parametric linear programming (mp-LP) was treated in [24, 25]. Recently, the method for solving multi-parametric quadratic programming (mp-QP) was proposed in [1]. The mp-QP can be solved by applying the first order Karush-Huhn-Tucker (KKT) conditions (see for example [26]). Consider the following almost general parametric quadratic program:

$$J(\theta)=\min_{U}\left\{\frac{1}{2}\theta^T Y\,\theta+\theta^T FU+\frac{1}{2}U^T HU\right\}$$
$$s.t.\ \ GU\le W+E\theta \qquad (1)$$
$$U\in\mathbf{U}\subseteq\Re^s,\quad \theta\in\mathbf{\Theta}\subseteq\Re^n$$

H is an *(s×s)* symmetric positive definite constant matrix, *G* and *E* are *(p×s)* and *(p×n)* constant matrices, respectively. *W* is a constant vector of dimension *p,* **U** and **Θ** are compact polyhedral convex sets of dimensions *s* and *n*, respectively. Transforming the QP problem Eq.1 into the standard multi-parametric programming problem is easy, once the following linear transformation is considered:

$$z\triangleq U+H^{-1}F^T\theta \qquad (2)$$

The QP in Eq.1 is then formulated to the following standard multi-parametric programming (mp-QP) problem:

$$V_z(\theta)=\min_{z}\frac{1}{2}z^T Hz\quad s.t.\ Gz\le W+S\theta \qquad (3)$$

Where H=HT>0, $S\triangleq E+GH^{-1}F^T$ and $V_z(\theta)=J(\theta)-0.5\theta^T\left(Y-FH^{-1}F^T\right)\theta$. Note that in the transformed problem, the parameter vector θ appears only on the rhs of the constraints. Solving the above optimization problem was addressed in [1], which is substantially based on the "Basic Sensitivity" theorem [28, 29].

Lemma: let $\theta_0\in R^n$ be a vector of parameters and (z_0,λ_0) be a KKT pair corresponding to Eq.3, where $\lambda_0=\lambda(\theta_0)$ is a vector of nonnegative Lagrange multipliers, λ, and $z_0=z(\theta_0)$ is feasible in Eq.3. Also assume that (see [26] for details) (i) strict complementary slackness (SCS) holds, and (ii) linear independence constraint qualification (LICQ) holds. Then in the neighborhood of θ_0, there exists a unique, once continuously differentiable function, $\left[z(\theta),\lambda(\theta)\right]$, where $z(\theta)$ is a unique isolated minimizer for Eq.3, and

$$\begin{pmatrix}\dfrac{dz(\theta_0)}{d\theta}\\ \dfrac{d\lambda(\theta_0)}{d\theta}\end{pmatrix}=-(M_0)^{-1}N_0,$$

$$M_0=\begin{bmatrix}H & G_1^T & \cdots & G_p^T\\ -\lambda_1 G_1 & -V_1 & & \\ \vdots & & \ddots & \\ -\lambda_p G_p & & & -V_p\end{bmatrix}, \qquad (4)$$

$$N_0=\begin{bmatrix}\emptyset & \lambda_1 S_1 & \cdots & \lambda_p S_p\end{bmatrix}^T$$

Where G_i, S_i and W_i denote the ith row of G, S and W, respectively, $V_i=G_i z_0-W_i-S_i\theta_0$, and $\emptyset$ is a null matrix of dimension *(s×n)*. Under the assumptions of theorem, near the feasible point $\theta=\theta_0\in\mathbf{\Theta}$, a first order approximation of $\left[z(\theta),\lambda(\theta)\right]$ can be obtained in the neighborhood of the KKT point:

$$\begin{pmatrix}z(\theta)\\ \lambda(\theta)\end{pmatrix}=-(M_0)^{-1}N_0(\theta-\theta_0)+\begin{pmatrix}z(\theta_0)\\ \lambda(\theta_0)\end{pmatrix} \qquad (5)$$

Proposition:

Let $H>0$ and the assumptions of previous lemma hold. Consider the optimization problem of mpQP given in Eq.3. Then, the optimal solution z and the associated Lagrange multipliers λ are affine functions of the optimization parameter θ.

Proof:

Assuming the Lagrangian $L = V_z + \lambda^T (Gz - W - S\theta)$, the first order KKT conditions for the mp-QP problem in Eq.3 are given by:

$$\begin{aligned} &\nabla L = Hz + G^T \lambda = 0, \\ &\lambda_i \left(G_i z - W_i - S_i \theta\right) = 0,\ i = 1,....,p \\ &\lambda \geq 0. \end{aligned} \tag{6}$$

The first equation comes directly from optimality and the second equation refers to the complementary slackness [26].

Recalling that H is invertible Eq.6 is written as

$$z = -H^{-1} G^T \lambda, \tag{7}$$

Let $\bar{\lambda}$ and $\tilde{\lambda}$ denote the Lagrange multipliers corresponding to the inactive and active constraints, respectively. For the inactive constraints $\bar{G}z - \bar{W} - \bar{S}\theta < 0$, therefore the complementary slackness yelds $\bar{\lambda} = 0$. For the active constraints,

$$\tilde{G}z - \tilde{W} - \tilde{S}\theta = 0 \tag{8}$$

Where $\tilde{G}, \tilde{W}, \tilde{S}$ correspond to the set of active constraints. From Eq.6

$$\tilde{\lambda} = -\left(\tilde{G}H^{-1}\tilde{G}^T\right)^{-1}\left(\tilde{W} + \tilde{S}\theta\right) \tag{9}$$

Note that $\left(\tilde{G}H^{-1}\tilde{G}^T\right)$ is invertible because of LICQ assumption. This $\tilde{\lambda}$ is an affine function of θ. Substituting $\tilde{\lambda}$ into Eq.7 yelds:

$$z = H^{-1}\tilde{G}^T\left(\tilde{G}H^{-1}\tilde{G}^T\right)^{-1}\left(\tilde{W} + \tilde{S}\theta\right) \tag{10}$$

Note that z is also an affine function of θ. □

Equations 5 and 10 show that given the solution z_0, λ_0 for a specific vector of parameters θ_0, one can obtain the solution $z(\theta), \lambda(\theta)$ for any parameter vector θ from Eq.5. The set of θ where Eq.5 remains optimal is defined as the critical region (CR0) and can be obtained as follows [27]. Let CRR represent the set of inequalities obtained *(i)* by substituting z(θ) into the inactive constraints in Eq.3, and *(ii)* from the positivity of the Lagrange multipliers corresponding to the active constraints, as follows:

$$CR^R = \left\{\bar{G}z(\theta) \leq \bar{W} + \bar{S}\theta,\ \tilde{\lambda}(\theta) \geq 0\right\} \tag{11}$$

Then CR0 is obtained by removing the redundant constraints from CRR as follows:

$$CR^0 = \Delta\left\{CR^R\right\} \tag{12}$$

Where Δ is an operator which removes the redundant constraints (see [22] for details). In the next step the rest of region CR^{rest}, is obtained:

$$CR^{rest} = \mathbf{\Theta} - CR^0, \tag{13}$$

By exploiting the procedure described in [1], equations 5 and 11-13 are repeated and a set of $z(\theta), \lambda(\theta)$ and corresponding CR^0 is obtained. The solution procedure terminates when no more regions remained, i.e., $CR^{rest} = 0$. Finally, to find a compact representation, the regions which have the same solution must be unified to give a convex region. The presented procedure is summarized in Fig.1 in seven steps. In Figs.2 and 3 the state-space partitioning algorithm and obtaining the corresponding optimal PWA control function are illustrated based on the given flowchart. As shown figures 2 and 3 at first the critical region corresponding to θ_0 (CR^0) is computed then the parameter space is divided by reversing one by one the hyper planes defining the critical region CR^0. Iteratively each new region CR^i can be subdivided in a similar way as was done with the initial region $\Theta = CR^i$.

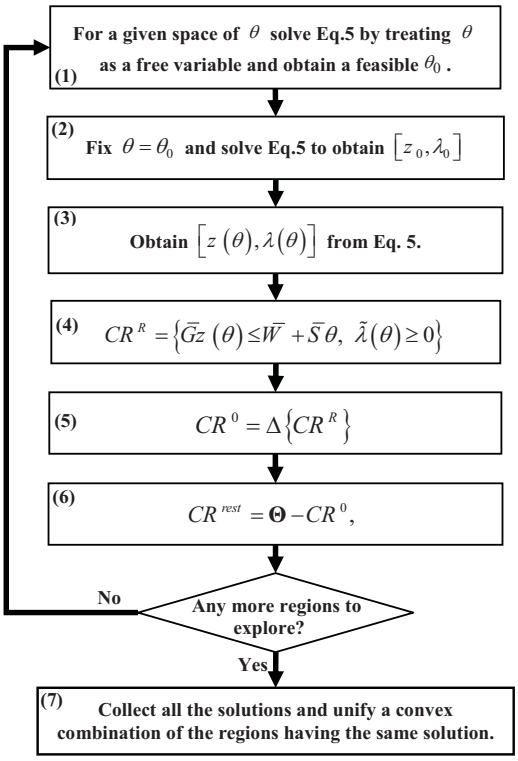


Fig.1: The flowchart of mpQP design steps.

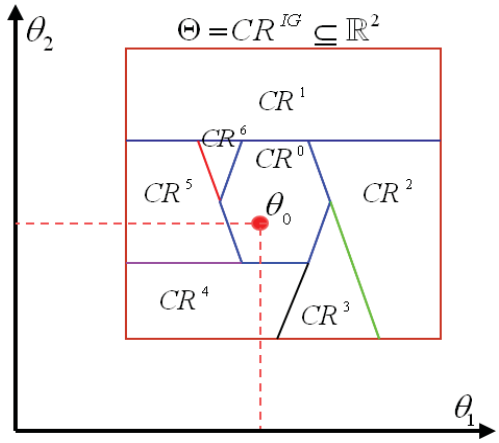


Fig.2: State space exploration strategy.

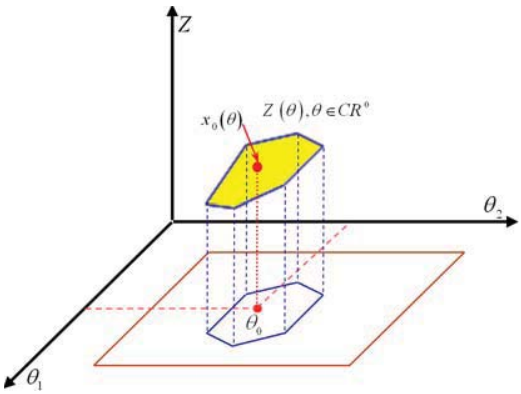


Fig.3: Corresponding affine optimal value for a specific partition.

3- A Unified Structure for Explicit Model Predictive Control

Consider the problem of regulating to the origin the discrete-time linear time-invariant system

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t), \end{aligned} \tag{14}$$

While fulfilling the constraints

$$\begin{aligned} y_{\min} &\le y(t) \le y_{\max}, \\ u_{\min} &\le u(t) \le u_{\max}, \end{aligned} \tag{15}$$

At all time instants $t \ge 0$. In (14) and (15), $x(t) \in R^n$, $u(t) \in R^m$, and $y(t) \in R^r$ are the state, input, and output vector respectively, and the pair *(A,B)* is stabilizable.

In (15), $y_{\min} \le 0 \le y_{\max}$, and $u_{\min} \le 0 \le u_{\max}$, are vectors of upper and lower bounds (more generally, we can allow only some components of the inputs or outputs to be constrained).

MPC solves such a constrained regulation problem in the following way. Assume that a full measurement of the state is available at the current time t. Then, the optimization problem is:

$$\begin{aligned} J(x(t)) &= \min_U \left\{ \begin{array}{l} x'_{t+N|t} P x_{t+N|t} + \\ \sum_{k=0}^{N-1} \left(x'_{t+k|t} Q x_{t+k|t} + u'_{t+k|t} R u_{t+k|t} \right) \end{array} \right\} \\ \text{subj.to} \quad & y_{\min} \le y_{t+k|t} \le y_{\max}, \quad k = 1,\dots,N \\ & u_{\min} \le u_{t+k|t} \le u_{\max}, \quad k = 0,\dots,M-1 \\ & x_{t|t} = x(t), \\ & x_{t+k+1|t} = Ax_{t+k|t} + Bu_{t+k|t}, \quad k \ge 0, \\ & y_{t+k|t} = Cx_{t+k|t}, \quad k \ge 0, \\ & u_{t+k|t} = Kx_{t+k|t}, \quad M \le k \le N-1, \end{aligned} \tag{16}$$

With respect to $U \triangleq \left(u_t^T, \dots, u_{t+k-1}^T\right)^T$, where $R=R^T>0$, $Q=Q^T \ge 0$, $P=P^T>0$, $x_{t+k|t}$ is the prediction of x_{t+k} at time t, and M is the control input horizon. When the final cost matrix P and gain K are calculated from the algebraic Riccati equation, under the assumption that the constraints are not active for $k \ge N$ in Eq.16 exactly solve the constrained (infinite-horizon) LQR problem for Eq.5 with weights Q, R (see also, [28] and [29]).

One possible choice is to set $K=0$ and P to be the solution of the discrete Lyapunov equation $P=A^TPA+Q$. However, this solution is restricted only to open-loop stable systems [29], since the control action is switched off after N steps. Alternatively, one can choose K and P as the solution of unconstrained, infinite horizon LQR problem, i.e. $N = \infty$, as follows:

$$\begin{aligned} K &= -\left(R + B^T PB\right)^{-1} B^T PA, \\ P &= \left(A + BK\right)^T P \left(A + BK\right) + K^T RK + Q \end{aligned} \tag{17}$$

Introducing $x_{t+k|t} = A^k x_t + \sum_{j=0}^{k-1} A^j Bu_{t+k-1-j}$ derived from Eq.14 into Eq.16 results the following QP problem:

$$\begin{aligned} J^*(x_t) &= \min_U \left\{ \frac{1}{2} U^T HU + x^T FU + \frac{1}{2} x_t^T Y x_t \right\} \\ & s.t.\ GU \le W + Ex_t \\ & U \in \mathbf{U} \subseteq R^s, s = mN \\ & x_t \in \mathbf{X} \subseteq R^n \end{aligned} \tag{18}$$

Note that six parameters (*H, F, Y, G, W* and *E*) are depended to known parameters Q, R and Eq.16.

To calculate these six matrices it is necessary to replace $x_{t+k|t} = A^k x_t + \sum_{j=0}^{k-1} A^j Bu_{t+k-1-j}$ in Eq.16 and reordering while x_t and $(u_t, \dots, u_{t+k-1})$ are the parameters and *(A, B, P, Q and R)* are known values.

Despite the fact that this is done off-line, when the horizon *(N)* is big (e.g., more than 5) performing this procedure needs several many matrix operation and parametric operation to obtain the *H, F, Y, G, W* and *E* matrices. This makes the design procedure highly complicated and because of the parametric manipulation it is hard to be done systematically.

In order to remove parametric manipulation and to reduce the off-line computation complexity, a set of simple closed form formulas for the six matrices *H, F, Y, G, W* and *E* has been calculated after a comprehensive algebraic manipulation (see appendix for detail). The resulting closed form formulas can be calculated systematically for any given horizon (*N*) with only numerical operations.

Let $U^*(x_t)=\left(u_t^{*T},...,u_{t+k-1}^{*T}\right)^T$ be the optimal solution of Eq.18 at time t. then, the first sample of $U^*(x_t)$ is applied to system in Eq.14:

$$u(t)=u_t^*(x_t), \tag{19}$$

The optimization (Eq.18) is repeated at time *t+1*, based on the new state *x(t+1)*, yielding a moving or receding horizon control strategy [19]. Comparing model predictive control on-line optimization problem (Eq.18) and multi-parametric programming problem (Eq.1) shows that the Eq.18 can be solved explicitly with substituting plant state vector *x(t)* as the parameter vector $\theta(t)$ in multi-parametric quadratic programming problem. Moreover, the Eq.18 can be simply transformed to standard form of mp-QP (Eq.3), exploiting the transformation $z=U+H^{-1}F^Tx_t$. All remaining steps for obtaining optimal control input $u(t)=u_t^*(x_t)$, are straight forward as shown in Fig.4.

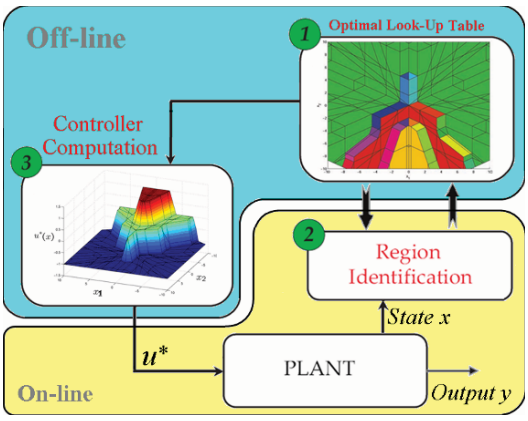


Fig.4: The complete structure of the EMPC.

In Fig.4 the structure of implementing the presented method (EMPC) is shown. This structure has two main parts, on-line and off-line. The off-line part has two databases, the first one contains control regions information, and the second one contains the corresponding PWA functions. The on-line part, i.e. region identifier, is used for recognizing the active critical region for current measured state at each sampling time. Note that this part is the most important part when one takes care of computational complexity in the practice.

4- Simulation Results

In this section two illustrative examples are studied to describe the approach through the paper. At first using a simple artificial example the whole presented method is exerted. Then, a practical servo system is considered to describe the way of extending the standard problem to the tracking problem using the presented EMPC method.

4.1- Simple illustrative example

Consider a simple second order discrete time system with *Ts = 0.1* second:

$$\begin{aligned} x(t+1)&=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}x(t)+\begin{bmatrix}1\\ 0.5\end{bmatrix}u(t)\\ y(t)&=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}x(t). \end{aligned} \tag{20}$$

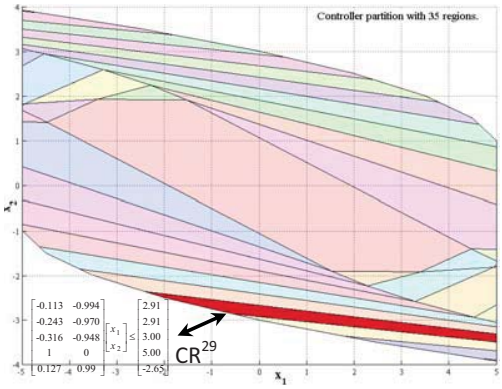
The constraints are $|u(t)|\le 1$ and $|y(t)|\le 5$.

The corresponding optimization problem of the form Eq.16 for regulating to the origin is given as follows:

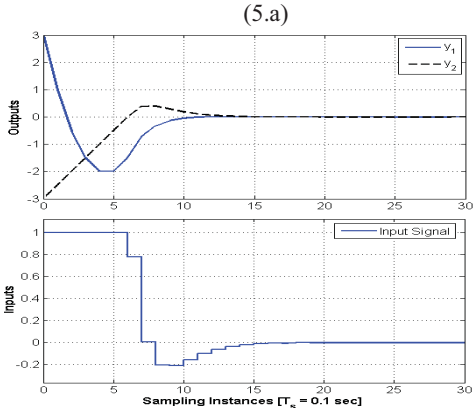
$$\begin{aligned} &\min_U\left\{x_{t+15|t}^T P x_{t+15|t}+\sum_{k=0}^{14}\left(x_{t+k|t}^T x_{t+k|t}+u_{t+k|t}^2\right)\right\}\\ &subj.to \quad -1\le u_{t+k|t}\le 1,\ \ k=0,...,14\\ &\qquad -5\le y_{t+k|t}\le 5,\ \ k=1,...,15\\ &\qquad x_{t|t}=x_t, \end{aligned} \tag{21}$$

Where $U=\{u_t,u_{t+1},...,u_{t+14}\}$, *N=15* and *P* solves the Lyapunov equation with *Q=diag(1,1), R=1* and *N=M=15*. The corresponding mp-QP problem of the form Eq.3 associated to the MPC problem can be simply calculated by using the given closed form formulas (see Appendix) and then the transformation $z=U+H^{-1}F^Tx_t$. The solution of mp-QP problem can be computed by Algorithm presented in Fig.1, and the corresponding 35 polyhedral partitions of the state-space are depicted in Fig. 5(a). The closed loop response for a worst case initial condition $x_0=[3\ -3]^T$ is shown in Fig.5(b). Also the resulting PWA control function and corresponding PWQ optimal cost

function are depicted in Fig.6. To illustrate the EMPC operates in practice, consider the starting point $x_0=[3\ -3]^T$. This point is substituted into the constraints defining the *CRs* and it satisfies only the constraint of CR^{29} that indicated in Fig.5 (a). The control action corresponding to CR^{29} is $u^{29} = 1$, which does not require any further calculation and it is same as the one obtained from the on-line MPC but without any optimization time.



(5.a)



(5.b)

Fig.5. (a) polyhedral partitions, of the state-space, (b) Closed loop response of the EMPC, $x_0 = [3\ -3]^T$.

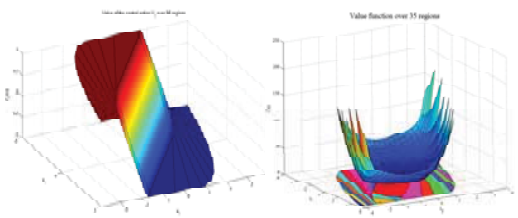


Fig.6. PWA control input and corresponding PWQ cost function.

4.2- Application of EMPC for Servo System

Here we have applied the studied EMPC method to an extensively used system as a direct-current motor in which the sampling frequency is almost high (i.e. 0.06 second) and the traditional MPC may face to trouble in the on-line optimization when the constraints are involved. The input of the system is the voltage applied to the motor *(V)* and the output is the shaft angle *(θ)*. It is obvious that the process has an integral effect, given that the position grows indefinitely whilst it is fed by a certain voltage. In order to obtain a model which describes the behavior of the motor, the inertia load (proportional to the angular acceleration) and the dynamic friction load (proportional to angular speed) are taken into account. Their sum is equal to the torque developed by the motor, which depends on the voltage applied to it. It is a first-order system with regards to speed but a second-order one if the angle is considered as the output of the process:

$$J\frac{d^2\theta}{dt^2}+b\frac{d\theta}{dt}=T_m \tag{22}$$

Where Tm, J and b are the motor's torque, total inertial and friction (damping) coefficient, respectively which are depend on the electromagnetic/mechanical characteristics of the motor. The inertia divided by the friction coefficient is referred to as the motor's time constant, $\tau=J/b$. So the model of the system becomes:

$$\tau\ddot{\theta}(t)+\dot{\theta}(t)=K\,u(t) \tag{23}$$

Where $K=\alpha/b$, and α is a constant gain which relates the input voltage *(u)* to the torque (T_m).

The EMPC controller is going to be implemented on a real motor with a feed voltage of *24 V* and nominal current of *1.3 Amp*, subject to a constant load. The Reaction Curve Method (RCM) is used to obtain experimentally the parameters of the motor, applying a step in the feed voltage and measuring the evolution of the angular speed (which is a first-order system). The obtained five parameters (*J, b, α, K* and *τ*) based on Reaction Curve Method (RCM) are given in table 1.

Table 1: Motor numerical parameters

J [kgm2]	b [Nm.sec]	α [Nm/volt]	K [(sec.volt)-1]	τ [sec]
50e-6	5.56e-5	1.39e-4	2.5	0.9

Choosing $x_1=\theta$, $x_2=\dot{\theta}$, and taking the sampling time of $T_s=0.06$ *second* one gets the discrete time state space model of the servo:

$$x(t+1)=\begin{bmatrix}1 & 0.1451\\ 0 & 0.9355\end{bmatrix}x(t)+\begin{bmatrix}0.00489\\ 0.06449\end{bmatrix}u(t)$$
$$y(t)=\begin{bmatrix}1 & 0\end{bmatrix}x(t). \tag{24}$$

The practical constraints which are needed to be satisfied during the operation are as follows:

$$|\theta(t)|\le 20\deg,\quad |\dot{\theta}(t)|\le 10\frac{\deg}{\sec},\quad |u(t)|\le 2volt \tag{25}$$

Here we need some modification to achieve output reference tracking. To this aim the state space model should be augmented.

Since the input which is necessary to keep the state at the reference point is not generally known, therefore, the dynamics are reformulated as a so-called δu formulation which guaranties achieving to the mentioned goal:

$$\begin{bmatrix}x(t+1)\\ u(t)\\ y_{ref}(t+1)\end{bmatrix}=\begin{bmatrix}A & B & 0\\ 0 & 1 & 0\\ 0 & 0 & I\end{bmatrix}\begin{bmatrix}x(t)\\ u(t-1)\\ y_{ref}(t)\end{bmatrix}+\begin{bmatrix}B\\ I\\ 0\end{bmatrix}\delta u(t)$$
$$y(t)=\begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x(t)\\ u(t-1)\\ y_{ref}(t)\end{bmatrix}. \tag{26}$$

The above augmented system is similar to adding an integrator in control loop to remove tracking error. Now if a regulator be designed for the augmented system, then the tracking will be attained for original servo system.

Let (A_u, B_u, C_u, D_u) and $\bar{x}$ be the augmented matrices and state vector respectively. Equivalently, the corresponding optimization problem of the form Eq.16 for tracking problem is given as follows:

$$\min_U \sum_{k=0}^{2}\left(\begin{array}{l}\left(y_{t+k|t}-r(t)\right)^T Q_y\left(y_{t+k|t}-r(t)\right)\\ +\delta u_{t+k|t}^T R\,\delta u_{t+k|t}\end{array}\right)$$
$$\begin{aligned} subj.to\quad & -20\deg\le x_1(t+k)\le 20\deg\\ & -10\deg/\sec\le x_2(t+k)\le 10\deg/\sec\\ & -2\le u(t+k)=x_3(t+k)\le 2,\ k=0,1,2\\ & \bar{x}_{t|t}=\bar{x}_t,\end{aligned} \tag{27}$$

Where $U=\{\delta u_t, \delta u_{t+1}, \delta u_{t+2}\}$, $r=y_{ref}$, $Qy=500$, $R=0.05$ and $N=3$. Like the regulation problem, one can transform the tracking problem into:

$$J^*(x_t)=\min_U\left\{\begin{array}{l}\frac{1}{2}U^T HU+\\ \left[x^T(t)\ \ u^T(t-1)\ \ r^T(t)\right]FU\end{array}\right\}$$
$$s.t.\ GU\le W+E\begin{bmatrix}x(t)\\ u(t-1)\\ r(t)\end{bmatrix} \tag{28}$$

The same algorithm can be used to obtain an explicit PWA solution $\delta u(t)=F(x(t),u(t-1),r(t))$.

The closed loop response of the explicit controller for varying reference signal and the input control effort are depicted in Fig.7 and Fig.8 for initial condition $x_0=[-12\ \ 5]^T$.

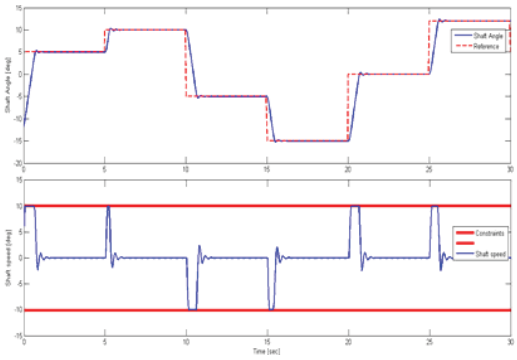


Fig.7. closed loop response of the explicit controller.

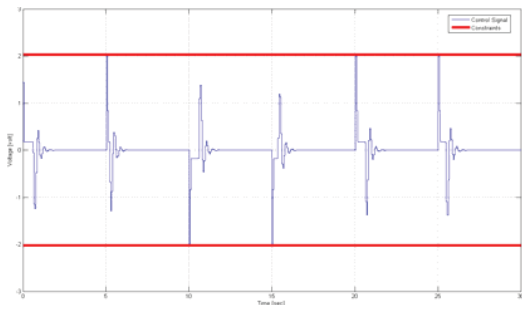


Fig.8. the corresponding input signal (explicit controller).

The solution of mp-QP problem was computed by Algorithm presented in Fig.1, and the corresponding polyhedral partitions of the state-space are 4-dimensional. In Figs.9 and 10 cutting the polyhedral partitions of augmented system on $x_3=x_4=0$ and $x_4=0$ are shown.

To illustrate how the EMPC operates in practice, consider the starting point $x_0=[-0.2\ \ -0.15]^T$ where the reference signal is $r(t)=0.2$ rad. This point is substituted into the constraints defining the critical regions (*CRs*) depicted in the Figs.9 and 10 and it

satisfies only the constraints of CR^{40}. The polyhedron which describes the CR^{40} is:

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ -0.99 & -0.14 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0.62 & 0.48 & 0 & -0.62 \end{bmatrix} \begin{bmatrix} x(t) \\ u(t-1) \\ r(t) \end{bmatrix} \leq \begin{bmatrix} -0.0858 \\ 0.3551 \\ 0.1745 \\ 2 \\ 2 \\ 0.3491 \\ -0.095 \end{bmatrix} \quad (29)$$

The control action corresponding to CR^{40} is:

$$\begin{aligned} \delta u^{40}(t) &= \begin{bmatrix} 0 & 0 & -1 & 0 \end{bmatrix} \bar{x} + 2 \\ &= -u(t-1) + 2 \end{aligned} \quad (30)$$

The control signal will be $u(t) = u(t-1) + \delta u^{40}(t)$, or $u(t) == 2$ which does not require any further calculation. The corresponding region (CR^{40}) is shown in Figs.9 and 10.

In order to comparison between explicit MPC and traditional online MPC the same scenario has been used for online MPC. The error signals $e_u = u_{EMPC} - u_{MPC}$ and $e_y = y_{EMPC} - y_{MPC}$ are shown in Fig.11.

According to the Fig.11, the results of two EMPC and MPC controller are almost the same. The main difference between these two methods is the computation burden that should be compared. To this aim, the total time which is needed for performing the given scenario (Fig.5) is computed for both methods in the MATLAB ver.7 platform using a Pentium IV processor. The total time when the EMPC is used is about 0.546 second, while the online MPC takes 5.047 second for its optimization. This comparison shows that the explicit method is almost 10 times faster than the online MPC. The important note is that when the problem becomes more complicated and faster (in sense of its dynamic) but no high dimensional, the computation time in EMPC remains almost fixed because of the fact that it needs only a simple affine function evaluation. But the online MPC will be more time consuming and even prohibitive. When the problem is high dimensional the EMPC won't be effective as much as the online MPC is now. This is because in the case of high dimensional problems the number of partitions might increase prohibitively. So as a future work and an active research direction this problem can be dealt with.

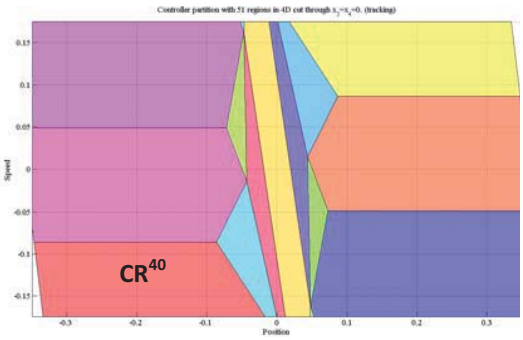


Fig.9. the 2D polyhedral partitions of the augmented system cut through x3=x4=0.

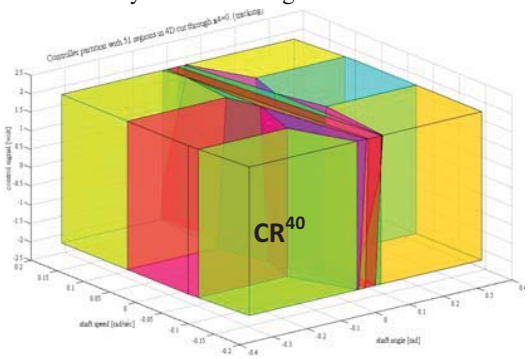


Fig.10. the 3D polyhedral partitions of the augmented system cut through x4=0.

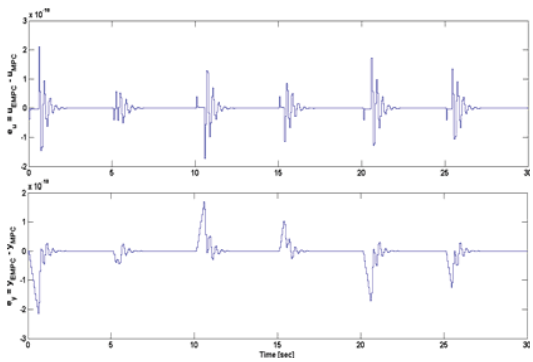


Fig.11. the error signals between EMPC and MPC

5- Conclusion

The explicit solution of the linear MPC as an optimization problem with a quadratic objective and linear constraints were studied. Based on the recent development in [1], it was shown that exploiting multi-parametric programming a complete map of the optimal control as a function of the states can be achieved. Also it was shown that this solution is piecewise affine feedback control law and the closed loop on-line operation of this structure is nothing but a simple PWA function evaluation. In order to complexity reduction of the design procedure a computationally simple closed form of matrices *Y, H,*

F, G, W and *E* were well established in this paper. This modification would reduce the computational complexity significantly. This nice characteristic allows the MPC to be able to apply for systems with high sampling rate (with no huge dimension till now), in the range of micro second, as claimed in some other literatures as well (see for example [30]). A complete unified structure was presented in this paper and the presented algorithm was verified with a simple illustrative example, in which all steps of EMPC and corresponding mp-QP problem were studied and then using a well known practical servo system the standard EMPC method was developed to deal with tracking problems.

Appendix

Substitution of $x_{t+k|t} = A^k x_t + \sum_{j=0}^{k-1} A^j B u_{t+k-1-j}$ into the Eq.14 and after a plenty of algebraic manipulation one would obtain the matrices (*Y, H, F, G, W* and *E*) as follows:

$$(1):\ Y = 2\left(A^N\right)^T PA^N + 2\sum_{k=0}^{N-1}\left(A^k\right)^T QA^k$$
$$(2): H = 2\left(\bar{R} + B^T \bar{H} B\right), \tag{31}$$

where

$$\bar{R} = diag\underbrace{\{R, R, ..., R\}}_{N\ times}$$

$$\bar{H}_{ij} = \begin{cases} \left(A^{N-1-i}\right)^T PA^{N-1-j}, \\ \qquad for\left(i = N-1\right) or\ \left(j = N-1\right) \\ \left(A^{N-1-i}\right)^T PA^{N-1-j} + \\ \sum_{k=\max(i,j)}^{N-2}\left(A^{k-i}\right)^T QA^{k-j}, \quad ow \end{cases}$$

and

$$(3):\ F = 2\left[f_0 B\ f_1 B\ ...f_{N-1} B\right], \tag{32}$$

where:

$$f_i = \begin{cases} \left(A^N\right)^T P, & for: i = N-1 \\ \left(A^N\right)^T PA^N + \sum_{k=i+1}^{N-1}\left(A^k\right)^T QA^{k-1-j}, & ow \end{cases}$$

$$(4): G = \begin{bmatrix} I_{s\times s} \\ -I_{s\times s} \\ 0_{2n\times s} \end{bmatrix}, \qquad (5):\ E = \begin{bmatrix} 0_{2s\times n} \\ -I_{n\times n} \\ I_{n\times n} \end{bmatrix}, \tag{33}$$

$$(6):\ W = \begin{bmatrix} U_{\max} \\ -U_{\max} \\ X_{\max} \\ -X_{\max} \end{bmatrix} \tag{34}$$

where:

$$U_{\max} = \begin{bmatrix} u_{\max} \\ \vdots \\ u_{\max} \end{bmatrix}_{s\times 1}, U_{\min} = \begin{bmatrix} u_{\min} \\ \vdots \\ u_{\min} \end{bmatrix}_{s\times 1}$$

Using these close forms will decrease the complexity of the design procedure and make it straightforward by removing several many intermediate high dimensional parametric matrix operations, e.g. multiplication, summation and reordering.

References

[1] Bemporad, A., Morari, M., Dua, V., Pistikopoulos, E., "The explicit linear quadratic regulator for constrained systems", Automatica, 38 (1), 3–20, 2002.

[2] M.V. Kothare, P.J. Campo, M. Morari, and C.N. Nett. "A unified framework for the study of anti-windup designs", Automatica, 30(12), 1994.

[3] J.M. Gomes da Silva Jr., S. Tarbouriech, “Anti-windup design with guaranteed regions of stability for discrete-time linear systems”, Systems & Control Letters, Volume 55, Issue 3, March 2006.

[4] Seong-Sik Yoon, Jong-Koo Park, Tae-Woong Yoon , “Dynamic anti-windup scheme for feedback linearizable nonlinear control systems with saturating inputs” Automatica, 44(12), pp: 3176-3180, 2008.

[5] Tingshu Hu, Andrew R. Teel, Luca Zaccarian, “Anti-windup synthesis for linear control systems with input saturation: Achieving regional, nonlinear performance” Automatica, 44(2), pp: 512-519, 2008.

[6] Qin, S. and Badgwell, T. “A survey of industrial model predictive control technology”, Control Engineering Practice, pp 733–764, 2003.

[7] E. B. Lee and L. Markus. Foundations of Optimal Control Theory. Wiley.New York 1967.

[8] Ali Jadbabaie. Receding Horizon Control of Nonlinear Systems: A Control Lyapunov Function Approach, PhD thesis. California Institute of Technology. Pasadena. Oct. 13, 2000.

[9] Richalet, J., Rault, A., Testud, J. L., Papon, J. "Model predictive heuristic control-application to industrial processes", Automatica, 14,413-328, 1978.

[10] CUTLER, C. R., and RAMAKER, B. L. "Dynamic matrix control-a computer control algorithm", Proceedings of the Joint Automatic Control Conference, 1980, USA.

[11] CLARKE, D. W., MOHTADI, C., and TUFFS, P. S. “Generalized predictive control—Part I. The basic algorithm”, Automatica, 23, pp.137-148, 1987.

[12] Manfred Morari, Miroslav Barić, “Recent developments in the control of constrained hybrid

systems", Computers & Chemical Engineering, Volume 30(10-12), pp: 1619-1631, 2006.

[13] Mario Vasak. Time optimal control of piecewise affine systems, PhD thesis. University of Zagreb, 13th July, 2007.

[14] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert. "Constrained Model Predictive Control: Stability and Optimality", Automatica, 36(6), pp:789–814, 2000.

[15] A. Bemporad, F. Borrelli and M. Morari, "The explicit solution of constrained LP-based receding horizon control", in Proceedings of the 39th IEEE Conference on Decision and Control, Sydney, Australia, pp: 632-637, 2000.

[16] P. Tondel, T.A. Johansen, A. Bemporad, "An algorithm for multiparametric quadratic programming and explicit MPC solutions", Automatica 39, pp: 489–497, 2003.

[17] ZADEH, L., and WHALEN, L., "On Optimal Control and Linear Programming", IEEE Transaction on Automatic Control, Vol. 7, pp. 45-46, 1962.

[18] F. Borrelli, M. Baotic, A. Bemporad, M. Morari, "Efficient On-Line Computation of Constrained Optimal Control", Proceedings of the 40th IEEE Conference on Decision and Control Orlando, Florida USA, December 2001.

[19] A. Bemporad, F. Borrelli, and M. Morari, "Model Predictive Control Based on Linear Programming—The Explicit Solution", IEEE Transactions on automatic control, 47(12), 2002.

[20] M. Baric, C. Jones, M. Morari, "Parametric Analysis of Controllers for Constrained Linear Systems", Proceedings of the 45th IEEE Conference on Decision & Control, USA, pp: 13-15, 2006.

[21] F. Borrelli, A. Bemporad, M. Morari, "Geometric Algorithm for Multi-parametric Linear Programming", Journal of optimization theory and applications, 118(3), pp: 515–540, 2003.

[22] Gass, S., and Saaty, T., The Computational Algorithm for the Parametric Objective Function, Naval Research Logistics Quarterly, Vol. 2, pp. 39–45, 1955.

[23] GAL, T., and GREENBERG, H., Advances in Sensitivity Analysis and Parametric Programming, International Series in Operations Research and Management Science, Kluwer Academic Publishers, Dordrecht, Netherlands, Vol. 6, 1997.

[24] GAL, T., and NEDOMA, J., Multi-parametric Linear Programming, Management Science, Vol. 18, pp. 406–442, 1972.

[25] GAL, T., Postoptimal Analyses, Parametric Programming, and Related Topics, 2nd Edition, de Gruyter, Berlin, Germany, 1995.

[26] A.V. Fiacco, Introduction to sensitivity and stability analysis in nonlinear programming. Academic Press, London, U.K., 1983

[27] E. Pistikopoulos, M. C. Georgiadis, and V. Dua, multi-parametric programming, Vol. 1, WILEY-VCH Verlag GmbH & Constraints. KGaA, Weinheim, 2007.

[28] Chmielewski, D., Manousiouthakis, V. "On constrained infinite-time linear quadratic optimal control", Systems and Control Letters, 29, 121–129, 1996.

[29] Scokaert, P., Rawlings, J. B. "Constrained linear quadratic regulation". IEEE Transactions on Automatic Control, 43, 1163–1169, 1998.

[30] P. Dua, K. Kouramas, V. Dua, E.N. Pistikopoulos, "MPC on a chip—Recent advances on the application of multi-parametric model-based control", Computers and Chemical Engineering 32, 754–765, 2008.

Journal of Control

A Publication of Iranian Society of Instrumentation and Control Engineers, Vol. 3, No. 1, Spring 2009.

Publisher: Iranian Society of Instrumentation and Control Engineers
Managing Director: Prof. Iraj Goodarznia
Editor-in-Chief: Prof. Ali Khaki-Sedigh
Tel: 84062317
Email: sedigh@kntu.ac.ir
Assistant Editor: Dr. Hamid Khaloozadeh, Dr. Alireza Fatehi
Executive Director: Dr. Hamid Khaloozadeh

Editorial Board:
Prof. A. Khaki-Sedigh, Dr. H. Khaloozadeh (Associate Prof.), Prof. I. Goodarznia, Prof. P. Jabedar-Maralani, Prof. A. Ghafari, Dr. H. Hassibi (Assistant Prof.), Dr. M.R. Jahed-Motlagh (Associate Prof.), Dr. K. Badie, Dr. Karim Safavi (Associate Prof.), Prof. R. Asgharian,Dr. H.R. Momeni (Associate Prof.), Prof. A. Vahidian-Kamyad, Prof. S. Khanmohammadi, Prof. S.K. Nikravesh, Prof. M. Shafiee, Dr. A. Fatehi (Assistant Prof.), Dr. B. Labibi (Assistant Prof.), Prof. B. Moshiri, Prof. S. Katebi, Dr. A. Maghsoudi-Pour, K. Falamaki (M.Sc.), M. Baradaran-Mozafari (M.Sc.), Gh. A. Ramazani (M.Sc.).

Advisory Board:
Dr. H.R. Momeni, Dr. N. Abedi, Prof. A. Ghaffari, Dr. A. A. Gharehveisi, Dr. M. Tavakoli-Bina, Dr. H.R. Taghirad, Dr. M. Bathaei, Dr. M.T. Hamidi-Beheshti, Prof. B. Moshiri, Prof. M. Shafiee, Prof. R. Asgharian, Prof. A. Khaki-Sedigh, Dr. B. Moaveni, Dr. R. Kazemi, Dr. S.A. Mousavian, Dr. A.H. Markazi-Davaei, Prof. M. Haeri, Dr. A.R. Khalili-Tehrani, Prof. H. Seifi, Dr. A. Kazemi, Dr. H. Khaloozadeh, Dr. M. Aliari-Shourehdeli, Dr. A. Fatehi, Dr. M.R. Akbarzadeh-Toutounchi, Dr. Mirabedini, Dr . H. Pedram, Dr . A. Harounabadi, Prof. A. Vahidian-Kamyad, Dr. M. Arvan, Dr. J. Heirani-Nobari, Dr. B. Labibi,Prof. F. Hossein -Babaei, Dr. K. Aghaei.

The ISICE Board of Director:
A. Sheri-Moghadam, Dr. K. Masroori, Dr. H.R. Momeni, Prof. B. Moshiri, Dr. F. Jafar-Kazemi, Dr. H. Khaloozadeh, A. Rastegari.

Address: Unit 241, 2nd floor, No.27, Mousavi Ave. Ferdowsi Sq. Enghelab St. Tehran, Iran.
P.O. Box: 15815-3595 Tel: (+9821) 88813002 Fax: (+9821) 88324979
http://www.isice.ir

A Publication of Iranian Society of Instrument and Control Engineers

Vol. 3, No. 1, Spring 2009

Persian Part

بررسی یادگیری تقویتی و خواص سیاست بهینه در مسائل جدولی با استفاده از روش های کنترل دیجیتال ۱
سیدمصطفی کلامی هریس، ناصر پریز، محمدباقر نقیبی سیستانی

طراحی موجک بهینه برای نویززدایی از طیف انرژی رادیوایزوتوپ های گسیلنده پرتو گاما جهت استخراج برهمکنش های غالب ۱۲
علی فیاضی، حسین احمدی نوبری

کاهش زمان نشست در کنترل فیدبک تاخیری سیستم آشوبی روسلر با سنکرون کننده PI ۲۳
حسن فاتحی مرج، رجب اصغریان، ناصر پریز

بکارگیری روش میانگین گیری مرتب وزندار (OWA) در ترکیب داده های ربات مین یاب ۲۹
محمدرضا بادلو، بهزاد مشیری، بابک نجاراعرابی

ارائه یک سیستم آشوبناک بعد بالای جدید همراه با یک نقطه تعادل و پایدارسازی آن با استفاده از کنترل کننده فیدبک حالت خطی ۳۷
علی ابویی، محمدرضا جاهد مطلق، زهرا رحمانی چراتی

English Part

A Combined DC-Filter and Optimized Modulation to Absorb DC-Link Oscillations of Cascaded H-Bridge Converters 1
Mohammad Tavakoli Bina, Bahman Eskandari

Utilizing Multi-Parametric Programming for Optimal Control of Linear Constrained Systems 12
Farhad Bayat, Ali A. Jalali, Mohammad R. Jahed Motlagh

www.isice.ir